THE FORM AND ACQUISITION OF FREE RELATIVES

A Dissertation Presented

by

MICHAEL CLAUSS

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2017

Linguistics

THE FORM AND ACQUISITION OF FREE RELATIVES

A Dissertation Presented

by

MICHAEL CLAUSS

Approved as to style and content by:

_____

Thomas Roeper, Co-chair

_____

Jeremy Hartman, Co-chair

_____

Rajesh Bhatt, Member

_____

Jill de Villiers, Member

_____

Seth Cable, Head of Department
Linguistics

DEDICATION

For my ʻohana, and for the workers of the world.

ACKNOWLEDGMENTS

Writing this dissertation was the unavoidable outcome of a series of questionable life choices and a lot of good luck. Since this is an acknowledgments section and not a memoir, I'll stick to discussing the good luck, which generally took the form of wonderful people I've been fortunate enough to have in my life.

My plan to work on acquisition in my PhD became solid some time between when I applied to UMass and when I started here. As such, when I was applying I didn't know much about Tom Roeper; but, during the first department Town Meeting I attended, in my first week at UMass, right after I said who I was and what I was interested in, Tom took me aside and suggested we talk some time. Since then he's been an indispensable font of ideas which I have worked with and which eventually became the major substance of this dissertation. Tom's advising has done a huge part in shaping what sort of linguist I am, and TAing for him taught me a lot about how to teach, especially how to teach experimental methods.

Jeremy Hartman came into the department a year and change after I did, so I always felt (perhaps unreasonably) that he was almost as much a peer as an advisor. Working with him on various projects while I've been here has been a great joy, and I hope to continue working on the projects we've started here long after I've gone some place else.

The first person I knew anything about, and the first person I talked to at UMass linguistics was Rajesh Bhatt; it surprises me looking back that this dissertation is the first project I've actually worked with him on. But, it was worth the wait - Rajesh's advice and training in understanding and writing about syntax has been crucial to whatever amount of success I've accomplished here. Beyond all this as well, Rajesh is my role

model for how to be an academic and live a good and fun life. Rajesh's sunny attitude and frankly surprising coolness is something for me to aspire to.

From the very beginning of my time at UMass, Jill de Villiers's perspective, and insistence on methodological rigor, greatly helped me in developing as an acquisitionist. Having her at LARC meetings and at workshops was always as delightful as it was necessary.

My other professors here at UMass also have played an invaluable role in molding me into what I am (for better and worse), particularly Kyle Johnson, Ellen Woolford, Magda Oiry, Luiz Amaral, Joe Pater, Ilaria Frana, Kristine Yu, Lyn Frazier, Brian Dillon, Seth Cable, Alice Harris, Erik Cheries, and Andrew Cohen.

My life outside of the department would have been miserable - and may well not have existed, for all practical purposes - without my dear friend Megan Somerday. Megan is my very coolest friend, a kindred spirit of a kind I've met very few times in my life, someone I went on poorly-planned adventures with, wasted countless hours watching YouTube with, shared art and music with, and most importantly someone who helped me build a social life in this town that felt a bit cold before she got here. Megan has done me only one unkindness over the years, and that was simply to leave Massachusetts before I did.

Nick LaCara and Jyoti Iyer also played a huge role in my survival here in Northampton and within the linguistics department. Nick, seeing shows with you, talking about and playing music with you, and basically just nerding around comparing Star Trek series and shouting Futurama lines at each other have been a huge part of my life here. Jyoti, you're my favorite neighbor, one of my favorite singing-partners, and the best of smoking buddies.

My cohort-mates Shayne Sloggett, Stefan Keine, Jérémy Pasquereau, and Hannah Greene – we were not always a cohort in more than name, but my first year would have gone all the worse without a group of bright-eyed geniuses like each of you. Jon Ander

Mendia, Sakshi Bhatia, Ivy Hauser, Yangsook Park, Alex Nazarov, Tracy Conner, Brian Smith, Ethan Poole, Hsin-Lun Huang, Andrew Weir, Elizabeth Bogal-Allbritten, Presley Pizzo, Thuy Bui, Rodica Ivan, and Georgia Simon have all been major parts of my life here at UMass who helped make the department a livable space. In addition to all of them, I've had the pleasure of meeting, however briefly, a number of wonderful visitors to UMass linguistics: Fernanda Mendes, Gustavo Freire, Laura Aldridge, Margreet van Koort, Becky Woods, Marlijn Meijer, Yu'an Yang, and Athulya Aravind (Athulya we'll count you as a visitor, just a very regular one).

I also must mention the other assorted strange characters I've interacted with in Northampton outside of the department: Ghazah Abbasi, Theresa Copeland, Jordan Kroll, Sonny Nordmarken, Avanti Mukherjee, Chris Bigelow, Monk Kelley, Fiona Riewer, Tiphaine Duprat, and Alexandra Levy. You've all been a huge and important part of my life, and of the funny little town we live in.

Not enough could ever be said to thank the staffs at The Roost, La Veracruzana, Fitzwilly's, Provisions, and most especially Ye Ol' Watering Hole and Beer Can Museum. "Invaluable" wouldn't begin to cover it.

In addition to all of the faculty here at UMass, the people who worked with me through my Masters at UH deserve a lot of credit; Kamil Deen in particular made me interested in and committed to acquisition work, and helped me better understand the nativist program, and he also remains one of my favorite people to see at conferences. William O'Grady and Yuko Otsuka were also a huge part of shaping my early work in linguistics - and, it's definitely worth saying, William's teaching did a lot to convince me of the importance of linguistic nativism (though maybe he would not enjoy hearing that).

My first linguist friends, at UH, were and are a huge part of my life, both academically and personally. Most especially, Joelle Kirtley is someone I love dearly, and my connection with hear in and after those days still has a very special place in my heart; leaving UH was hardest because I was further from her. My days at UH were also brightened by

my friendships with James Grama, Hanbing Feng, Laura Berbusse, Melody Ross, Mary Walworth, Jessica Seid, Nala Huiying Lee, Hyunah Ahn, John Van Way, Aaron Waldrip, Jinsun Choe, Elaine Lau, Tobias Bloyd, Amanda Hamilton, and Mayumi Oiwa; as well as my extended ʻohana in Hawaiʻi: Sai Bhatawadekar, Kim Schauman, Jay Purvis, Aashish Hemrajani, Emily Krause, Mike Lee, Hannah Rosenberg-Jones, Wayne Tanaka, and Letani Peltier; mahalo for all of your kōkua.

My Seattle family also need mentioning: John Appiah Duffell, my oldest, best, and strangest friend, the Mario to my Luigi and so much more; Tara and Ian Adams, Mac and Matinay Justice, Kurt Litsinger, and Katy Anderson, are all wonderful reminders to me every time I see them that my home town can still be home even from so far away.

Finally and most importantly, my family is, more than anyone else, to blame for the sort of person I am. My father is the model of academia on which I built myself and, eventually, my career in linguistics. My mother provides me constant inspiration for how to live a good life of work/life balance and constant singing. My sister, aside from being by far the coolest person I've ever met, is living the artist's dream that I always wanted to. And my brother has given me all of the practice in arguing that I've needed to be an effective academic.

ABSTRACT

THE FORM AND ACQUISITION OF FREE RELATIVES

SEPTEMBER 2017

MICHAEL CLAUSS

B.A., WESTERN WASHINGTON UNIVERSITY

M.A.., UNIVERSITY OF HAWAI'I AT MĀNOA

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Thomas Roeper and Professor Jeremy Hartman

This dissertation examines the syntax of Free Relatives (FRs) in English at different stages of first language acquisition. The goal is to provide a theory of Free Relatives that explains phenomena in adult and child FRs, is feasibly learnable by a child, and reflects principles expressed in theories of Universal Grammar based on the Minimalist Program (Chomsky 1993, 1995, 2005).

The central empirical concern is the difference between the distribution of Wh expressions in FRs vs. Wh questions in English, the difference in grammaticality between *Charles wondered dish what Sebastian made* and **Charles ate what dish Sebastian made* (*Wh-NP). To explain this and related phenomena I develop an analysis of FRs building on recent work by Caponigro (2003), Donati (2006), and Cecchetto and Donati (2011, 2015); I propose that *Wh-NP is a consequence of FRs being derived from question-like Wh clauses by Head Movement of the Wh word to a higher D head. This movement is motivated by a requirement that definite D always be overtly realized; *Wh-NP arises

because Left Branch movement is barred in English. I describe this in terms of a feature [±Overt].

Further, I show that while the basic form of FRs is acquired early by children (Flynn and Lust 1980, Guasti and Shlonsky 1995), experimental given here data shows that the syntax deriving *Wh-NP is developed considerably later. I explain this in terms of development of the functional structure of definite descriptions, in particular requirements on overt Determiners. I compare this with the development of other types of nominals in English. I then discuss recent insights on the nature of Parameter Setting by Biberauer and Roberts (2012) and describe ways in which a simple Acquisition Device can deal with learning FRs by generalization of features of the nominal domain.

Finally, I discuss the issue of Wh-ever FRs, which do not show *Wh-NP. I give experimental evidence from adults and children, as well as naturalistic data, that this is because Wh-ever FRs are not definite descriptions, and so do not require the movement of the Wh word to a higher D head.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

CHAPTER 1

ACQUISITION PROBLEMS IN MINIMALISM

Knowledge of language is, in general, knowledge of form-meaning mappings. In the realm of syntax, this consists minimally of two elements: the set of strings allowable in a given language, and the allowable meanings for any given string. The latter can be shown to have a subtype, which makes specific reference to the allowable strings in a language.

The first is type of knowledge is simple to demonstrate: a speaker knows the possible surface strings of their language independently of their meanings; or, put another way, a speaker knows the set of forms which are allowed to be mapped onto some meaning describing the world, as distinct from knowledge of the real world to which sentences might refer. This was most clearly illustrated by Chomsky's (1957) example comparing a well-formed but non-sensical sentence to an ill-formed one, as in (1).

(1) a. Colorless green ideas sleep furiously

b. *Sleep ideas colorless furiously green

The first sentence must be included among the set of sentences in the speaker's language, as it is on the surface grammatical, and it could be mapped onto a meaning, but the resulting meaning does not seem to describe anything in the real world. The second sentence cannot mean anything, as it is ill-formed in the first place.

The second type of knowledge involves the disambiguation of different types of sentences. This takes two forms. One is where certain sentences are globally ambiguous between two meanings, while a rule of interpretation disambiguates another, similar

sentence. This is shown in (2), where the sentence in (a) is ambiguous between a reading where every professor impresses her own students, and another where there is some woman whose students were impressed by every professor. The sentence in (b), however, unambiguously means that there is one student's professor who impressed every student, not that every student was impressed by her own professor.

(2) a. Every professor impressed her students

b. Her professor impressed every student

Both of these sentences are grammatical on the surface, but a general rule regarding the interpretation of pronouns and quantifiers gives rise to a variable ambiguity - the type of sentence in (2a) is ambiguous, and the type in (b) is not.

There is another way in which we see knowledge of language appearing as knowledge of disambiguation, where the ambiguity of surface grammatical strings relies on whether or not similar sentences are or aren't grammatical. As an example of this, take simple attachment ambiguities: (3a) allows for two readings, one where the student possesses a set of binoculars, and one where Sebastian is using binoculars. The apparent paraphrase in (3b), where the PP *with binoculars* is topicalized, only allows the latter meaning.

This is parallel to the similar pair of sentences in (4): the sentence in (a) is unambiguous, only allowing the "low" reading where the painting is by Charles; the same syntactic alternation as in (3) applied here gives rise not to disambiguation, but surface ungrammaticality.

(3) a. Sebastian saw the student with binoculars

b. With binoculars, Sebastian saw the student

(4) a. Sebastian liked the painting by Charles

b. *By Charles, Sebastian liked the painting

The topicalization in (4b) violates the constraint against movement out of a complex nominal in English (Ross 1967), and thus the sentence is ungrammatical. The low reading of the PP in (3a) cannot be preserved in (3b), as this would involve the same sort of ungrammatical movement. We can illustrate this in terms of the location in a structure of a trace relative to an NP, as in (5).

(5) Complex NP islands

a. *[By Charles]$_i$ Sebastian liked [the painting t_i]

b. *[With binoculars]$_i$ Sebastian saw [the student t_i]

So, it seems that in fact rather than being two different phenomena, (3-4) are two instances of the same phenomenon, where a particular structure is blocked and thus a particular meaning is blocked; it just happens to be the case that the string in (3a) has a second structure which could be associated with it, generating another meaning.

Phenomena such as these tell us something about the depth and reality of speakers' knowledge of the formal properties of language: these phenomena are only possible if speakers make use of their syntactic knowledge both in the sentences they produce and accept, and the set of possible interpretations they give to acceptable sentences. Further, this tells us something about the nature of the types of linguistic knowledge exhibited in (1-2), which is that they may not be different types of knowledge at all, but rather two manifestations of knowledge of linguistic **structures**: certain structures are available and others aren't, and any given structure derives a particular meaning, which may or may not be pronounced the same as other structures with different meanings (and which may or may not have a meaning that "makes sense" in the real world).

This gives rise to an interesting question about the acquisition of language: What are the cases where a child learning a language determines the possible meanings of potentially-ambiguous sentences based on what they learn about their language's syntax, and what are the cases where they deduce properties of their language's syntax

based on associating strings with meanings-in-context. Take, for example, the examples above: previous research suggests that the sort of disambiguation in (3) is learned early by children (Otsu 1981). We could ask, do children learn that (4b) is ungrammatical from the difference in observed meanings of (3a-b) , or do they come to know that (3b) is unambiguous because they already know that (4b) is ungrammatical? Are children inducing island constraints from the grammar they observe, or are they applying island constraints which they already know (perhaps innately) to a particular type of sentence to constrain meaning?

Here I will discuss questions of this sort involving a case of ambiguous sub-sentential strings and the ways in which they are disambiguated, namely Free Relative clauses (FRs) and embedded Wh questions in English (6).

(6) a. **Free Relative**

Sebastian stole [what Charles read]

b. **Wh question**

Sebastian wondered [what Charles read]

I will be concerned with the ways in which these two differ in their form and interpretation, and how children come to develop a grammar which generates these differences. In particular, I will investigate what sort of learning mechanism can allow a child to converge on an adultlike grammar of the type assumed in the Minimalist Program (Chomsky 1993, 1995). The rest of this chapter will concern the relationship between Minimalist syntax and problems of acquisition.

For the sake of clarity, the following two definitions will be used to describe specific acquisition problems and linguistic knowledge within Minimalism:

(7) An **acquisition problem** is any grammatical property which must be learned for an individual language to achieve knowledge of the target language.

(8) A **Minimalist system** refers here to knowledge of the set of linguistic properties assumed to be innate given Minimalist assumptions about grammar.

The purpose of this chapter is to describe ways in which a Minimalist system provides possible solutions for individual acquisition problems that allow learners to identify their target grammar.

The first section describes the goals of Minimalism as a theory of syntax, and characterizes the broad nature of acquisition problems in Minimalism. The second section discusses the issue of parameters in Minimalism and other theories of linguistic knowledge as a way of describing variation between grammars. The third section describes specific ways in which a Minimalist system (especially properties of syntactic derivations in Minimalism) can provide solutions to some basic acquisition problems. The fourth section describes one particular acquisition problem which will be the topic of the following chapters: the issue of Wh Movement, and properties which unite and distinguish constructions of this type, particularly in English.

1.1 Goals of Minimalism, and Too-Many-Analyses

Early work in the Principles and Parameters framework (Chomsky 1981), as well as earlier work within generative frameworks (Chomsky 1965) focused on UG as a solution to a learning problem, with the former work emphasizing the role of Parameter setting as a goal of a learner (see papers in Roeper and Williams 1987). Later work under the aegis of the Minimalist program de-emphasizes the role that UG plays in restricting possible grammars, and shifts this burden onto interfaces (Chomsky 2005:8-9). An early attempt at escaping the importance of “Principles and Parameters" per se - the way a certain set of explicit Principles govern representations, and a set of Parameters define variation - can be seen in the quote below.

> To yield the correct results, the “least effort" condition must be interpreted so that UG principles are applied wherever possible, with language-particular rules used only to “save" a D-structure representation yielding no

> output .... UG principles are thus “less costly" than language-specific principles. We may think of them, intuitively, as “wired-in" and distinguished from the acquired elements of language, which bear a greater cost. (Chomsky 1995:140).

That is, roughly, the linguistic experience of the learner requires that they hypothesize a grammar which can generate the natural utterances which they here (and, others which bear some similarity to the ones they’ve heard), and UG restricts the sort of representations which are admissible to accommodate the experienced language. Under this view, the acquisition process consists of experiencing utterances and finding the best way to represent all of them, within a single grammar, given the native limitations on possible grammars.

The second element of this viewpoint, which is emphasized more in later Minimalism (Chomsky 2005) and more explicitly defines what the important factors outside of UG, is that rather than UG principles doing the greater work restricting possible grammars, necessities of interface systems restrict most possible grammars as being ruled out on more practical grounds. This keeps the intuition from the above quote that certain innate properties of the mind make certain grammars more “costly" than others, but gives flexibility to what could be a possible grammar allowed by UG in the first place.

However, acquisition questions necessarily still come up in the Minimalist program, and the role of UG doesn’t go away completely. The existence of an innate linguistic module necessarily means that there is a single particular way in which structures are built, and thus a way in which UG restricts an acquisition path - a way in which UG provides a limited set of solutions to the problems which a learner encounters . In this section I discuss what constitutes such a question and how it might be solved in Minimalism.

1.1.1 Acquisition problems in Minimalism systems

While later Minimalist work backs away from questions of acquisition to some extent, focusing rather on why a Minimalist system is ideal from an evolutionary point of view (Chomsky 2013; Hauser, Chomsky, and Fitch 2002), the perspective discussed in Chomsky (2005) does give rise to a specific acquisition program. The “design" of the broad faculty of language from this perspective is outlined in (9).

(9) **Three Factors of Language Design**

1 The narrow faculty of language (UG, taken to consist of binary Merge);

2 Experience of the learner (The linguistic input);

3 Other properties of the mind and language use (“Third-Factor" properties).

To some extent this is simply a restatement of any theory of Universal Grammar: specific languages are learned through exposure to them, and language necessarily interacts with other cognitive systems (it must be perceptible by some medium, speakers must be able to actively assign representations of sentences as they hear them, and expresses concepts which are somehow familiar to speakers), so in fact the second and third factors are necessary for any theory of linguistic knowledge. Any theory of UG then is simply a definition of the first factor.

The Strong Minimalist Hypothesis of Chomsky (2005) is that UG - the cognitive process unique to linguistic knowledge - consists solely of the syntactic operation Merge, which combines a pair of syntactic objects to create a new syntactic object (and which may apply recursively). However, even if we wish to be strictly Minimal about our theory of syntactic knowledge and say that Merge is the only primitive of the language faculty, it is insufficient to say that it simply consists of binary Merge - properties of Merge must be described.

At least two properties of the syntax seem to be necessary addenda to this definition of Merge: first, Merge applies to syntactic objects with structure. That is, while the

first two objects Merged in a derivation (say, two words) may have no internal structure, the output does have a structure, and any further instances of Merge will maintain this structure. That is, if an object {a,b} is merged with an object {c}, the output is the structured {c,{a,b}}, not the unstructured {c,a,b}. This is more or less trivial, but it is important to note because (a) it makes **structural** linguistic knowledge a part of the Language Acquisition Device, and (b) it allows for a distinction between a simple object X^0 and a complex object XP which contains it (that is, a Head vs. a Phrase), a distinction which we may refer to in the formation of different rules, either in the sense of patterns specific to a language or of more general constraints on the application of Merge. Further discussion here will make important use of this distinction, particularly where it concerns constraints on movement.

A second property which must be defined for Merge, and is more difficult to pin down (but which also may make reference to the Head-Phrase distinction), is a theory of Labeling: how is the syntactic category of the output of Merge determined? That is, when XP and YP are merged, is the output an XP, a YP, or other? What is the "name" of the output of any given instance of Merge? While there is ongoing theoretical debate on the exact formulation of a Labeling algorithm (Collins 2002, Chomsky 2013, Cecchetto and Donati 2015, Chapter 2 of this dissertation), most analyses assume at least that the label of the output of any instance of Merge is the label of one of the items being merged. That is, if XP and YP are merged, the label of the ouput cannot be ZP. Along with the Head-Phrase distinction, this aspect of Merge will inform much of the rest of the discussion here.

We could also posit that other properties must be attached to Merge - for example, constraints on when it can and must apply (a theory of Agree), or when a sequence of Merges (ie., a derivation) can or must end (a theory of Spell-Out); however, most of the discussion here will concern these first two - what the ways are in which innate knowledge of the head-phrase distinction and a necessary theory of labels can guide the ac-

quisition process, and how much of the acquisition of syntax (and properties of specific constructions) can be described simply in terms of this knowledge.

1.1.2 The Too-Many-Analyses problem

Since within Minimalism, UG is assumed to be simple and unitary, and properties of mind generally are going to be subject to either individual variation or no variation at all rather than varying across speech communities, any feature of a particular language in the common sense (some pattern specific to American English, Malawian Chichewa, etc.) is necessarily due to features acquired by a given speaker from caretakers and companions during the acquisition period – the Input. So, learning involves structuring observed utterances into a grammar permissible by UG (and optimal in terms of interface conditions).

A grammar as viewed by Minimalism is, derivationally, a very restrictive system: There are very few operations (perhaps only one), and any operation must be motivated by the presence of formal features. But, it is still representationally quite free. Two surface-similar patterns seen in two different languages, which have essentially the same meaning and recognizably similar linear orders, may have vastly different syntactic representations, each of which is derived by step-by-step binary Merge. We could call this the “Too-Many-Analyses" problem; Chomsky (2005:8-9) alludes to it specifically, saying that an “anything goes" output of UG is the necessary correlate of the Strong Minimalist Hypothesis (and thus the importance of interfaces in restricting grammars).

To illustrate this, consider two examples: Subject-Verb-Object order may be derived quite differently in different languages. In English, it is essentially underived; the subject is generated pre-verbally, and moves string-vacuously away from the verb. In German, the verb originates to the right of the object, moves into the left periphery, and the subject moves in front of it (10). Similarly, externally headed post-nominal Relative Clauses may be generated either by movement by the head noun out of the clause or movement

of some non-overt material within the clause, such as a “matching" nominal which goes unpronounced (11).

(10) Analyses of *John ate an apple*

a English: [$_{CP}$ ... [$_{TP}$ John$_i$ [$_{vP}$ t$_i$ ... ate an apple]]]

b German: [$_{CP}$ John$_i$ ate$_v$ ... [$_{vP}$ t$_i$ ... an apple t$_v$] ...]

(11) Analyses of *the boy John knows*

a Movement analysis: [$_{DP}$ the [$_{CP}$ boy [John knows t$_{NP}$]]]

b Deletion analysis: [$_{DP}$ the [$_{NP}$ boy] [$_{CP}$ ~~boy~~ [John knows t]]]

All of these derivations involve recursive application of Merge, but in different orders and applying to different lexical items and constituents, and so all of them produce representations which are plausible Minimalist analyses. As such, a child acquiring English may potentially entertain any or all of these during the acquisition period, and this would be evidenced by their performance on the sort of phenomena which linguists use to distinguish them in a given language. For example, a child learning German would posit a V2 representation based on things like object topicalization or the form of embedded clauses. Ultimately, these sorts of input properties must guide the child to a grammar which yields a narrower set of representations and, thus, form-meaning mappings.

Complicating things further is the fact that a child may need to learn a grammar which does not always choose between the potential representations: English may allow, in some cases, for both the English syntax given in (10) as well as the “German", V2 type (Vikner 1995). Further, English does seem to need both analyses of headed RCs to obtain the full range of allowable sentences and denotations (Bhatt 2002, Hulsey and Sauerland 2006). This can have different sorts of results for learners: grammars which

allow multiple possible representations, speech communities with non-uniform representations of similar sentences, and protracted acquisition periods as the “right" representation is slowly picked out.

Instances of each of these have been observed: Hulsey and Sauerland (2006) give evidence that both of the structures of RCs in (11), must be available, and in most cases there is no need to choose between the two; Han, Lidz, and Musolino (2007) argue that the syntax of verb movement in Korean is underdetermined by the input and that this results in adult speakers of Korean not consistently sharing the same syntactic representations; and, the language acquisition literature shows many examples of grammatical patterns which are not entirely clear from the input showing delayed acquisition. Below I discuss some of these in greater detail.

1.1.3 Too-Many-Analyses in acquisition

The issue of “Too-Many-Analyses" - the fact that a Minimalist System allows a number of possible representations for surface patterns - is a topic, explicitly or not, in many generative studies of language acquisition. Here I will mention a few.

The issue of the multiple possibilities for Relative Clauses available within and across different languages is taken to be an explanation for certain non-adultlike behavior by Labelle (1990, 1996). In particular, Labelle points out that children acquiring French produce Resumptive Pronouns which are illicit in adult French, as in (12) from Labelle (1996).

(12) *Celle-là que le papa* ***lui*** *montre un dessin*
that one that the father **to her** shows a drawing
‘The one that father shows (her) a drawing’

Labelle argues that this is evidence for French children giving an analysis of RCs which lacks movement of the nominal head of any kind (a third option in addition to the head movement and head matching analyses in (11) above). Labelle points out

(1996:71fn) that this is a pattern attested in some non-standard varieties of adult French (as well as being attested cross-linguistically), and so it must be considered that children producing utterances like (12) are making use of a grammar which is generable by a Minimalist System, but which is not the target grammar they are acquiring.

Belletti (2009, Ch. 9) gives an overview of how acquisition data and cross-linguistic comparative data can work together to explain certain correlations within and across languages. Without necessarily adopting all of the assumptions Belletti makes about acquisition and the form of the grammar, we can understand this within the context of the Too-Many-Analyses problem discussed here.

Two case studies discussed by Belletti which are of interest to the current discussion are the case of acquisition of post-verbal subject positions in Italian and the acquisition of clitic placement in French. In the former case, Belletti discusses data from Belletti and Leonini (2004) and Belletti, Bennati, and Sorace (2007) looking at second language learners of Italian. They look at two aspects of Italian which are not found in the second language learners' L1: null subjects and post-verbal subjects where the subject is focused. They find that while the learners produce null subjects where expected fairly reliably, they are very unreliable in restricting post-verbal subjects to just subject focus cases.

Note that it's not the case that these studies didn't observe VS patterns, only that they didn't correspond to a particular semantic/pragmatic condition, namely subject focus. Based on this, Belletti (2009) argues for an analysis of Italian where VS is available because (a) there is no requirement in general that the canonical subject position be filled (akin to the *pro*-drop parameter or an EPP feature), and (b) there is an available low Focus projection above *v*P; this allows for the subject to move to the Focus position and not all the way to the specifier of IP, just in case there is an available FocP position, as in the syntax in (13a). Since the L2 speakers are not restricting VS to subject focus

conditions, Belletti argues for an L2 syntax like (13b), where the subject simply remains within vP; in both cases, the verb simply moves past it to create a VS order.

(13) Analyses of VS (*Gianni called*)

a. L1 Italian

[$_{IP}$ called [$_{FocP}$ Gianni [$_{vP}$ t_G t_v]]]

b. L2 Italian

[$_{IP}$ called [$_{vP}$ Gianni t_v]]

What generates this distinction is variation in terms of selectional factors: the L1 Italian representation involves an IP which embeds a FocP, whereas the L2 representation involves embedding of only a v_P. What's crucial about this is the difference between L1 and L2 Italian is that the availability of null subjects and of the VS linear order is wholly transparent from the overt syntax, while the target structure of VS sentences, with its extra FocP projection, is not immediately visible; thus, while learning the availability of null subjects is trivial, learning the particular syntax for Italian VS is much more difficult, resulting in a distinction between what is learned easily by the L2 speakers and what is not. This is then a case where the number of possible representations of a string is increased by the number of possible representations of it which may be generated. It is also a case where what varies between the two possible representations is the availability of a particular functional head - see the following section for a broad discussion of this issue and its importance to a Minimalist theory of acquisition.

The second case study discussed in Belletti (2009), that of clitic placement in L1 vs. L2 French (based on data from Hamann and Belletti 2004), finds that while L1 speakers acquiring French do not experience problems with the acquisition of either (weak) pronouns or clitic placement, for L2 learners of French clitics are allowed to vary between pre-verbal and post-verbal positions. Assuming that cliticization involves movement

of a Head onto the verb, Belletti (2009) analyses these errors by L2 learners as a misanalysis of clitics as phrasal pronouns. The logic is roughly as follows: clitics must be morphologically attached to a verb, and so may only occur in the pre-verbal position, while independent pronouns may occur in the canonical object position after the verb, or be moved above the verb. The relevant representations are sketched in (14): for L1 French speakers, the clitic moves as a head onto the verb, while for the L2 speakers the pronoun moves as a phrase, essentially scrambling past the verb.

(14) Analyses of preverbal pronoun in *le-laisse* ('allow it')

a. L1 French

$[_{IP}\ [_{V}$ le-laisse$]\ [_{vP}\ t_{v}\ [_{DP}\ t_{D}\ ]\ ]]$

b. L2 French

$[_{IP}\ [_{DP}$ le $]$ laisse $[_{vP}\ t_{v}$ t$_{DP}\ ]]$

This is another case where surface strings allow multiple possible relations, in this case varying between whether a syntactic object (*le*) is analyzed as a Head (a clitic) or as a Phrase (a phrasal pronoun). The consequence of this choice is whether it moves preverbally necessarily (for a clitic) or optionally (for a pronoun). And again, as with the above cases, we see this is a pattern where the acquisition of the target grammar can be difficult.

These three cases are instances where surface strings mask distinctions in whether things move, where things move, and whether syntactic objects are complex (Phrases) or simplex (Heads), decisions which have consequences of various sorts relating to available variation of form and interpretation within the language. Throughout the discussion here we will see other instances where exactly these choices, and their opacity in surface strings, leads to difficulties in acquisition paths, and where the choice made has consequences for particular grammars.

These are also cases where Minimalist systems allow for subtle variation which must be determined for any learner; this thus requires a theory of how a learner makes these sorts of choices when acquiring a target grammar, and what the properties of the Language Acquisition Device are that allow this variation. The next section discusses theories of variation in terms of Parameters and Parameter settings as a solution for this.

1.2 Parameters and variation across grammars

Under the view of discussed in work like Chomsky (2005) linguistic knowledge, it is typically taken to be the case that the First Factor is simply Recursive Binary Merge, however that may be defined (the Strong Minimalist Hypothesis). However, evidence of any amount of variation between languages (never mind the considerable amount of variation which is actually observed) necessitates that the Second Factor, linguistic experience, has some effect on linguistic knowledge. For any of the patterns discussed in the previous section (object clitics, constituent order, relativization), a child learning a language must arrive at a specific grammar allowing for only some possible structures, based on particular patterns they observe in the input. Thus, a theory of acquisition must be linked to a theory of variation. A broad way to describe this under a related set of frameworks is as Parameters and Parameter-setting. Here I will discuss a theory of Parameters which restricts variation (as much as possible) to properties of the Lexicon (Borer 1984, Chomsky 1995).

1.2.1 The lexical view of parameters

Under traditional Principles and Parameters views, in addition to universal properties of and constraints on syntax, the First Factor also must consist of a limited set of Parameters, a list of binary options allowing some particular set of structures, each of which learner must choose a setting for. The inclusion of such as list is generally undesirable if one is trying to make innate syntactic knowledge strictly Minimal.

The nature of cross-linguistic variation and the identification of specific target grammars from among the possibilities offered by UG over the course of acquisition, typically referred to as Parameters and Parameter-Setting, can be characterized in a number of ways. One way is what may be called a “Menu Options", or “Switch Board" theory (Chomsky 1981), where there is a set of binary choices that are be made by a learner, with different permutations giving rise to different observable grammars; a learner may choose to set a “switch" in their mental representation which makes their language Verb-Final, or Verb-Second, or Poly-synthetic, etc.. Another way to describe Parameters and variation is as a sort of “toolbox," where UG offers a number of grammatical functions which may or may not be used in any given grammar, with the learner picking them out from UG to parse sentences in their input.

The view taken by most generative approaches to the acquisition of syntax is that the primary goal of acquiring a specific language is the setting of Parameters (Borer and Wexler 1987, Rizzi 2005, Yang 2003, Snyder 2007). Views on the exact nature of Parameters and Parameter Setting vary somewhat across different approaches and theories: there are open questions about whether Parameters apply to the narrow syntax vs. some other part of linguistic knowledge, whether there are only a small number of “Macroparameters" or a larger number of “Microparameters", and whether Parameters are set early and firmly (Borer and Wexler 1998), or gradually (Yang 2003, Snyder 2007), are all open questions. While in general, Minimalism does not make necessary predictions for the latter two questions, the Strong Minimalist Hypothesis requires that Parameters do not apply to the narrow syntax, as this consists purely of binary Merge.

So, in (strong) Minimalism, Parameters must be put somewhere besides the narrow syntax. One way to do this would be to posit that Parameters apply to the interfaces rather than the narrow syntax. For example, left- or right-headedness is a property of a linearization algorithm which applies to Phonological Form rather than in the narrow syntax, and interpretation of the scope of quantifiers is a property of semantic compo-

sition and so it applies to Logical Form. Another option is to put Parameters within the lexicon - which is, in fact, the only piece of knowledge of any given language which conceptually **must** be learned entirely and not given “for free" by any internal system.

I will take the view, originating in Borer (1984) (and endorsed from a Minimalist standpoint in Chomsky 1995), that the primary locus of parametric variation is in the lexicon, rather than the narrow syntax or the canonical “Interfaces" (The phonological and semantic interfaces). I will also take the view that, particularly given this assumption, there must be a larger number of micro-parameters, given the amount of variation observed within and across languages in the form of different (but related) constructions.

Some amount of expounding on the nature and consequences of this view will be necessary as background for the pursuing discussion. First, I assume that parametric variation largely consists of the distribution of certain formal features across the lexicon, such as Φ features associated with morphology of case and agreement (of the sort discussed in Borer 1984) or features which require movement into certain positions (for example, the feature which requires that there be a constituent moved to subject position). This will give rise to the Micro-parametric nature of variation.

Secondly (and relatedly), I will assume that while lexical items contain a kind of hierarchical structure, it is not of the same sort as that of syntactic structure (see discussion in Harley and Ritter 2002 and Béjar and Rezac 2009 for formalizations of this idea). The structure is thus is not necessarily composed in the syntax (contra the assumptions of Distributed Morphology; Marantz and Halle 1993) - though it may sometimes be.

As the distribution of formal features is a property of the lexicon, and thus necessarily subject to a lot of potential idiosyncrasy, we predict to observe exceptionality within languages in the morpho-syntactic domain of a similar sort to those we see in the morpho-phonological domain, and that both might be derived by the extent to which particular lexical items are specified for their formal representation and the extent to which they

default to more general patterns. That is, morpho-phonological patterns of the kind seen in (15) will have some parallels in exceptional morpho-syntactic patterns.

(15) a Languages vary in respect to whether stress is in general lexically assigned (Japanese, Russian) or entirely predictable (Finnish, Hawaiian), and some languages may have a “default" stress pattern which may have lexical exceptions (English).

b English has default forms for morphemes like PLURAL (/-z/) and PAST (/-d/), but allows for specific lexical exceptions (*feet*, *caught*).

An advantage of the lexico-parametric view is that the difference between this sort of exceptionality/generalization and generalization within specific realms within syntax (say, left- vs. right-headedness in nominal vs. verbal domains) is that one can make reference to parametric settings within the **substantive** lexicon - what one might call the lexicon *per se*, the lexicon in the traditional sense - and the **functional** lexicon - the set of functional heads in a language (one might also refer to this in terms of closed- vs. open-classes of lexical items). The things we conventionally would call Parameter settings apply to items in the functional lexicon, and thus have systematic correlations with syntactic patterns (of various levels of generality), and the things we would conventionally call exceptions apply to (sets of) substantive lexical items, and thus have very limited correlations to syntactic patterns. The discussion below will explicate properties of functional lexicons in a way which restricts possible grammars and yet allows for parametric variation.

1.2.2 The functional lexicon

One way in which the distinction between the functional and substantive lexicons provides useful tools for syntactic acquisition can be seen in the fact that the substantive lexicons of different languages seem to vary a great deal more than the functional lexicons; while not every language necessarily makes use of every possible functional

lexical item, languages vary much more dramatically in their substantive lexicons, with certain words appearing to be “untranslatable", in the sense that there is not a simple, single-word translation of one word into another language.[1]

In discussing the restricted nature of the functional lexicon, Cinque (2013) gives evidence for this from specific examples of plausible morphemes expressing mental states which do not seem to occur in natural languages, as shown by some of his examples given in (16).

(16) Non-existent morphemes (Cinque 2013:51)

a. He fight-*af* (= he is afraid of fighting)

b. I say-*am* you are wrong (= I am sympathetic in saying you are wrong)

Cinque points out that non-existent verbal fear morpheme -*af* in (16a) would be parallel to verbal morphemes expressing desire in some natural languages. While both desire and fear may reasonably be thought to be innate parts of human cognition, only the former seems to appear in human verbal systems.

It is not necessarily entailed, of course, that native Concepts be meaningfully identical to the native Functional Heads which introduce them as predicates in a derivation. For example, the cognitively available concept of Theory of Mind, or the representation of other belief states than that of the speaker, may or may not be distinct in substance from the denotation of an epistemic predicate which quantifies over possible worlds. Similarly, the concept of uniqueness, which we may well want to assume is native, does not have to bear particular resemblance to definite descriptions, but we probably also want a natively available functional ι predicate. I give examples of these sorts of dis-

[1]This sort of idea is not common in the literature on linguistics, outside of pop-linguistic articles with a somewhat Whorfian perspective on language variation; these tend to either correlate to some cultural convention, or to be derived from compounding or synthetic morphology, or to be some non-compositional idiom; a representative list could be seen as the list here:
https://www.rocketlanguages.com/blog/20-of-the-worlds-most-beautiful-untranslatable-words/

tinctions in (17); (a-b) give typical formalizations of these types paired with informal descriptions of the concepts they are associated with, while (c) gives a broad description of the concept of sympathy with a formalism for what the non-existant Generalized Sympathy predicate would have to look like.

(17) Concepts and Predicates

a. Theory of Mind

Concept: One's own set of beliefs and assumptions are not necessarily shared by others.

Predicate: $[\![\text{might}_{epis}]\!]^w = \lambda$P . $\exists$w' $\in$ R(w)[P is True in w']

b. Uniqueness

Concept: An entity may be uniquely distinguished from all others

Predicate: $[\![\text{DEF}]\!]^w = \lambda$p . ιx[p(x) in w]

c. Sympathy

a. Concept: An individual sympathizes with another because of an event or state

b. Predicate: *$[\![\text{SYMP}]\!] = \lambda$x . λe . x is adversely affected by e, and sympathize(speaker)(x)

Beyond the broad intuition that these concepts and predicates may not be identical, research in general and linguistic cognitive development seems to give specific evidence making exactly these distinctions. An extensive body of research exists discussing the relationship between cognitive Theory of Mind development and its linguistic correlates, and there is indication that the two follow different (but interacting) developmental paths (de Villiers 2007 and citations therein), but both are likely innate (Onishi and Baillargeon 2005).

Further, both benefaction and sympathy have been argued to be innate concepts which human children make use of in their social-cognitive development (Warneken and Tomasello 2009), but only benefaction seems to appear in verbal-inflectional systems (Cinque 2013). Thus, the fact that there does not seem to be any such functional predicate as that in (17c) provides evidence that the set of concepts expressed in the functional lexicon need not be, and indeed cannot be, identical to the set of concepts available to children learning a language.

While the functional lexical items expressing speaker belief states and uniqueness are closely related to cognitive concepts, other native concepts do not seem to have corresponding functional heads; so the functional lexicon and the inventory of native concepts are clearly not co-extensive. At the same time, the former must in some sense be limited by the latter. If we are following Cinque's line of thought that natural languages' functional lexicons are limited to a particular native set, we would not expect a functional head which expresses some concept which itself is non-native. For example, we wouldn't expect a functional head expressing the concept of earning a degree, expressed as a verbal affix -*gra* (18).

(18) Non-native concept: Earning a degree

$[\![\text{GRAD}]\!] = \lambda x\, \lambda p\, \lambda e$. agent(x)(e) and p(e) and $\exists$e'[cause(earn.degree)(p)(e)]

Charles *gra*-studied (Charles earned a degree by studying)

While we can freely speak about someone earning degrees by particular means, by the assumptions made here we don't expect that languages will use this sort of morphology to express it.

1.2.3 Minimalism, Markedness, and Parameter Setting

Various perspectives on the language faculty, acquisition, and variation challenge the notion that an innate set of Parameters needs to be built into the grammar. In particular, one family of critiques comes from the nature of representational markedness in

syntax. Newmeyer (2004) in particular discusses the issue of markedness as a problem for Parameter Setting models: various issues in cross-linguistic variation might prompt us to assume certain Parameters have a marked value (whether preposition stranding is preferred or dispreferred, what the least marked constituent order is, etc.), but do not show delay in acquisition (2004:213-214). He builds on a model proposed by Hawkins (2004) which predicts delays in acquisition for certain patterns relative to others, and describes syntactic markedness based on principles of processing (in other words, “Third Factor" properties).

An advantage of this sort of idea is that certain delays in acquisition can be accounted for in terms of processing which correlate to typological variation: delays in the acquisition of Object Relative Clauses, which are sensitive to syntactic structure (O’Grady 1997; Friedmann, Belletti, and Rizzi 2009) can be seen as paralleling the cross-linguistic implicational hierarchy on Relative Clauses (Keenan and Comrie 1977). Since there is no apparent single Parameter on the accessibility of different syntactic positions for Relative Clause formation, this acquisition path could be expressed as an instance of an implicational hierarchy based on ease of processing (Newmeyer 2004, Hawkins 2004).

However, the same model runs into problems based on the sort of data discussed above, some of which Newmeyer himself presents: in particular, he cites V2 word orders in German apparently being learned early by children as evidence against a Parametric conception of markedness (2004:214). This would seem to contradict another principle Newmeyer discusses, Minimize Forms (Hawkins 2004:38), which pressures learners to (a) minimize the number of idiosyncratic representations across the grammar and (b) minimize the complexity of individual representations. Given standard assumptions of the syntax of V2, even constructing a simple SVO sentence in German involves Verb movement and Subject movement, as well as learning variation between SVO and SOV clauses.

While a theory of acquisition and variation must take into account markedness phenomena which relate to Third Factor properties, not all points of acquisition or delays in acquisition (or variation for that matter) can be understood in this way. Rather, this could be framed as a reflex of the extent to which a language's particular properties can be identified using simple Minimalist tools. For example, despite the somewhat complex representations necessary to generate V2, simple tools will necessarily generate such representations.[2]

(19) Learning V2

a. If *John eats rice* and *John will rice eat*, then

John eats rice → [John [$_{aux}$ eats$_v$] [rice t_v]]

b. If *John eats rice today* and *Today eats John rice*, then

John eats rice → [John$_i$ [$_{aux}$ eats$_v$] [t_i rice t_v]]

This is to say that, hearing that SVO alternates with S-Aux-OV (19a) will indicate to the learner that the Verb moves to the position where Aux is pronounced, and hearing SVO alternating with Adv-VSO (19b) indicates that the subject moves above this position. These are the essential pieces of V2, and they are easily inferable by the child despite the opacity of simple SVO word order.

In the following chapters, I will discuss other instances of crosslinguistic variation, and the steps in the acquisition of the relevant constructions, which do and don't allow for these sorts of easy inferences based on Minimalist syntactic assumptions. I will show that, where these tools can be simply applied, the acquisition of the relevant properties

[2]This assumes that there is some notion of uniformity between clauses across all of a language like German; in principle nothing might rule out a grammar for German where SVO sentences are represented identically to those in English - ie., one where the underlying order is not Verb final. This sort of grammar would be somewhat unwieldy, though, as it would have to include an exceptional rule denoting something like "VP is head final except where there is no overt Auxiliary and nothing moves above C." Such a rule could be formulated, but one might reasonably ask why the learner would postulate that rather than assuming VO patterns always derive from OV.

comes early, whereas in the cases where they can't, acquisition is delayed. Further, I will show that this is the case despite no apparent variations in markedness of the type Newmeyer discusses, and that it conforms with particular types of micro-parametric variation patterns within typologies and within acquisition paths. The domain of this investigation will be Free Relatives (FRs), and their relation with other Wh Movement Constructions and the nominal domain in English.

1.3 What is gained by (Minimal) innateness

An important, and oft-cited, example of a way in which innate properties of linguistic structure observably guide acquisition is constituency: nothing in a General Cognitive theory of language acquisition requires that linguistic rules make reference to hierarchical syntactic structure, and yet there is ample evidence to suggest that they do (Chomsky 1959; Crain 2012; Crain, Koring, and Thornton 2016). That is, given a pattern which seems to show a syntactic dependency or evidence of what one might consider Movement by any definition, a learner will have to postulate that the relevant rule makes reference to a structure.

This is commonly described in terms of the construction of English polar questions in sentences involving Relative Clauses (Chomsky 1968): sentences like (20a) can only correspond to question forms which move an auxiliary chosen because it is structurally highest, rather than a linear rule or some other type of rule.

(20) a. The eagles [that can fly] will eat

b. Will eagles [that can fly] eat?

c. *Can eagles [that fly] will eat?

Evidence that children can and must make reference to constituency structures in their linguistic representations has been given based on a number of experiments based on question formation of the type in (20) (Berwick et al 2011) and syntactic disjunction

(Gualmini and Crain 2005), among many others. This is counter to assumptions which are necessarily made in a naive General Cognition/Usage-Based system (see discussion in Crain et al. 2016).

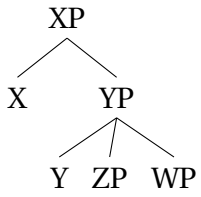
So, we have some simple, but crucial ways in which innate rules about syntactic structures restrict possible grammars, and one which is common to almost all generative models and certainly any which are descendants of the “Extended Standard Theory" of Chomsky (1977, et seq). However, the current discussion concerns whether there are such properties which come necessarily along with the assumptions of Minimalism. We can indeed identify, at least in principle, some such properties.

1.3.1 The things you can’t do

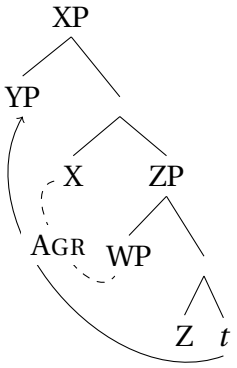
One simple example of a property of a Minimalist system that any learner will have to assume is binarity: if all derivations consist only of iterations of binary Merge, a learner would never posit structures which must be derived from some other type of Merge, resulting in ternary branching. Another, more complex example, concerns internal Merge (ie. Movement). We generally assume that any instance of Internal Merge (ie. Movement) is necessitated by some Agree relationship between the thing which moves and some head close to the position to which it moves (Chomsky 1995). If this is the case, the learner would also not posit structures where Agree relationships trigger movement of something which is not part of the Agree relationship. These are both shown in (21): the tree in (a) involves ternary branching, and is thus non-generable by Merge. The tree in (b) shows an Agree relationship between X^0 and WP, but movement of YP, which is not part of any Agree relationship. This movement is thus unmotivated by Agree, and so again this relationship would not be postulated by a Minimalist learner.

(21) Impossible structures

a. Ternary branching



b. Unmotivated movement



This sort of bias is not trivial; it can ban certain grammars which are unattested in natural languages, such as purely-linear rules, or some pattern of agreement which would derive from (21b), or, more strikingly, grammars which violate island constraints and other phenomena which are derivational in nature. If the etiology of island effects is based the sort of derivation which would be involved in generating island-violating relationships, an innate Minimalist syntax will prevent any grammar from being postulated which does not obey island constraints.

Further, the fact that the structures in (21) cannot be generated requires that children make specific hypotheses about the representations which guide their interpretation of them, and the generalizations they make. Given that (21b) is not possible, a child must make an assumption that movement of a phrase corresponds to a syntactic relationship between the moved XP and a head local to its new position. In other words, once a child determines that there is a trace of an XP in the representation of a string, they will make

hypotheses like the ones in (22), something which will determine the formal features they postulate for the relevant lexical items.

(22) If [... XP_i Y ... t_i ...], then Agree(Y)(XP);

If Agree(Y)(XP) then ∃ F such that Y = [... F ...] and XP = [... F ...]

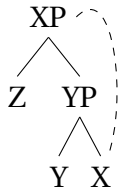
The other properties of Merge generate other important limits on derivations which will be crucial to solving the specific acquisition problems discussed here.

As an example, recall from above that the distinction between simple and complex syntactic objects (Heads and Phrases) is visible to Merge, and so other constraints may merge from this distinction. We can apply this to rules about Labeling and Head Movement. Specifically: assume that the label of any node in a Tree must project from some node the former immediately dominates, and that only a Head can move to another Head position, never a Phrase.

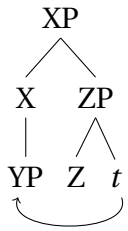
To make this slightly more concrete, the trees in (23) show two more impossible representations according to these assumptions: in (a), the label of the root node skips a node and comes from a node further down the Tree, rather than coming from a daughter node; in (b), YP adjoins to X^0, an instance of a Phrase engaging in Head movement, which is impossible.

(23) More impossible structures

a. Illicit labeling



b. Illicit Head movement



All of these are examples of representations which are impossible in a Minimalist system, and so a Minimalist learner would never posit them. Thus, even with strongly Minimalist assumptions we can articulate ways in which the Acquisition Device restricts children's possible hypotheses about representations in their language: sentences which reflect derivations like those in (21 and 23) will not be observed either in children's speech or comprehension, or in any natural adult language.

Though testing the prediction that these structures should not be attested in observable patterns requires very subtle types of methodological tools - for example, testing whether or not a moved phrase has moved to a head position is quite opaque on the surface - we can describe what some such tools might be looking for exactly. The ban on ternary branching predicts that children will not allow consequences of C-command (eg. binding) of an XP over a YP to the exclusion of an intervening ZP; and the restrictions on Phrasal vs. Head movement makes predictions about learners' behavior regarding distribution and selection of different types of constituents: an environment selecting an XP will not allow a string generated by the movement in (23b); and so on.

For example, to begin to understand the role of requirements of Head Movement and Labeling as guides in the acquisition process, we could construct a similar guideline to grammar hypothesizing like (22) that allows learners to posit, null syntactic objects which provide labels necessary in some syntactic environment (eg., if a string appears in a DP position, there must be a D^0). This would be characterized as in something like (24):

(24) **Head Positing Formula**

If [α [β ...] ...] whereα selects XP, then [α [$_{XP}$ ...] ...]

If [α [$_{XP}$ ...] ...], then [α [$_{XP}$ X^0 ...] ...]

That is, if β occurs where an XP is selected for, β must be an XP; and, for any XP, there must be an X^0 which is accessible to the labeling algorithm.

Chapters 3-4 will discuss particular ways in which some of these hypotheses can be tested, and evidence that these principles are indeed obeyed by children during various stages of acquisition.

1.3.2 Non-Minimalist and non-Nativist approaches

Any theory of Principles and Parameters - either the robust, articulated sort of UG assumed in Government-Binding theory or the more abstract UG of a Strong Minimalist approach - takes innate linguistic knowledge to be essential to language acquisition; that is, it is a Nativist explanation. The alternative to this is a non-Nativist approach, which posits, essentially, that only the second two factors play a role in language acquisition: the experience of the learner and general properties of the mind (or the properties of the mind which are absolutely necessary for language, like working memory and auditory processing) interact in such a way that the learner makes sufficient generalizations about their target language that they demonstrate adultlike linguistic behavior.

Since Minimalism does so much to restrict the substance of the First Factor, it starts to resemble in some ways this “Usage-Based" approach (Tomasello 2000, inter alia). Even under this sort of approach, it has been suggested that apparent restrictedness of output grammars could be understood as results of the “Third Factor" applying pressure to the acquisition process (Hawkins 2004, O’Grady 2005). Given simplicity of the innate language faculty, and the resulting explosion of potential representations (as discussed above), one could reasonably ask what the native Language Faculty actually contributes to the acquisition process that isn’t covered by other two Factors.

Several answers to this question have been suggested in the literature: one is that the set of possible functional predicates postulated above does much of the work in guiding acquisition; another is that the ability for (certain types of) recursive representations (Roeper 2011) guides learners into particular representations and not others. Another line of thought allows certain specific Principles in the GB sense into a generally Minimalist syntactic system, like principles of Theta theory (Pearl 2016) or the Binding conditions (Lukyanenko, Conroy, and Lidz 2014; Deen and Timyam forthcoming), which largely have to do with properties of the interfaces (the latter two could plausibly be described in terms of the syntax-semantics interface).

Another approach which adds innate material to the Minimalist first factor is the Maturational approach of Borer and Wexler (1987), which posits that in addition to any particular properties of syntax which a learner may have instant access to from the moment acquisition begins, there are particular operations which are native to the human Language Faculty, but which do not become available to speakers until somewhat late in development. This hypothesis is deployed to explain facts like children’s apparent delay in acquiring, for example, A-movement constructions like Passivization. This, like classic Principles-and-Parameters approaches, is contrary to Minimalist assumptions, as it gives extra functionality to the first factor in addition to Merge, in the form of speci-

fied operations like A-movement (Borer and Wexler 1987), or multiple instances of Agree (Wexler 1998).

Whether or not any of these approaches is on the right track, there is one major contribution of a Minimalist system to the Acquisition of a given language, and that is the role of formal features and derivational properties discussed in the previous section in yielding licit representations. A possible representation of a given sentence must involve a structured relationship between lexical items and formal features associated with them, and licit binary relationships between atomic parts of the structure which could have been derived by iterations of Merge. That means a learner will not posit grammars which produce structures which cannot be derived this way, or which involve no relationships (or impossible relationships) between formal features in the structure.

Of course, the possibility of Second and Third factor explanations for any given phenomenon must be considered. Thus, for any given observation about some stage of acquisition, we must ask: is this a result of limits of children's exposure to relevant structures (a Second Factor or input-based explanation), a property of children's performance limitations relative to adults (a Third Factor or processing-based explanation), or the state of the child's grammar at that time. Specifically, Second Factor explanations will refer to cases where children's exposure suggests that they should *expect* certain structures over others, or that their production closely mirrors rates of production of relevant patterns by caretakers. An assumption I will make here is that the burden of proof is always on Second and Third Factor explanations rather than grammatical explanations. The statements in (25) generalize this idea as it will be implemented here.

(25) An extra-grammatical explanation for an acquisition phenomenon succeeds relative to a grammatical explanation only in the case that:

(i) there is specific evidence observed in children's input that correlates closely with children's performance (Second Factor explanations);

(ii) there is evidence in favor of a specific hypothesis about the non-grammatical mental processes which give rise to non-adultlike behavior (Third Factor explanations); or

(iii) there is no conceivable Minimalist system which will generate the observed patterns.

The intuition behind (25) is that (Minimalist) grammatical explanations are flexible enough to accommodate a wide variety of possible grammatical stages, so in the absence of specific evidence for extra-grammatical explanations there are a great number of possible descriptions of pre-adult grammatical states.

It is important to note about (i-ii) that these refer to what one might call *purely* non-grammatical explanations - that children perform in ways which are purely reflective of their input, or that their behavior makes no particular reference to grammatical states. This is not an evaluation of explanations for particular grammatical states. Nothing excludes the possibility that children's acquisition path reflects properties of input frequency, or is restricted in some way by cognitive development more broadly. Rather it states that, absent of other specific hypotheses, non-adultlike behavior will be assumed to reflect a non-adultlike grammar. Further, this does not choose between particular grammatical explanations; any given explanation for an acquisition stage is subject to evaluation against other explanations. This simply proposes that grammatical explanations should be the first guess.

1.4 Wh Movement Constructions

There is a large literature on the set of constructions which involve movement of a constituent to a clausal specifier position and, in many cases, Wh morphology, broadly described as Wh Movement Constructions (Chomsky 1977). We can ask in general what the full set of constructions which could reasonably be included within this set as a nat-

ural class, but a partial set of the constructions traditionally considered to make up this class can be seen in (26).

(26) Wh Movement Constructions

a. Wh Questions

What did Charles eat?

b. FRs

Sebastian made what Charles ate

c. Headed Relatives

Sebastian made the food which Charles ate

d. Topicalization

The boiled eggs, Sebastian made

e. Tough Movement

The boiled eggs were tough to make

It is notable that not all of these require Wh morphology, and the latter two do not allow it:

(27) a. Sebastian made the food (that) Charles ate

b. *Which boiled eggs, Sebastian made

c. *The boiled eggs which eggs were tough to make

In general we can ask how a child acquiring English determines that these constructions must form a natural class. For the first three, we could say that the use of Wh morphology prompts the learner to generalize properties of their derivation and representation. However, this obviously does not account for the learning of Topicalization or Tough Movement, especially since there are other constructions which superficially look similar but have a different set of properties and likely different representations.

For example, Topicalization could easily be misanalysed as scrambling, which differs in its derivation and semantic properties in subtle ways from English topicalization (Kidwai 2000), and Tough Movement could be misanalysed as a Control construction (akin to an adjective like *eager*).

In fact, there is ample evidence that children experience difficulty making the latter choice - a child may analyze individual lexical items as having a Tough-type interpretation or Control interpretation, and the decision (C. Chomsky 1969 et seq).

(28) Sebastian is easy to please

a. Tough reading: It is easy to please Sebastian

b. Control reading: Sebastian has an easy time pleasing people

Even for Wh questions and relatives of different types, subtle differences in syntactic and semantic properties give rise to a complex set of problems for the child. Even if we assume that in some sense Wh movement is a primitive option available to the syntax, subtleties such as the distribution of different Wh items across the different constructions provides a problem for the learner to determine the places where the representations of FRs, questions, and headed relatives converge or diverge. Take (29-30) for example: FRs are unique among these three in not allowing *which*, and neither FRs nor headed RCs allow a phrasal Wh expression like *what bottle*

(29) a. Which did Sebastian make?

b. *Charles ate which Sebastian made

c. Sebastian made the food which Charles ate

(30) a. What (bottle) did Sebastian bring?

b. Charles drank what (*bottle) Sebastian brought

c. *Charles drank the wine what (bottle) Sebastian brought

Constructions like Wh-movement, which are broadly cross-linguistically attested, raise questions about what the set of native requirements of syntactic representation and derivation are. Work in syntax has attempted to describe common patterns like islandhood, which are acquired early, as results of Minimalist derivations, either by properties of the application of Agree (Rizzi 2013) or of Spell-Out (Chomsky 2000, 2001). This is a desirable goal in maintaining a Minimalist theory of UG, but it relies on properly articulating the relevant properties and the representations of the relevant constructions.

The same could be said for (surface) properties which are not universal: while differences between specific languages on which structures are allowed, in a Minimalist system the barring of specific expressions must make reference to the ways in which they are derived and represented in Minimalist terms.

A Minimalist theory of acquisition, then, involves a learner identifying the representations of strings in their language which block the illicit representations, and conform to general principles of syntactic representation of the kind discussed above. In particular, for FRs, the concern involves the finding the optimal way to unite the syntax of Wh movement with the syntax of the nominal domain in English. The first essential step in this is to label Wh strings as nominals, which can be done using the Head Positing Formula in (24) to posit the existence of a D head for a Wh string, as in (31)

(31) **Positing an FR**

If [V [$_{\beta}$ what ...] ...] where V selects DP, then [V [$_{DP}$ what ...] ...]

If [V [$_{DP}$ what ...] ...], then [V [$_{DP}$ D^{0} what ...] ...]

The rest of this dissertation will look in depth at the sorts of surfaces differences in (29-30) and discuss in particular the ways in which children must pick out peculiar properties of FRs based on patterns scene in Wh questions and nominals and the formula in (31), and evidence that shows the path they take in making these discoveries.

CHAPTER 2

THE FORM OF FREE RELATIVES AND THE NOMINAL DOMAIN

This chapter concerns the distribution of specific Wh expressions in different constructions. Specifically, I will attempt to derive the restricted distribution of Wh expressions in FRs, especially with reference to Wh words with nominal complements (Wh-NP), and discuss the nature of restrictions on “Pied Piping" of complex Wh expressions in different Wh constructions.

The primary focus will be on the blocking of Wh-NP in FRs (*Wh-NP), but not in Wh questions, as shown in (32).

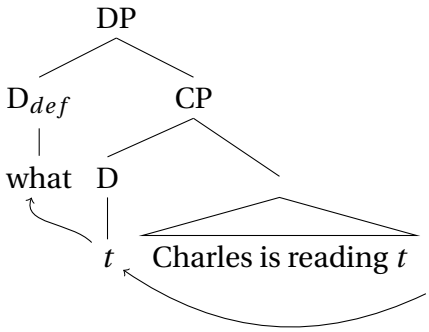
(32) ***Wh-NP**

a. Sebastian wonders [what (**book**) Charles is reading]

b. Sebastian wrote [what (***book**) Charles is reading]

Building on arguments from Caponigro (2002, 2003), and Donati (2006) and Cecchetto and Donati (2011, 2015), I argue that this effect is a result of the the structure for FRs in (33), which involves an FR being derived from a Wh question, and the Wh word moving from the Specifier of CP to adjoin to a type-shifting D^0 head, to fulfill the requirement that English definite descriptions have a phonologically overt determiner.

(33) English FRs



I show how this syntax can derive not only *Wh-NP, but also can derive other restrictions on Wh expressions in FRs.

The chapter is divided as follows: The first section describes the different distributions of Wh expressions in FRs and Wh questions, and provides the essential facts which derive the syntax in (33). The second section lays out the details of the syntax, and shows how it derives different properties of FRs and Wh questions. The third section discusses the broader issue of Pied Piping, by comparing the distribution of Wh expressions in FRs and questions vs. Headed Relatives. The fourth section discusses cross-linguistic variation of these properties, and the relationship between this variation and other types of variation in the nominal domain.

2.1 Wh expressions in FRs

The set of Wh strings allowed varies by construction in English, and Wh expressions allowed in FRs vary across languages. This section deals primarily with the distribution of different Wh expressions in FRs versus Wh Questions in English; it also compares English with some other languages where FRs involve Wh words (Wh-FRs). It is worth noting that other languages have different sorts of morphology associated with FRs; some of the facts described here do apply in these languages, while others don't. We will generally see that, where Wh expressions appear in FRs, only a subset of the Wh words used in questions may appear in FRs, an observation first made by Caponigro (2003).

Given the variation of things which might broadly be construed as FRs, a firm definition of the phenomenon is called for. Caponigro proposes a definition which restricts FRs to Wh-FRs (2003:8), but the similar properties of other types of relativization call for a slightly broader definition. Thus, I propose the working definition in (34), in comparison with (Externally) Headed RCs.

(34) A **Free Relative** is any nominal which contains a CP which restricts the referent, but not an NP restrictor external to the CP.

FR: $[_{DP}$ D $[_{CP}$ ...]]

Headed RC: $[_{DP}$ D ... $[\text{NP}_i]$ $[_{CP}$ ... t_i ...]]^1

Thus this definition of FRs includes both English-style Wh FRs, correlatives including both those with Wh morphemes (Turkish) and without (Hindi-Urdu), and Internally-Headed Relative Clauses, which have a notional nominal head which is within the CP rather than outside of it (Korean, Lakhota). This chapter will deal primarily with the form of Wh-FRs in English, with some reference to data FRs in other languages which inform the place of English FRs in a greater typology. See Appendix A for a discussion of Wh words in the derivation of Headed RCs.

2.1.1 Telling FRs from Wh questions in English

English FRs have the property of having the surface appearance of embedded Wh questions, but the distribution of nominals, rather than of question clauses. The sentences in (35) and (36) show this: Strings like *what Charles is reading* can appear in either nominal positions (such as complement of *write*) or clausal positions (comple-

[1]This operates on the assumption that CP restrictors in headed RCs which are not linearly adjacent to them have been moved; thus this definition can describe Headed RCs of various kinds. Note also that this definition does not itself distinguish between movement and non-movement analyses of headed RCs, it only assumes that some sort of trace exists within the CP. For languages where headed RCs include a resumptive pronoun, assume the trace is equivalent to this pronoun for at least these schematic purposes.

ment of *wonder*). Other strings, which are unambiguously DPs (*this NP*) or question CPs (*whether S*), can only occur in one of the two.

(35) Nominal positions

a. Sebastian borrowed this book.

b. Sebastian borrowed what Charles was reading.

c. *Sebastian borrowed whether Charles was reading this book.

(36) Clause positions

a. *Sebastian wonders this book.

b. Sebastian wonders what Charles was reading.

c. Sebastian wonders whether Charles was reading this book.

The obvious difference between the two is in their semantics: FRs like the one in (35b) are definite descriptions, denoting entities, and questions like (36b-c) denote sets of propositions (by standard question semantics; Hamblin 1973, Karttunen 1977, inter alia). A parallel difference is seen in subject positions: certain predicates, like *get better*, allow for only nominal subjects, while others, like *be obvious*, require clausal subjects. As with the examples above, (37-38) show that Wh strings can appear in both positions.

(37) FR subjects

a. Sebastian's book got better

b. What Sebastian was writing got better

c. *That Sebastian wrote poetry got better

(38) Clausal subjects

a. *Sebastian's book was obvious

b. What Sebastian was writing was obvious

c. That Sebastian wrote poetry was obvious

Thus, the **external** syntax of FRs and questions - the environments which select them - provides the necessary contexts to examine properties of their **internal** syntax. In the next section I will use this diagnostic to look at some of the differences between the two.

2.1.2 The distribution of Wh expressions in questions and FRs

Looking in slightly more detail at Wh expressions in FRs vs. in Questions reveals a number of morpho-syntactic differences between the two constructions. In particular, FRs are more restrictive as to which Wh expressions are available, including a restriction against Wh-NP sequences, disallowing pied piping of prepositions, and disallowing particular Wh words, all of which are permitted for matrix and embedded questions (Larson 1987, Donati 2006, Chomsky 2013).[2]

(39) Wh words in questions

Sebastian wonders { what (book) / who / which (book) / how much / whose book / %to whom / where / why } Charles was reading

[2]We could add *when* to the list of words that fit both positions, though there isn't a reasonable interpretation for *borrow when Charles was reading*.

(40) Wh words in FRs

Sebastian borrowed { what (*book) / %who / *which (book) / *how much / *whose book / *to whom / where / *why } Charles was reading

In all of these cases we see that phrasal Wh expressions which are allowed in questions (Wh-NP, *how much*, Wh words with pied-piped prepositions or possessa) are generally barred in FRs. But this is only part of the issue - *which* and *why* are disallowed in FRs, and *who* in FRs is rejected by many speakers.[3] So while complexity of the Wh expression is clearly a factor, there is something else at work in picking out the Wh words allowed and disallowed in FRs.

In addition to facts like (39-40), morphosyntax can disambiguate so that only FR readings are available. For example, FRs but not typically Wh-questions allow the *-ever* morpheme on Wh words (Bresnan and Grimshaw 1978).[4] However, in other ways, such as the constraint against PP pied piping, Wh-ever behaves like FRs

[3] *Who* FRs are accepted by a set of speakers without a particular geographic or generational boundary. See footnote in Chomsky (2013:46).

[4] Bresnan and Grimshaw (1978:334) give the example *I'll inquire whatever he is selling* as ungrammatical, but some speakers accept this, as equivalent to *Whatever is he selling?* Bresnan and Grimshaw note that the latter sort of *ever* floats (*What is ever the matter?*), but do not note that there seems to be some gradience in the acceptability of their example of embedded Wh-ever in terms of the selecting predicate. However, the badness of pied piping with Question Wh-ever (41c) reinforces their point that Wh-ever is not typically available in embedded Wh questions.

(41) a. Sebastian borrowed whatever Charles was reading.

a'. %Sebastian wondered whatever Charles was reading.

b. *Sebastian kissed to whomever Charles was reading.

b'. *Sebastian wonders to whomever Charles was reading.

To build a complete theory of the distribution of Wh expressions in English, a distinctions must be made between Wh-NP sequences and other complex Wh expressions. To be clear about the terminology I am using here, I make the distinction in (42), whereby Wh-NP is a specific type of pied piping.

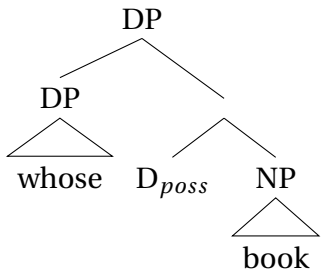
(42) Types of complex Wh expressions

- **Wh-NP**: Constituents consisting of a Wh word with an NP restrictor (What/which NP, How much/many NP), particularly when they are moved together.
- **Pied piping**: Any instance of a Wh word moving with a constituent which contains it (including Wh-NP, PP pied piping, and Whose-NP pied piping).

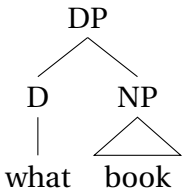
By these definitions, the notion of pied piping covers all instances of phrasal Wh items which move together, Wh-NP refers to a particular sort of set. This captures the notional difference between the movement of a Wh word and its maximal projection (the smallest XP containing the Wh word) and movement of a larger XP which contains the Wh word.

(43) Varieties of Pied-piping

a. Possessum pied-piping



b. Wh-NP



It is also necessary to capture the full range of variation between different Wh movement constructions; for example, headed relative clauses disallow Wh-NP but allow other types of pied piping, Wh questions allow all (or most) kinds, and FRs disallow all kinds, as shown in (44).

(44) Wh-NP in different constructions

a. Sebastian borrowed what (*book) Charles was reading.

b. Sebastian wondered what (book) Charles was reading.

c. Sebastian borrowed the book which (*tome) Charles was reading.

Appendix A contains a discussion of the peculiarities of RCs in this regard, and a broader theory of pied piping. Here I will remain focused on the specific constraints on FRs.

What we see broadly is that if a Wh expression is used in an FR, it can also be used in Wh questions. That is, there is no Wh word that occurs just in FRs; if we imagine that there was a language English' that had an additional Wh word reserved for non-human animates, *whanimal*, we would still not expect a pattern like (45).

(45) a. Sebastian kept whanimal Charles brought.

b. *Sebastian wondered whanimal Charles brought.

The absence of such patterns cross-linguistically was observed by Caponigro (2003), and described as in (46), led to the analysis of Wh FRs as being derived syntactically, as well as semantically, from Wh questions. Table 2.1 (from Chierchia and Caponigro 2013) shows cross-linguistic data from typologically diverse languages, English, Italian, and Mixtec, which demonstrate the robustness of this generalization.

(46) **Caponigro's Generalization**:

If a language uses the Wh-strategy to form both Qs and FRs, the Wh-words found in FRs are always a subset of those found in Qs. Never the other way around. Never some arbitrary relation between the two sets.

Table 2.1: Wh expressions allowed in FRs crosslinguistically

	What	Who	When	How	How much	why	Wh-NP
English Q	✓	✓	✓	✓	✓	✓	✓
English FR	✓	%	✓	✓	*	*	*
Italian Q	✓	✓	✓	✓	✓	✓	✓
Italian FR	%	%	✓	✓	%	*	*
Mixtec Q	✓	✓	✓	✓	n/a	✓	✓
Mixtec FR	✓	✓	✓	✓	n/a	*	*

Not every language has FRs which involve Wh morphology; languages like Hindi-Urdu, Korean, and Italian Sign Language have FRs whose morphemes are distinct from Wh morphology. Given that these do not appear to be derived out of Wh questions, and don't involve Wh morphology, it is trivially true that Caponigro's Generalization does not hold for these languages; and in fact we see that the restrictions on types do not generally have the same sorts of restrictions as FRs in languages like English or Mixtec. Below I will provide some discussion to the restrictions that do seem to exist in other FR systems.

2.1.3 Variable ambiguities

Much of the data presented above focuses on the fact that only particular substrings (such as *what Ben brought*) are ambiguous between Wh questions and FRs. However, it is important to note - and will be very useful both here and in comparing adult and child linguistic knowledge (as in Chapter 3) - that certain ambiguities persist at the sentence level. In particular, when a predicate allows both clausal and nominal arguments, we see an ambiguity between two readings: a clausal, or question reading, and a nominal, or referential reading.

(47) Ambiguous positions

a. Ben saw {that Molly was not having fun/the snake}.

b. {That Molly was not having fun/The snake} surprised Ben.

c. Ben saw what Molly brought.

d. What Molly brought surprised Ben.

(48) Disambiguation

a. Ben saw {what gift/whatever} Molly brought.

b. {What gift/whatever} Molly brought surprised Ben.

As Wh-NP is available for questions (clauses) but not FRs (nominals), it follows that for sentences like (47c-d), the ambiguity is one of category (after Donati 2006; Cecchetto and Donati 2011, 2015, inter alia): referential readings come about when the category of the string is DP, and question readings come about when the string is a CP.[5] Thus we get the disambiguation between (47c) and (48a); in the absence of Wh-NP, nothing in

[5]While I will go on to suggest that DP is more or less the exact right description, I will not spend a lot of effort here in choosing between CPs or other clause-level labels (ForceP, SpeechActP, etc.), though there may be good reason to suggest that in at least some cases it is one of these rather than CP per se.

the sentence forces a choice between CP and DP interpretations, whereas in the latter, *Wh-NP blocks a DP interpretation.

The next question which must be addressed is, what are the two readings for a sentence like (47c)? Generally, there is evidence that at least standard FRs are definite descriptions (or more particularly, definite descriptions derived from questions; see Caponigro 2003, Chierchia and Caponigro 2013) of entities with the property denoted by the gapped clause, while questions are propositions or sets of propositions (Hamblin 1973, Karttunen 1977, inter alia). A pseudo-formalism of each of these readings is in (49).

(49) a. ⟦see [$_{DP}$ what Molly wrote] ⟧→ see the x such that Molly wrote x

b. ⟦see [$_{CP}$ what Molly wrote] ⟧→ see for which x, Molly wrote x = T

Basically, (49a) is true as long as the subject sees the thing with the property of Molly having read it; (49b) requires that the subject come to know, by seeing, the answer to the question 'what did Molly write?'.[6] Certain conditions can piece these two readings apart; if one sees something with the property of having been written by Molly, without one knowing that this property belongs to it, then only the DP reading is true. And crucially, in just these cases, ambiguous strings are deemed true, while strings which can only have the syntax of questions are false, as (50) shows.

(50) Situation: Charles sees several books laying open. Among them is the book Sebastian wrote, though he cannot tell which book is which.

a. Charles saw what Sebastian wrote = True

b. Charles saw what book Sebastian wrote = False

[6]Not all uses of *see* [$_{CP}$] seem to involve literal seeing; *I see that Charles is here* is typically acceptable in the case that the speaker **hears** Charles without seeing him. So perhaps a better way to describe this meaning would make reference to evidentiality. The paraphrase would then be something more like "receive direct evidence of the answer to the question 'what did Molly write?'"

In this sort of case, in the presence of a Wh-NP sequence, the FR reading becomes unavailable, and so (49b) can only be false; however, the definite description reading holds, so the ambiguous sentence in (49a) that can correspond to an FR structure can be evaluated as true. I return to this issue in the following chapter.

So, we can describe complete knowledge (and thus complete acquisition) of FRs, and in particular the *Wh-NP effect: knowledge of the effect (or the properties of the grammar which bring it about) involve not only non-production of illicit FRs, but also properly rejecting (50b) as true in a situation such as this – in other words, knowing that the string in (50b) cannot be a definite description.

2.1.4 Some exceptions to Wh-NP

The focus of this chapter will be on a representation for FRs which derives effects like Wh-NP - specifically, what we might call “simple" FRs, of the type seen above. However, it is worth noting now that there are similar constructions in English which do not all show the same types of restrictions on Wh expressions. Specifically, relatives with Wh-ever morphology allow a wider range of Wh expressions, notably including Wh-NP - though Wh-ever relatives are sill restricted in some ways, as shown in (51).

However, a point that must be made about Wh-ever is that it does in fact allow for Wh-NP, and in fact allows for other Wh expressions not allowed in standard FRs (51). English also has two constructions with similar form and distribution to that of simple FRs, but which are less restrictive in the Wh expressions they allow. The first is relatives with Wh-ever morphology (51); the second is FRs which look quite similar to simple FRs but carry an implication of the cardinality of the referent being small (52), observed by Grosu (1996). Neither of these show *Wh-NP.

(51) a. Molly will eat whatever snack Ben brings

b. Molly will eat however many snacks Ben brings

c. ?Molly will meet with whoever’s work Ben is interested in

d. *Molly will meet with to whoever Ben sent a gift

(52) a. Molly will eat what snacks Ben has

b. ?Molly will eat how much Ben has

c. ?Molly will meet with who's work Ben is interested in

d. *Molly will meet with to whom Ben sent a gift

Both of these types of constructions have distributions similar to that of standard FRs, but allow a broader range of Wh expressions. Other than the fact that the -ever morpheme doesn't typically occur in Wh questions, these do still conform to Caponigro's Generalization. But, there still exists the question of why they do not share all the syntactic properties of FRs. This fact has led to some question about whether or not these really do constitute FRs as such (Battye 1989, Cecchetto and Donati 2011). I will examine these in greater detail in Chapter 4, and argue that they are indeed a type of FR, with a syntax minimally different from the one proposed below.

2.2 The representation of FRs

We in general want to assume that something like the distribution of Wh items in a given construction is a necessary result of some basic mechanisms of the grammar and some necessary properties of the atoms of the construction (say, Wh words). In the case of FRs, the pieces which must interact are Wh words, Wh movement, and the syntax of the nominal domain; FRs are nominals which involve (at least apparent) Wh movement, and allow a limited number of Wh words. Here I will examine the ways in which they interact, and what essential properties of the Grammar constrain this interaction.

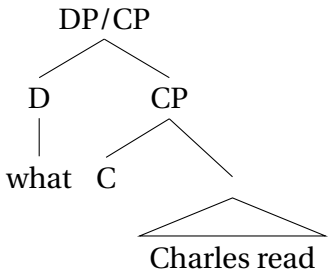
To begin, I will adopt two basic assumptions about the syntactic form of FRs: (i) that in the type of FRs under discussion here, the Wh word occupies the D^0 which is the head of the entire FR (following Donati 2006; Cecchetto and Donati 2011, 2015); and that (ii) FRs are syntactically (as well as semantically) derived from Wh questions, and consist of

a +Wh CP dominated by a DP with a definite operator (following Caponigro 2002, 2003; Chierchia and Caponigro 2013).

The first idea is taken as a way of deriving *Wh-NP in FRs. Cecchetto and Donati (2015) argue that *Wh-NP effects occur because, for a Wh structure to be in a nominal position, the Wh word must be in a position where it can “relabel" the entire substructure as a DP; in the structure in (53a), either the C Head or the D Head can label the structure, resulting in this string having the distribution of both DPs and CPs (as discussed above).[7]

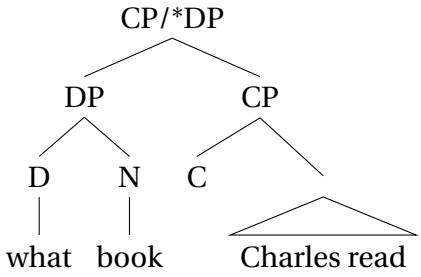
The syntax they assume is shown in (53-54). Something important to note here is that, at the stage in the derivation shown here, the label of the root note is undetermined: a syntactic structure is only labeled when it is merged with another. The tree in (53) has both a D Head and a C Head in positions which are visible to Agree, and thus may provide the label for the given trees. However, in (54), a phrasal Wh item is moved, and an adjoined XP cannot relabel a structure in which it moves; so, only C can project its label, and the associated string only has the distribution of a CP.

(53) The structure of ambiguous Wh strings



[7]This is in fact quite similar to the “Head Analysis" of FRs explored in older work by Bresnan and Grimshaw (1978), Larson (1987), and later by Izvorski (2000), all of which have the label of the FR related directly to the Wh expression which has moved. Izvorski in particular has a “Move-then-Project" analysis which is extremely similar to Cecchetto and Donati’s Relabeling analysis, but does not have a restrictive system for relabeling which accounts for effects like *Wh-NP. These eariler discussions typically concerned Wh-ever FRs; this will be revisited in Chapter 4.

(54) The structure of unambiguous Wh-NP strings.



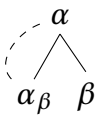
This elegantly derives *Wh-NP from a single supposition: that movement of simplex syntactic objects (ie. heads) – but *only* simplex objects – may “relabel" a structure. This possibility of a system of **relabeling** syntactic objects by movement is driven by an algorithm for probing/labeling, taken to be a primitive feature of the grammar. This system comes from two steps: a general Probing Algorithm which labels objects created by External Merge, and an addendum which labels objects which have been changed by Internal Merge

(55) **Probing Algorithm** (Cecchetto and Donati 2015:39)

The label of a syntactic object $\{\alpha, \beta\}$ is the feature(s) that act(s) as a probe for the merging operation creating $\{\alpha, \beta\}$.

This works simply enough: If α has a feature which selects β, Merge creates an object labeled α. This is schematized in (56).

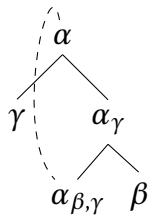
(56) Labeling of new object



The object created by Merge in (56) is motivated by a selectional feature (ie. a Probe) on α; thus by (55), α projects its label onto the new object. In less simple cases, a Probe may have more than one selectional feature; in this case, the lower object projects its

label through two instances of Merge.[8] This is, roughly, a case of labeling in a Spec-Head Merge. This is schematized in (57).

(57) Label projection

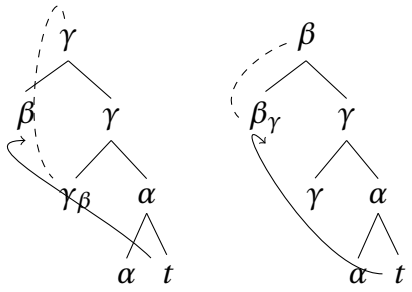


Since α selects for two different objects, its label projects through two instances of Merge, to create the larger object in (57).

In the case of Internal Merge, the system of C&D allows for labeling ambiguities. An assumption made implicitly by C&D is that an instance of Internal Merge may be driven by a Probe in either the outer or inner realm of the operation; that is, the label may come from either the thing which moves, or the place to which it moves. The former is a case, in their terms, of Relabeling. This is schematized in (58).

(58) Internal Merge and labeling

External labeling Internal (re)-labeling

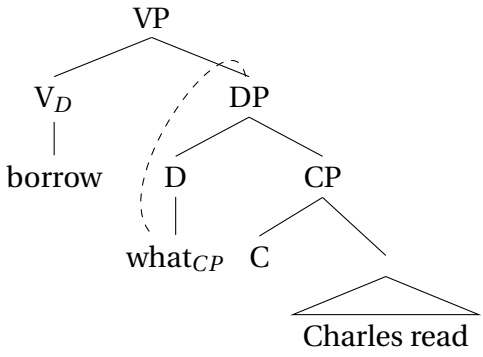


[8] C&D use the example of a transitive verb for this (2015:41), wherein an object {read,book} will be labeled as a Verb, but still have a selectional feature seeking an external argument; this is thus incompatible with other theories of Argument selection like that of Kratzer 1996; however, these assumptions may be adopted slightly to accommodate this model. I will not draw this out in detail here.

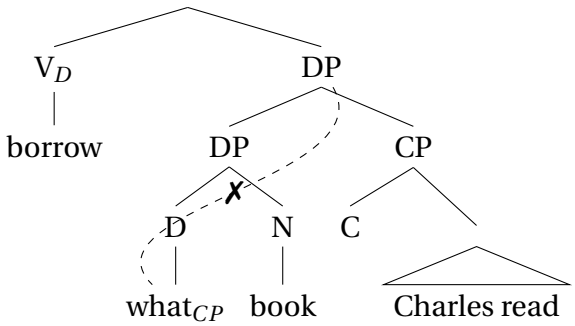
Here is where the schema for the ambiguity of FRs and questions of the type in (53) comes in. Under the assumptions of C&D, a question is a case of “external" labeling: C selects for a Wh word, and so the result of internal merge of *what* remains a CP. An FR, on the other hand, is a case of Internal relabeling: *what* in this case is a D which selects for a CP (2015:47), and so the result of Internal Merge is a DP.

In this framework, the selecting environment for this ambiguous structure will choose either an Externally-labeled CP in the case of (embedded) questions or an Internally relabeled DP in the case of FRs. *Wh-NP is simply the outcome of the structures in (53-54) and restrictions on where labels come from: the Wh word in an FR is able to project its label (and thus merge with a predicate which selects for a DP) only if it is moved alone (59), and not if it is moved within a larger DP (60).

(59) Selecting for D



(60) *Selecting for D



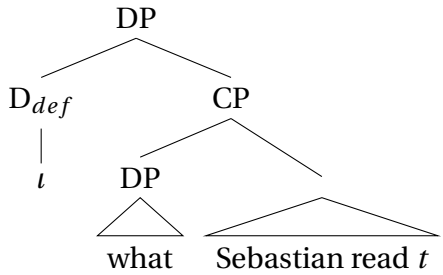
While the syntactic assumptions of this approach have an intriguing variety of consequences and applications (as C&D’s discussion displays), a concern is the question of

how the resulting structure is interpreted. In particular, we might ask of the structure in (53), what accounts for the different interpretations of *what Charles read* when the only difference is the syntactic feature which is selected? C&D make it clear that in their framework, the only difference between an FR and a question (where they are ambiguous) is whether the D^0 is a probe or the C^0 is a Probe, and thus the conditions under which their Probing can succeed.

However, assuming a compositional semantics, if FRs and embedded questions have all the same syntactic elements, this cannot be the whole story. There is no clear mechanism for how the projection of a label in itself would contribute to semantic composition, even if the categorial features themselves might have some semantic value, and so this syntax on its own doesn't have a way to derive the different meanings discussed above. A second factor, regarding the semantic pieces of an FR, must be added to the analysis.

In attempting to derive the relationship between Wh words in questions and FRs, Caponigro (2002, et seq) proposes that FRs are in fact derived from Wh questions, and denote an entity which provides an answer to the question from which it is derived (So, [$_{DP}$ *what Sebastian read*] denotes the thing which answers the question *What did Sebastian read?*). This is done by the inclusion in the representation of FRs of a definiteness operator ι which type-shifts the question structure. Caponigro assumes the operator to be a covert D Head, and thus results in a structure like (61), based on Caponigro et al (2012).

(61)



The Operator Chierchia and Caponigro (2013) make use of to derive definite descriptions from questions semantically operates on the assumption that a “short answer" for a question denotes a set of entities (eg., the set of things Sebastian might have read) rather than sets of propositions (Sebastian read War and Peace, Sebastian read Anna Karenina, etc.). If this is not a tenable assumption about the nature of “short answers" (see Weir 2014), another formalization would be needed. In this case, the CP would be more akin to the CP in a headed RC, and simply denote a property (ie., ‘What Sebastian read’ = the set of things Sebastian read). This formulation of D_{rel} is used in Caponigro et al (2012).[9]

Each of these has an equivalent result, with the full DP denoting the maximal element of the set of things which Sebastian read. I show this in (62): whether CP denotes a set of propositions or a set of questions, composing with D_{rel} will give a definite description with the semantic output in (62).

(62) a. $[\![CP]\!] = \lambda P \exists x[P = \lambda w[\text{Sebastian read x in w}]]$

b. $[\![D_{rel}]\!] = \lambda Q[\iota x[TP(Q)(x)(w)]]$

c. $[\![DP]\!] = \iota x[TP(\text{what Sebastian read})(x)(w)]$

$= \iota x[\text{Sebastian read x in w}]$

In either case, D^0 is an unpronounced element which semantically type-shifts a CP denoting a set to a maximal element. While this does the necessary semantic work, the

[9]Some evidence that the derivation from a question might be desirable comes from Tlingit FRs, where the Wh word comes with an overt Question operator; these seem to specifically be semantically equivalent to Wh-ever FRs (Seth Cable, pc.):

(i) [**daa sá** agawdzi.ée] xaxáa nooch.
what Q he.cooks I.eat.it
‘Whatever he cooks, I eat (it).’

Whether or not this is equivalent syntactically and semantically to English FRs would require much more in depth study, but prima facie it could be seen as an instance of an FR being derived directly from the syntax of Wh questions specifically.

syntax in (61) on its own does not derive *Wh-NP, or any other difference in the internal syntax of FRs versus questions; the syntax of the CP is unchanged, and the Wh expression is in the same position as it would be in any other Wh-CP. However, integrating the idea that there is a D^0 present in the syntax which semantically type-shifts the construction with the basic syntactic assumptions of Cecchetto and Donati (2010, 2015) can address the issue of interpretability in the latter, while still embracing the idea that the Wh word in an FR is the head of the entire DP.

Two observations about English inform how this could be implemented: that English definite descriptions generally require D^0 to be overtly realized (63), and that left branch extraction is barred in English (64).

(63) a. *Sebastian likes book

b. Sebastian likes books (≠ Sebastian likes**the** books)

(64) a. *What did Charles read [*t* book]?

b. *Whose did Charles read [*t* book]?

c. *The boy whose Charles read [*t* book]

In the case of simple nominals like those in (63), the requirement of overt phonological realization of definite D^0 may be satisfied with the determiner *the*; in its absence, a generic interpretation may be given to a bare plural, yielding (63b). Otherwise, the lack of overt determiner is simply ungrammatical (63a). In (64), Wh words move away from their NP complements, which is licit in some languages but bad for all Wh constructions in English.

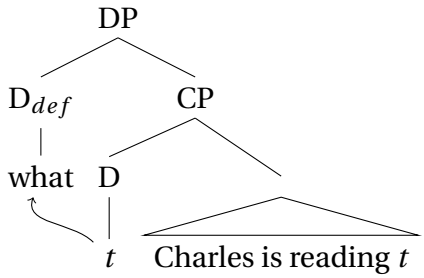
Taking these things together with the idea that FRs involve Wh movement, are definite descriptions, and must involve **Head** movement of the Wh expression, we can derive *Wh-NP simply: an FR involves the definite D^0 which selects and type-shifts a CP (as in 62), and the requirement that definite D^0 be overtly pronounced is satisfied by moving

the Wh word to D^0. Since movement of the Wh word without its restrictor is barred in English, and in general a Phrase cannot move to adjoin to a Head, Wh-NP cannot occur without either failing to pronounce definite D^0 or performing an illicit movement operation.

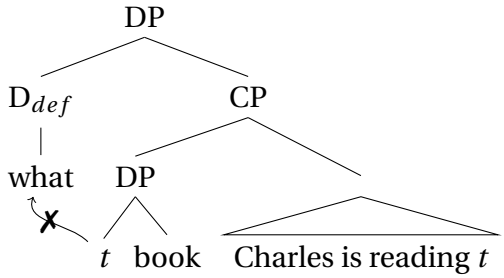
The trees in (65) sketch out what this would look like. The crucial idea here is that the Wh word must move to fill the definiteness operator (65a). In the presence of Wh-NP, movement of the Wh word would involve left branch movement which is unavailable in English, and thus generates the ungrammatical representation in (65b). Moving the entire Wh expression is not a possible syntactic operation as it involves adjoining a Phrase to a Head (65c).

(65) Free Relative Clauses with external Definiteness operator

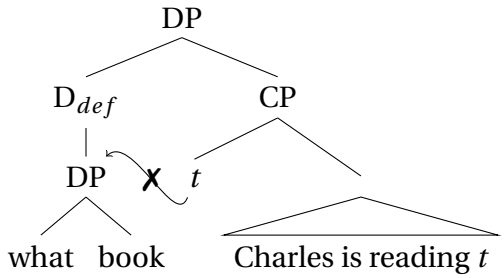
a. Licit structure



b. *Wh-NP: left branch movement



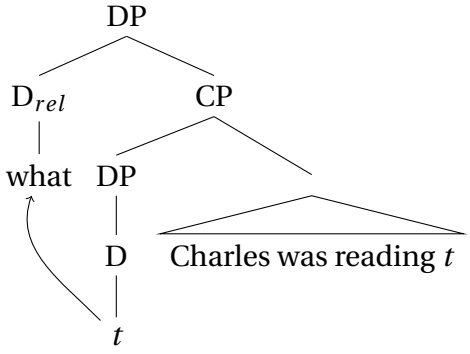
c. *Wh-NP: impossible Head Movement



This analysis takes advantage of the intuition of C&D and Donati (2006) that *Wh-NP arises because the Wh word in FRs is the categorial label of a DP, and thus must be a Head, while clarifying the semantic relationship between the two constructions as well as not requiring the assumption that a moved object may probe the XP from which it originates, both desirable outcomes. The rest falls out from the English-specific facts that D^0 must be pronounced if it is definite, and that Left Branch movement is barred.

One might intuitively think that this is a strange sort of movement, as *what* is apparently moving as both a Head and a Phrase (violating the Chain Uniformity, per Chomsky 1995); however, this is easily dealt with in a way that does not create the same problems: nothing prevents movement of a Head out of a Phrase (indeed, this is more or less the definition of Head movement). Further, while the version of Wh head movement proposed in Donati (2006) violates at least some versions of the Head Movement Constraint (Travis 1984, Donati 2006:27), the type of movement here does not, as no intervening Heads are passed over.

(66) Movement out of DP



While the DP containing *what* has moved to a Phrasal position, a sub-part of it has moved to a Head, so the relevant chains of movement are still uniform; further, while the Head movement of *what* passes several nodes, its path does not pass over any intervening Heads, so the Head Movement Constraint is not itself violated.

We might still worry that to avoid violating Chain Uniformity we are constructing a non-binary tree with the trivial DP, each of these concerns is in any case only relative to some other particular theoretical assumptions; picking out which of these several hypotheses and constraints are the right ones is non-trivial, and would rely on a much wider amount of empirical and theoretical observation.

2.2.1 Other Complex Wh

The *Wh-NP facts described here are of course only a single instance of the restrictions of certain Wh expressions being barred in FRs. We can look at a few of these and speculate, in the current framework, on what might account for the distributional facts. Two of these seem to fall out from fairly simple facts about the morphology of Wh words; another class takes slightly more in-depth analysis but do not seem to create a real problem.

An assumption I will make for all of these cases is that *what* is essentially a “default" Wh word, the spellout of a Wh morpheme where no other features are present. This idea is supported by the fact that *what* or an equivalent is typically used in Wh “scope marking" constructions cross-linguistically (McDaniel 1989, Dayal 1996, inter alia).[10]

[10]The idea that *what* in these cases is simply a pronunciation of the bare Wh feature is adopted specifically by Cheng (2000), who proposes that Wh scope marking is an instance where only the Wh feature from a Wh word is moved to the highest Spec in cases like (67), while the contentful Wh word is left in the lower clause. This assumption is not specifically necessary for my account here.

(67) Wh-scope marking

a. German

Was *glaubst du* *[**wann** daß sie gekommen ist]?*

what think you when that she come is

'When do you think she is coming?'

b. Hindi-Urdu

Rām ***kyā*** *sochtā hai [Sītā* ***kis-se*** *bāt karegī]?*

Ram what think is Sita who-with talk do.FUT

'Who does Ram think Sita will talk to?'

Accompanying this assumption is the assumption that any other Wh word contains some other semantic content, though this content may either be the result of *what* composes with other functional material in the syntax or a more complicated lexical entry containing multiple features (ie. *who* = [+Wh,+Animate, ...], etc.). These assumptions will play a part in the analyses below.

2.2.1.1 Who and Which

Two cases of interest, particularly because they seem so directly in opposition to the facts about simple *what* discussed here, are the badness of FRs involving *which*, and those involving *who*.

(68) a. (Looking at the gifts,) Charles asked which Sebastian brought.

b. (Looking at the gifts,) *Charles opened which Sebastian brought.

(69) a. Charles wondered who Sebastian invited to the party

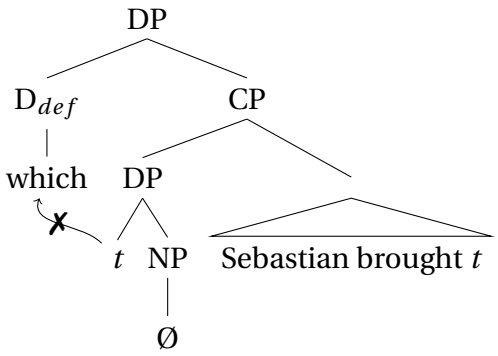
b. %Charles talked to who Sebastian invited to the party

Questions, embedded or otherwise, always allow *which* and *who*, but FRs consistently bar *which* (68b), and for a subset of speakers of English (Chomsky 2013, Patterson

and Caponigro 2016), *who* is not permissible in FRs (69b).[11] The analysis for these must take into consideration the variability of judgments on (69), but both facts can be derived from simple assumptions about these Wh words.

First we can look at *which*; since *which* is necessarily "D(iscourse)-linked", it is infelicitous without reference to a discourse-specified set; it must always mean something like 'which of the ones' (thus the need for some context like the parenthetical in 68). We can then assume that *which* necessarily carries with it an unpronounced variable NP; this thus makes the structure equivalent to the structure of Wh-NP, and blocks movement of the Wh word away from its null complement.

(70) D-linked FR structure



Since *who* FRs are acceptable to some speakers but not all, we will have to assume that there is variation in the representation of *who* across English speakers. This is another case where there are multiple possible representations of surface strings, and in this case we must assume that some speakers choose one while some choose another.

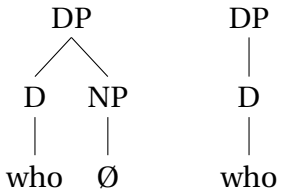
To get at what the choice would be, we can think of the sort of variable structures of pronouns suggested by Déchaine and Wiltschko (2002), where languages vary by how much structure there is to an apparently-simple pronoun. In particular, they suggest that pronouns in some languages involve a simplex Head, and others include a ϕ^0 Head

[11] Patterson and Caponigro (2016) point out that *who* FRs in English are given slightly better judgments in Object position, than in Subject position, and are improved more when they are "parallel," in the sense of being Object FRs in an Object position. They discuss a number of speculative explanations for this fact, but do not give a general account of the markedness of *who* FRs.

(as well as more complex possible pronouns); they also suggest that some languages (such as English) include pronouns of multiple types. While adopting all of Déchaine and Wiltschko's analysis is not necessary for this, something like this distinction will produce the necessary distinction here. A version of this is shown in (71), which translates the Déchaine and Wiltschko ϕP analysis into a simple D^0 vs. a DP with a null NP (bearing a person feature).

(71) Pronoun types

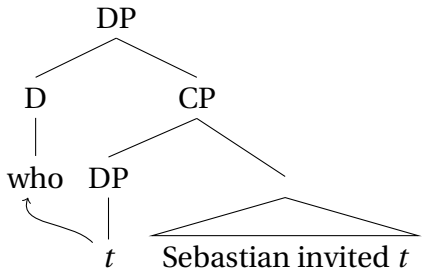
Complex *who* Simple *who*



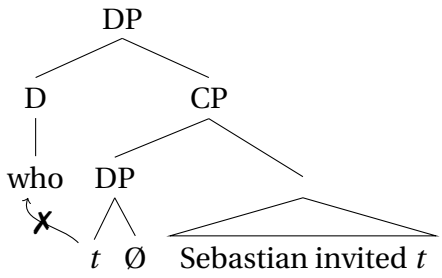
This is equivalent to the distinction between *what* and *which*: the former consists of only a single morpheme, and the latter includes a null complement, so that moving the pronounced portion would cause the same sort of Left Branch violation seen with Wh-NP. This is shown in (72).

(72) Representations of *who*

a. Simple *who* English: FR possible



b. Complex *who* English: FR impossible



Thus, the ungrammaticality of *which* and *who* in FRs can be reduced to the same source, and where *who* FRs are acceptable it can be explained by an alternate representation of *who* as a simplex element.[12] Further, the variability of representations of *who* receives some support, if one accepts elements of Déchaine and Wiltschko's proposal that English pronouns in general include multiple types (2002:419), though not necessarily entirely simple pronouns like "simple *who*" in (72).

This is an instance of exactly the sort of micro-parametric variation discussed in the previous chapter; the possible lexicons of English include two possible representations of *who*, and their behavior in the syntax varies based on this choice. Any evidence in favor of this analysis of *who* is thus evidence in favor of this lexical view of Parameters.[13]

Here we must assume that there is limited data in the input for learners to choose between them. We might assume that the marginal status of *who* FRs would be the best evidence to choose between the two, and children will have no access to such negative evidence, this would be a good reason to suspect this is true, unless other, positive evidence could be found to make this choice possible.

---

[12] The above analysis of *which* predicts that people who accept who-FRs should dislike them where they are D-linked. So, (i) should be bad. We can see something similar in the crashing badness of (ii).

(i) *Among the guests, Charles talked to who Sebastian invited

(ii) *I encourage who among you is without sin to cast the first stone

[13] It is worth noting that much of Déchaine and Wiltschko's argumentation is based on the binding properties and argument/predicate distinction of the different types; these are difficult to assess for Wh words in general, so a learner might have no good evidence to make this decision.

A third example which can be described in terms of *Wh-NP is *how much/many* FRs, which are blocked in English and are marginal in Italian (Donati 2006, Chierchia and Caponigro 2013). Note the apparent morphological distinction in (73-74) between English and Italian *how many*.

(73) English *how many*

a. I wonder **how many** (eggs) Sebastian has

b. *I ate how many (eggs) Sebastian has

(74) Italian *quanti* (Donati 2006)

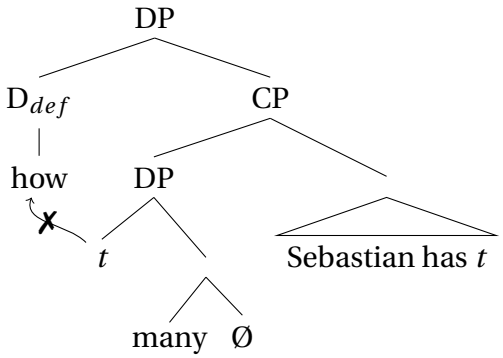
a. *Mi chiedo **quanti** biscotti hai preparato*
me wonder how.many cookies you.have made
'I wonder how many cookies you made'

b. *%Ho mangiato quanti biscotti hai preparato*
I.have eaten how.many cookies you.have made
'I ate how many (the amount of) cookies you prepared'

The unacceptability of the FR in (73) can be straightforwardly understood in terms of *Wh-NP - it involves an overt NP, creating a complex Wh expression of the same type as those discussed above. But, in English, *how many* FRs are bad without the NP as well as with it.

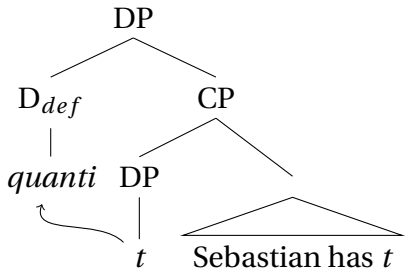
The important facts to consider here are that (a) while the Italian *quanti* FRs are marginal but accepted by some speakers, the English *how many* FRs are entirely bad; and (b) while the English quantity Wh expression is two words on the surface, the Italian equivalent is apparently a single morpheme. A syntax for English *how many* which would block these FRs is shown in (75); as above, the Wh word has a complement which it cannot move away where Left Branch movement is barred.

(75) *How-Many FR



We can imagine, then, that the same sort of variation we see in *who* FRs in English: Italian allows for the apparently simplex *quanti* to be analyzed syntactically as either bi-morphemic or mono-morphemic. Based on the sort of syntax for *who* that allowed FRs proposed above, we could speculatively propose a version of Italian syntax which allows *quanti* FRs as in (76).

(76)



The types of FRs discussed in this section have in common that surface complexity blocks Wh expressions from occurring in FRs in a language like English; in the next section I'll discuss FRs with Wh expressions where surface complexity does not generally seem to be an issue, but where other issues seem to arise.

2.2.1.2 Adjunct Wh in FRs

Another case, which has been discussed in much of the previous FR literature, is adjunct FRs. These seem to have the distribution both of DPs and PPs.

(77) a. Sebastian went [$_{PP}$ where Charles told him to go]

b. Charles checked out [$_{DP}$ where Sebastian was staying]

(78) a. Sebastian left [$_{PP}$ when Charles told him to]

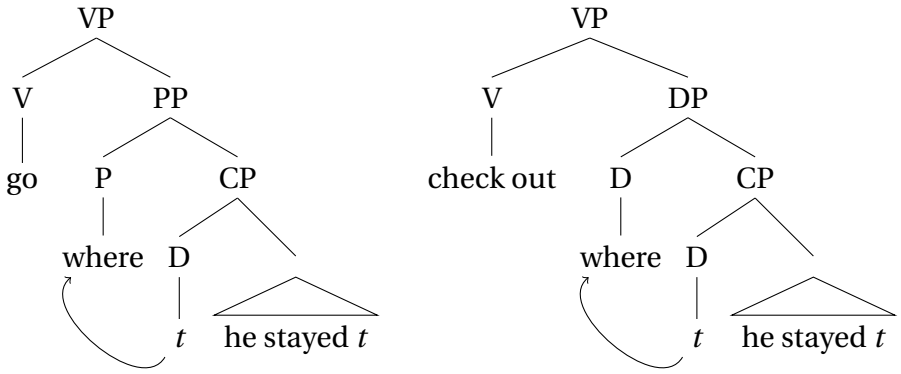
b. Sebastian waited until [$_{DP}$ when Charles was supposed to arrive]

Work by Iatridou, Anagnostopoulou, and Izvorski (1999) and Izvorski (2000) explores a syntax for these similar to the Relabel analysis of Cecchetto and Donati (2015) for FRs in general: the Wh word is of some non-D category (Adv, P, etc.), and its movement results in the entire structure being labeled as such. This is partially based on the idea that categories like P don't typically select clauses (Izvorski 2000:9), but the strings in (77-78) can correspond to non-FR strings like *after Charles told him to go*. Further, this doesn't give an obvious account of the categorial variability in these examples.

Conversely, the present analysis of argument FRs extends to these quite naturally; assuming that *when* and *where* are both simple Wh expressions, we can adopt a version of the analyses of Bresnan and Grimshaw (1978) and Larson (1987), in which the Wh expressions merge to a categorial P^0 position where necessary.[14] The resulting structures are given below.

[14] Bresnan and Grimshaw in fact assert that entire NPs move to this position, but this is neither necessary for the present purposes nor fitting with the assumptions made here about merging complex syntactic expressions with heads. Larson adopts an analysis where these involve a PP with no visible categorial head (1987:246). In §2.2.2 I discuss some ideas related to the nature of Type Shifting operators in the syntax which requires that Larson's analysis be rejected in favor of the one given here, though empirical motivations for choosing between the two for present purposes are unclear.

(79) FRs with Adjunct Wh



In the case where the resulting FR is a PP, it may not be necessary that the Wh word adjoins to P^0. There are apparent cases in English where P^0 may be silent, but these are infrequent, and seem to all be derived from having overt prepositions (as in *go upstairs* or *live downtown*), so we might treat these as exceptional. If we generalize that P^0 in English generally must be pronounced, we would have the structure in (79); if not, we would have a structure in which *where* remains in the specifier of CP. For these, the two options may be equivalent, as there don't seem to be cases in locative or temporal FRs of certain Wh expressions being blocked.

2.2.1.3 The absence of *why* FRs

There seem to be special restrictions on relatives denoting reasons. the languages presented in Table 2.1 above show no languages which allow FRs headed by *why*, and these may not exist at all cross-linguistically, as noted (but not explained) by Caponigro (2003). Further, in English headed RCs, *why* as a relative pronoun is quite marked; if the head of the RC is a word like 'reason', the non-Wh relative pronoun *that* is preferred. This does not seem to be the case for other adjunct relatives, which allow not only FRs but also for Wh relative pronouns, as the contrast in (80a-b) shows.

(80) Adjunct Relative Pronouns

a. Sebastian's absence was the reason {that, %why} Charles left

b. Sebastian's house was the place {that, where} Charles first drank

Note also that, to the extent that it is available, a relative clause with *why* allows something like a sluice, which is quite bad in all other types of relatives in English, including other relatives headed by 'reason'. In (81), eliding the clausal element following the noun is blocked entirely if the relative pronoun *that* is used, and is still ungrammatical with a Wh relative pronoun except for *why*.

(81) Sebastian was reading, and...

a. ... that's the reason {why, *that} ~~Sebastian was reading~~.

b. ... *that's the place {where, that} ~~Sebastian was reading~~.

c. ... *that's the thing {which, that} ~~Sebastian was reading~~.

We can think of this in selectional terms: Relative-like things, denoting reasons (which are inherently epistemic in nature) must be CPs. Since CPs in general allow sluices, we can get sentences like those in (82), which are semantically more or less identical to those in (81). However, there is no immediately apparent reason for the crashing badness of (82c).

(82) a. That's why

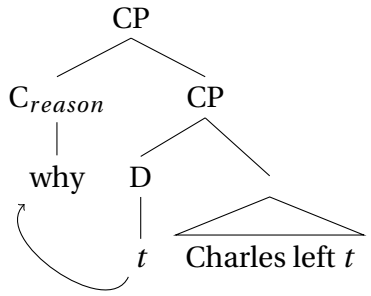
b. That's where

c. *That's which

Given all of this, we can think of the badness of FRs with *why* in selectional terms. While argument FRs are DPs, and other adjunct FRs are PPs, *why* FRs would necessarily

be adjunct CPs, and by the analysis adopted here would have to involve a complementizer, denoting a reason (equivalent to *because*), which selects a Question-CP. This would yield a representation like (83).

(83) *Why* FR



This syntax is plausible in the framework adopted here: an FR is generated by embedding of a CP, with a simplex Wh word moving from a specifier position to adjoin to the selecting head. However, the relevant type of selection does not seem to occur in English - *because* does not take question complements (84) - so the selectional environment which would produce (83) does not occur.

(84) *Sebastian left [because (of) whether Charles wanted him to]

Further support for the idea that the blocking of *why* FRs is a selectional or semantic fact comes from languages with non-Wh FRs, like Hindi-Urdu, which lack correlative pronouns which correspond to reason Wh words (85); while Hindi-Urdu generally has Wh-Relative pairs which are morphologically related, there is no such match for *kyō* 'why'.[15]

[15]In point of fact, there is a morphological mate, *jyō*, but this means 'how' or 'the way', rather than 'the reason' (less commonly used than another manner pronoun *jaise*, which is the morphological mate of the question manner word *kaise*). Thus while there is an apparent morphological exception to the generalization, the semantic generalization here stands.

(85) Hindi-Urdu Wh-relative pairs

a. *Āmir* ***kab*** *āegā*

Amir when will.come
'When will Amir come?'

b. ***Jab*** *Nādyā us-se kahegī, Āmir āegā*

When Nadya him-with will.tell Amir will.come
'Amir will come when Nadya tells him'

There are pairs like (85a-b) for most Hindi-Urdu Wh words, but there is no equivalent pair for the reason Wh word, and such a meaning would have to involve something like *jo kāran* 'which reason'.

2.2.1.4 PPs, pied-piping, and stranding

English also bars FRs with pied-piped PPs (86); this sort of pied-piping is available, though often dispreferred, in matrix questions, though is fairly marked in embedded questions. In all cases, prepositions may be "stranded" in their base position instead (87).

(86) PP pied-piping

a. [To what] did Sebastian change the subject?

b. Sebastian wondered [to what] he should change the subject.

c. *Charles was an expert in [to what] Sebastian changed the subject

(87) Preposition stranding

a. What did Sebastian change the subject [to *t*].

b. Sebastian wondered what he should change the subject [to *t*].

c. Charles was an expert in what Sebastian changed the subject [to *t*].

The syntax assumed here generally bars sentences like (86c), as the Wh word does not move to D^0. To fulfill the requirement that D^0 be realized by *what* in these cases, we would get strings like those in (88), which are blocked in matrix and embedded questions as well.

(88) a. *What to did Sebastian change the subject?

b. *Sebastian wondered what to he should change the subject

c. *Charles has a strong opinion about what to Sebastian changed the subject

These sorts of patterns are only allowed in “Swiping" patterns; the reason for this exception is unclear.

(89) Sebastian will change the subject, but I don’t know what to

In any case it is clear that the badness of FRs like (88c) comes from the same source as the badness of questions like (88a-b).

2.2.2 Some consequences: Overtness and interfaces

It should be noted that this analysis makes a specific assumption about the mapping between syntax and phonology for particular lexical items, namely that some given item may contain a feature which requires that it is pronounced (or perhaps bars it from being pronounced). Phenomena like this are seen in a number of cases, ranging from “EPP" phenomena requiring particular specifiers positions be filled (usually involving phrasal movement), and *do* insertion phenomena fulfilling a requirement that T^0 be pronounced in certain contexts.

The Move-to-D^0 analysis for FRs proposed in this section is more akin to the *do* insertion type of overtness requirement, as it involves phonological content given to a particular head. We could express this requirement in terms of a feature on functional heads, [±Overt], such that any head which is [+Overt] must have a phonological exponent.

There are two strategies seen in English nominals to satisfy [+Overt]: either having a determiner with its own associated phonological content, like *the*, or moving of another head to D^0, as in FRs. This pair of strategies also seems to be seen in the verbal/inflectional domain in English: cases of *do* insertion satisfy a requirement that T^0 (or some similar head) be pronounced in certain conditions, and Auxiliary movement satisfies a requirement that C^0 be pronounced.

(90) a. ***Do* support**:

Charles [$_T$ *(did)] not read the book.

b. **Auxiliary movement**:

[$_C$ Will] Charles *t* read the book?

These sorts of phenomena generally support the notion that both types of solutions for [+Overt] could be seen in various places in the grammar.

This analysis of FRs also has consequences for the syntax-semantics interface. The analysis takes advantage of the semantic assumption of Caponigro (2003 et seq) based on Partee (1987) and Chierchia's (1998) treatments of variable semantics of nominals, among others, that things which do not have the outward appearance of DPs (FRs, which resemble CPs, and bare plurals, which resemble NPs) may appear in DP positions by way of a type shifting operator. Other work on bare nominals (such as Cheng and Sybesma 1999) assumes that apparently bare NPs must always involve some additional *syntactic* structure. The syntax proposed here assumes that the type shifting operation for FRs correlates not just to a semantic type shifter, but a categorial D^0, which allows the CP-like string to occur in a nominal position.

As a way to limit the distribution of type shifting operators, Chierchia (1998) proposes the Blocking Principle, which restricts type shifting in the nominal domain to those cases where there is no Determiner available which could provide the relevant semantic function. His formulation of this principle is as in (91):

(91) **Blocking Principle** (Chierchia 1998:360):

For any type shifting operation τ and any X, *τ(X) if there is a determiner D such that for any set X in its domain, D(X) = τ(X)

That is, A type shifter τ cannot be a part of a (semantic) derivation if there is a determiner which could yield the same semantic output. We could restate this in a more general way, which may carry over to other conditions where there must be a "last resort" semantic operation, and in a way which makes the relationship to the syntax more clear:

(92) Wherever possible, any semantic type shift operator must carry a syntactic category label.

The thrust of (92), as with Chierchia's Blocking Principle, is that there is a preference for semantic operators to be lexical items. The reformulation in (92) has two advantages: one is that it makes specific that this ought to arise because the narrow syntax prefers that categorial selectional requirements of various positions be met, in addition to any interpretive requirements imposed at LF. The last resort addition at LF of purely semantic type shifters is strongly dispreferred, and set aside for instances where the grammar has no access to lexical items which fill the requisite semantic role, or where c-selection is not at-issue.[16] The second advantage of (92) over the Blocking Principle is that it deals with cases like FRs, where there are clear syntactic consequences of the semantic type shift, as discussed here.

---

[16]An example of this may be conjunctions, which seem to need to be of "flexible type" (Partee and Rooth 1983, Heim and Kratzer 1998). Aside from semantic issues, conjunction of predicates of the same type may still be sensitive to syntactic categorial issues, as suggested by the questionable acceptability of (i).

(i) ?Julia is ambitious and a Catholic

Both predicative nominals and adjectives are apparently of type <e,t>, but their conjunction seems anomalous. This suggests that type shifting of conjunctions may be sensitive to both semantic and syntactic type.

Something like (92) may in general be desirable for syntactic analysis in general; analyses of various phenomena rely on semantic effects being manipulated by the presence and absence of functional heads which may be phonologically null. This ranges from argument structure operators like v (Kratzer 1996), to projections in the left periphery affecting the type of clauses (Rizzi 1997), to other operations within the nominal domain (Cheng and Sybesma 1999). In each of these cases, it is assumed that there is an interaction between syntax and interpretation which is mediated by syntactic material which contains categorial features. This sort of assumption becomes more tenable in general in a system that prefers, if not requires, that semantic operators have syntactic labels.

In the present case, English simple FRs are typically restricted to positions which select DPs, and (by hypothesis) null D^0 is allowed, and so the ι type shifter is part of the syntax as a D Head, by which the string can be both labeled a DP and interpreted as a definite description.

2.2.3 What [+Overt] is

An obvious question raised by this analysis is what it means for a functional head to require Overtness. I define it here in terms of a feature associated with particular functional heads: Definite D^0, among other heads, has the feature, and so a PF which does not have a pronounced correspondent for definite D^0 is ungrammatical. But how and where is this requirement satisfied?

Given that the feature is defined in terms of its *phonological* consequences, and does not itself affect meaning, it is reasonable to say that its success or failure must be evaluated at PF. The only other option would be to say that the narrow syntax itself knows when a lexical item has phonological content or not, which we may or may not want to say.

There are a couple of possible ways through this issue. One is to wholly adopt the notion that Head movement is not done before PF spell out (Chomsky 1995); this allows for [+Overt] to only make its presence known once the issue of pronounceable material in general is invoked by the grammar. The only real issue here is whether PF movement is sensitive to all the same syntactic constraints as the narrow syntax (ie., restrictions on Left Branch movement). This is not an overly onerous prediction, although we might think of Head Movement as more strictly local than phrasal movement, in that it tends to obey the Head Movement Constraint.

The other option is to say that Head Movement happens in the narrow syntax, and the syntax of English may generate structures with and without Head Movement of *what*. It is then just the case that structures without the movement would crash at PF, and so speakers would reject sentences which require null D^0 (bare singular NPs, Wh-NP in FRs). While choosing between these two options isn't easy, in the following chapter (§3.4.1) I will discuss some issues in acquisition that might lead us to prefer this latter option.

2.3 Variation: Left Branch Movement and Overtness

The syntax of English FRs presented here derives effects like *Wh-NP from a few general grammatical principles (properties of selection and labeling, properties of compositionality), and particular properties of Wh movement and the syntax of DPs which are subject to cross-linguistic variation. Languages which differ from English as to whether determiners and specifiers may move away from their sisters, or whether any particular Determiner must be overt, should allow for different constraints on which Wh expressions may be involved in FRs. And we do indeed see that this is the case. Below I show a few such examples.

2.3.1 Movement

Recall that part of the analysis of English FRs makes use of the fact that “Left Branch" movement in English is generally barred in all Ā movement, shown again in (93) with Wh questions, headed relatives, and clefts; this applies to both movement of Wh possessive from its possessum and movement of *what* from its restrictor.

(93) a. *Whose did you read [*t* book]?

b. *What did you read [*t* book]?

c. *This is the man whose Charles read [*t* book]

d. *It’s Sebastian whose Charles read [*t* book]

However, these sorts of patterns are not blocked in all languages. Hindi-Urdu and some Slavic languages allow exactly this sort of movement (Bošković 2005, Izvorski 2000). The syntax for FRs sketched above makes predictions about what sort of variation is available: languages which allow Wh-NP should either allow Left Branch movement or not require definite D to be filled (or both). We see this in a couple ways: Russian, which has Wh FRs (94), and Hindi-Urdu (95), which has non-Wh correlatives, both allow left branch movement and do not bar complex relativized expressions in FRs.

(94) Wh-NP FRs in Russian (Izvorski 2000:47)

Mne est’ [{kakuju, č’ju} knigu čitat’]

me.DAT is what, whose book read.INF

‘There is someone’s/some kind of book I can read’

(95) Rel-NP Correlatives in Hindi-Urdu (Bhatt 2003)

[jo CD sale-par hai], Rām voh CD kharīdegā

RP cd on.sale is Ram that cd will.buy

‘The CD that is on sale, Ram will by that CD’

An exceptionally striking version of this variation is seen in Italian Sign Language (LIS; Branchini and Donati 2009, Cecchetto and Donati 2015), which has FRs which optionally allow movement of the relative pronoun (glossed *PE*) away from the Head of a relative, shown in (96).

(96) LIS relatives (Cecchetto and Donati 2015:81)

a. *[PIERO CONTRACT SIGN DONE PE] GIANNI FORGET*

Piero Contract sign ASP PE Gianni forget
'Gianni forgot what contract Piero signed'

b. *[PIERO [PE CONTRACT] SIGN DONE] GIANNI FORGET*

Piero PE contract sign ASP Gianni forget
'Gianni forgot the contract Piero signed'

LIS seems to be an example of an FR being formed by "Wh Head Movement" in the sense of Donati (2006), but with an NP restrictor, *and* with left branch movement, exactly the sort of FR predicted to exist by this type of analysis.

(97) Movement of *PE* in LIS

[$_{DP}$ [$_{CP}$ PIERO [$_{DP}$ *t* CONTRACT] SIGN DONE] [$_{D}$ PE]]

In LIS, the relative morpheme is allowed to either stay clause internal or move to the periphery of the relative, appearing to act as the head of the DP, and the presence of an NP does not block this movement. So, the movement that would be equivalent to movement of *what* moving away from an NP in an English FR is allowed, and thus FRs are allowed to have an NP.

2.3.2 FRs and Definite Determiners

Another point of variation which is observed which provides strong support for the approach taken here comes from FRs in languages which may include an overt D^0, including Wolof, Spanish (Caponigro 2002), and Bulgarian (Izvorski 2000). In Wolof, the

definite determiner *ki* optionally follows the clausal portion, in the same position it occurs in a simple nominal (Caponigro 2002). In Spanish, we see the non-canonical determiner *lo* which is also used in DPs lacking an overt NP.

(98) Wolof FR (Caponigro 2002)

door-naa $[_{DP}$ *ki nga begg (ki)]*

hit-1SG REL you love DEF

'I hit who/the one you love'

(99) Spanish *lo* (Caponigro 2002)

a. $[_{DP}$ *lo que tu crees] no es cierto*

DEF REL you believe not is certain

'The things you believe are not certain'

b. $[_{DP}$ *lo bueno Ø]*

DEF good

'The good (things)'

These are specific types of language we expect to see in view of the assumptions that (i) FRs are type-shifted Wh expressions and (ii) this type-shifting involves CPs being relabeled as DPs. In particular, these example seem to be specific proof-of-concept for the principle in (92) above, as they are instances where a non-nominal category (a CP) is semantically type shifted and overtly labeled as a DP, both make use of an otherwise available D, *ki* in Wolof and *lo* in Spanish.

The strategy employed in this analysis of something Moving to D^0 to express definiteness has been observed in other languages has been observed in other languages as well: Cheng, Heycock, and Zamparelli (2017) give examples from Italian (from Longobardi 1994) and Bangla where a Noun moves within a DP either in a definite case (Bangla) or to show type of definiteness (Italian).

(100) Bangla definiteness

a. *ek Ta chabi*

one CL key

'A key'

b. *chabi Ta*

key CL

'The key'

(101) Italian "rigid designator"

a. *Carlo voleva la sua casa*

Carlo wanted the his house

'Carlo wanted a private house'

b. *Carlo voleva casa sua*

Carlo wanted house his

'Carlo wanted his own house'

In (100), moving the noun past the Classifier to where the numeral had been gives a definite reading. In Italian, moving the noun past the possessor into the position of the definite article gives a "stronger" definite reading. Both of these are cases where material other than D^0 (in this case, N^0) moves to phonologically realize a (type of) definiteness morpheme, exactly what I have argued happens in the case of English FRs.

Another case which may involve other move-to-fill operations in the nominal domain appears the marking of plurality in different languages; in Welsh and Turkish, the plural morpheme does not appear in enumerated nominals (Cheng and Sybesma 1999).

(102) a Turkish

kiz-lar 'girls'

*uc kiz-(*lar)* 'three girls'

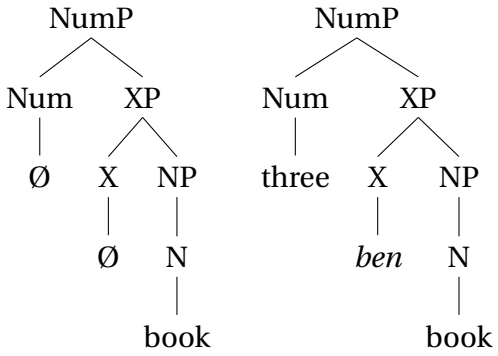
b Welsh

merch-ed 'girls'

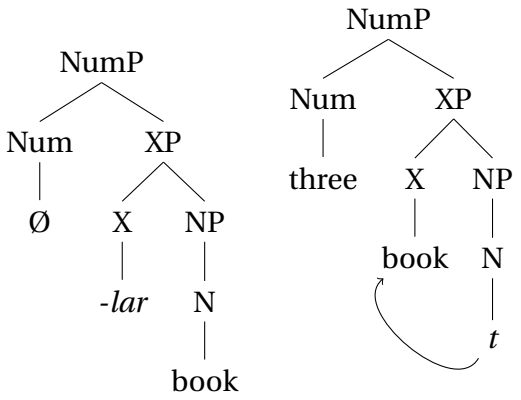
*tri merch-(*ed)* 'three girls'

A generalization by Greenberg (1963, cited in Cheng and Sybesma 1999:517) is that languages in general must express countable plurality by some morphological means. In languages like Cantonese, nominal classifiers do this; Turkish and Welsh are prima facie exceptions to this, but this may be explained in the present discussion in terms of a similar solution to moving to D^0 to realize definiteness; in simple Plurals, the position associated with plurality must be filled when appearing with a numeral. In Mandarin, this is satisfied by the presence of a numeral classifier (103); in the case of Turkish, this is (by hypothesis) satisfied by movement of N to the Mead associated with plurality (104). I will remain agnostic about what this head is and simply refer to it as X.

(103) *(san ben) shu* '(three CL) books'



(104) *uc-(lar) kiz* '(three) books'



All of these data show rich crosslinguistic variation with regard to apparent movement within the nominal domain which seems to work to fulfill various overtness requirements. This sort of variation fits into the general framework proposed here to account for syntactic facts about FRs in English and other languages: variation in what can and has to move to fulfill certain requirements interacts with the form of the morphemes associated with relatives (relative pronouns and Wh words) to restrict the types of FRs which may exist in a given language.

2.3.3 Variation and the acquisition problem

The sorts of cross-linguistic variation observed in this section regarding the syntax of Wh movement and nominal syntax, and its consequences for variation in the possible surface forms of FRs, lends support to the way syntax which derives *Wh-NP in English which I have proposed in this chapter. However, they also lead to a complex acquisition problem for FRs: *Wh-NP can be learned only when all of the involved parts are acquired. A child learning English who has not learned that Left Branch movement is blocked, or that D^0 must always be overtly filled in definite descriptions, has a grammar which should allow Wh-NP in FRs. In the next chapter I will present data that there is indeed a point in the acquisition path for English where this is the case, and that specifically the acquisition of Overtness requirements in English can derive this.

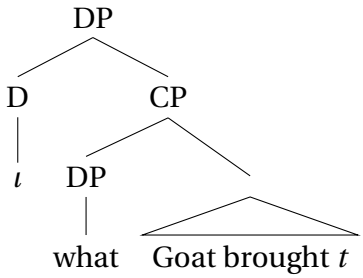
CHAPTER 3

THE ACQUISITION OF FR SYNTAX

In Chapter 2 I presented an assortment of data leading to an analysis deriving surface properties of English FRs, in particular the barring of Wh words with nominal complements (*Wh-NP), from the interaction of Wh movement and the syntax of definite DPs in English. Here I will examine FRs in first language acquisition, looking at early production and differences between children's and adults' interpretations of FRs. I show that children know the basics of FR syntax (that they are nominals derived out of Wh clauses) quite early, but that the syntax leading to *Wh-NP in adult English is not present until late in development. I show that this can be understood in terms of a theory of parameters and parameter setting, especially within the nominal domain, based on ideas discussed in the previous chapters.

The first section discusses early production of FRs and other Relatives. The second discusses the phenomenon of medial Wh answers in acquisition, and an experiment which demonstrates children have knowledge of the semantic distinction between FRs and Wh questions. The third section discusses an experiment which shows children do not know that Wh-NP disambiguates between FRs and Wh questions. The fourth section discusses how these data demonstrate that children's representations of FRs involves a similar derivation to that of adults, but does not involve adjunction of the Wh word to D^0, as shown in (105).

(105) Early FR syntax



The fifth section relates these findings to other issues in the acquisition of the syntax of the nominal domain in English and cross-linguistically.

3.1 Early knowledge of FRs and their pieces

Some previous findings in acquisition suggest that from a very early age children have adultlike knowledge of the form of Free Relatives. Children acquiring English begin producing adultlike FRs as early as 3;6 (Flynn and Lust 1980), their first form of relativization in production (106).

(106) Cookie Monster eats [what Big Bird pushes]

Given the analysis of FRs developed in Chapter 2, we can describe the acquisition problem for FRs as involving learning (i) the properties of Wh movement in English (conditions on Pied Piping, the constraint against Left Branch movement, etc.), (ii) the properties of Definite DPs (in particular that D^0 must contain phonological content if it is definite), and (iii) that FRs are definite DPs derived from Wh clauses. Here I'll present data from previous research which shows that children acquire (i) and (iii) quite early.

3.1.1 Complex Wh in production

While children seem to produce Wh-NP sequences from fairly early, observable instances of illicit Wh-NP seem rare-to-nonexistent. Data from the Adam and Sarah corpora (Brown 1974, MacWhinney 2000) shows children producing Wh-NP sequences in

embedded questions from around the same age as standard FRs (107). Instances of *how many* and *how much* appear at a similar time range.

(107) Which-NP

I don't know which day we have these Julia, 3;5

(108) How much/many

a. How many monkeys? Sarah, 3;2

b. How much it is all dese meats? Adam, 4;0

Questions with *whose* create a slightly more complicated picture; Gavruseva and Thornton (2001) provide apparent evidence of movement of possessors out of complex Wh expressions. In a production experiment, they elicited a number of examples of the type in (109a) from children ages 4;5-6;0. They also elicited some examples of the types in (b-c), but these are much more rare, and the former is almost entirely restricted to a single child.

(109) Elicited Left Branch movement (Gavruseva and Thornton 2001:250-251)

a. Who do you think [Ø's coin] is in the box?

b. Whose do you think [Ø's hat] is on the skis?

c. Whose do you think [Ø ball] is on the skis?

Two things are important to note about their findings: first, the instances of non-pied piping of the types in (109) are entirely restricted to long distance questions of exactly this type (out of complements of *think*). Second, among these, the plurality of produced sentences (46%) *did* involve pied-piping. The former fact is striking because this differentiates these sorts of non-adultlike English sentences from instances of Left Branch extraction in other languages, like Russian and Serbo-Croatian (Gavruseva and Thornton 2015, Bošković 2005).

(110) Russian Left Branch movement (Gavruseva and Thornton 2015)

Kakuju$_i$ Tania prochitala [t$_i$ knigu]?
which Tania read book
'Which book did Tania read?'

(111) Serbo-Croatian Left Branch movement (Bošković 2005)

Kakva$_i$ si kupio [t$_i$ kola]?
what.kind are bought car
'What kind of car did you buy?'

Gavruseva and Thornton (2015:234) suggest that French actually does show a pattern more like this Child English data, in that something like Left Branch movement is allowed only for a limited set of Wh words, namely quantifiers and Wh possessives. However, the syntax is distinct even on the surface from the type of LB movement in Slavic (and other languages) in a couple of ways: in the case of possessors, the movement is in fact from the right of the possessum, not the left; second, in the case of the quantity Wh word *combien*, it moves away from a PP rather than an NP. It is also, like Russian and Serbo-Croatian, not limited to long-distance movement.

(112) Quasi Left-Branch movement in French

a. *combien est-ce que tous les enfants ont lus [t de livres]?*
how.many Q all the children have read of books
'How many books have all the children read?'

b. *[De qui] a-t-il pris [le chapeau t]?*
of who has.he taken the hat
'Whose hat has he taken?'

Neither of these is movement of a Wh word from an NP complement, and in fact both have the character of movement of or away from an adjunct.[1]

Below I will give additional argumentation that the data from Gavruseva and Thornton (2015) does not represent a Left-Branch stage in English development.

3.1.2 Early production of FRs

Addressing the question of why FRs are produced earlier than headed RCs, Guasti and Shlonsky (1995) make the case that it is because the representation of FRs (or at least, a possible representation) is clearer from the surface than that of headed RCs. The argument is essentially as follows: while Wh movement is overtly present in FRs, a headed RC underdetermines the analysis, and could be analyzed as involving overt movement of the head, a null operator, copy-deletion movement, or any other possible derivation which would produce the target representation. Wh movement-derived RCs are only obvious in the RCs with pied piping; so, while determining a Wh-movement-based analysis of FRs is quite easy for children, determining an analysis for headed RCs is relies on a very particular kind of RC (one involving pied piping).

While PP pied piping is produced early in questions, at least for learners of French, where it is mandatory (Guasti and Shlonsky 1995), it is encountered extremely rarely in RCs (Labelle 1990); further, children learning English (where preposition pied piping is not mandatory) show poor understanding of sentences which use it (Guasti and Shlonsky 1995); thus, a piece of input which may be necessary to determining the analysis of headed RCs is missing, or when present not fully understood, by learners of English and French.

---

[1] In fact, an English sentence resembling (112a) is considerably better than canonical Left Branch movement, though still marginal. This is especially with *how much*:

(i) ?How much have the children read of War and Peace?

(ii) *How many did the children read books?

Labelle (1990, 1996) points out that in headed RCs, children acquiring French overproduce resumptive pronouns (113a) as well as producing doubled heads of RCs (b)[2], but do not produce anything of this kind in FRs (1996:77fn). They also produce RCs without overt heads, as an apparent escape from the need for pied piping (114).

(113) a. *celle là que le papa lui montre un dessin*
this there which the papa CL shows a drawing
= This one that papa shows a drawing (to him)

b. *sur [la boîte que la petite fille est débout sur la boite]*
on the box that the little girl is standing on the box
= On the box that the girl is standing on the box

(114) *Sur [à la petite fille Ø que le monsieur il montre un dessin]*
on to the little girl that the man he shows a drawing
= On the little girl (to whom) the man shows a drawing

All of these can be generalized as instances of avoiding canonical pied piping, and avoiding overt head-gap dependencies in the absence of a Wh word. Guasti and Shlonsky explain this in terms of an inability to use null operators before a certain age (for maturational reasons): FRs allow children to use an overt Wh operator to bind a gap, whereas adultlike headed RCs require an operator which children do not have access to.[3].

Possessive pied piping is produced and experienced early, though rarely. Below are the two examples in the CHILDES corpus, along with a sample of the input. Again, no instances of illicit pied-piping or non-pied-piping are observed.

---

[2]While Guasti and Shlonsky (1995) point out that this is extremely typologically rare, a type of this exists in English, though in a type of legalese or literary register which is inaccessible to most children. See the discussion in the appendix to Chapter 2. That said, even when this is taken into account, the extreme rarity of this type of construction cross-linguistically is of note, especially as it informs Guasti and Shlonsky's maturational argument.

[3]But, they do assume children have access to a null D^0 (Guasti and Shlonsky 1995:273); see the analysis for FRs in Chapter 2.

(115) Production

a. and there's a boy whose name tracey and there's a boy whose name Carl (4;6) (Hall et al. 1984)

b. Momma know whose that surprise was? (4;4) (MacWhinney 2000)

(116) Adult production

a. Mother: Is that the doll whose tummy hurts? (2;4) (Bates et al. 1988)

b. Investigator: we've friends whose wee boy is called Scott. (3;1) (Henry 1995)

To summarize the findings here: children's production of FRs seems to show early mastery, particularly in comparison to headed relatives, where children are resistant to prepositional pied piping, and commit errors like illicit resumption and, in rare cases, head doubling, which are not found in the input. This can be understood either in terms of the relative ease of assimilating the syntax of FRs with other Wh movement (due to the overt Wh word) or the maturational path which does not make the necessary syntactic components for headed RCs availability until fairly late.

3.1.3 Exhaustive interpretations

To determine whether or not children know the syntactic properties which distinguish FRs from Questions, we must first establish that children understand that the two constructions have two different semantic values. As FRs denote definite descriptions rather than questions (as discussed in the previous chapter), showing similarity between children's interpretation of FRs and other definite descriptions would be evidence that Wh strings can be both FRs and Questions - that is, both DPs and CPs.

A type of evidence in favor of this comes from comparing studies of children's knowledge of exhaustivity in different constructions. Many studies have shown that children up to around age 6 or 7 have difficulty interpreting definite descriptions with the as exhaustive (Maratsos 1974, Karmiloff-Smith 1979, inter alia). More recent studies

show similar performance for FRs, specifically that the age of onset for exhaustive FRs is roughly the same as that for prototypical definite DPs (Modyanova and Wexler 2007, Caponigro et al 2012); this is illustrated in (117), based on the data from an act-out task in Caponigro et al (2012).

(117) **Instruction:** Can you give me {what's, the things} in the bucket?

Adults: give all the things.

Children: give some of the things.

Conversely, exhaustivity in Wh questions does not seem so delayed, and children begin giving exhaustive responses to Wh questions closer to 4 (de Villiers and Roeper 2011).

Further, the manner of errors in exhaustivity also differs between questions and definite descriptions; de Villiers and Roeper note that errors in exhaustive question responses are typically singleton responses rather than "insufficient plurals" (say, two out of four possible answers); this does not seem to be mirrored in performance with definite descriptions.

(118) Who is wearing a sweater?

Adults: Julia, Celia, and Phoebe

Children: Julia

Not: Julia and Celia

Based on the different time frame and nature of these behaviors, we can conclude that children know that the exhaustivity requirement of FRs tracks not with the Wh word itself but with the semantics of definite descriptions, and that FRs are the latter.

Chapter 4 contains further examination of the interpretation of exhaustiveness for different constructions by children and adults, especially regarding Wh-ever FRs; for

the present discussion the distinction between simple FRs and Wh questions is what is important. In the next two section I discuss new experimental data which looks at other ways in which children show knowledge of semantic (if not always syntactic) knowledge of the difference between FRs and questions.

3.2 FRs and Medial Answers

A variety of different studies over the years have found that, in certain syntactic and pragmatic contexts, children will answer not the matrix Wh word, but rather the embedded, “indirect" question word. This is seen both as a phenomenon of interpretation in experimental contexts (de Villiers and Roeper 2011, and citations therein), and of natural production (Thornton 1991). This phenomenon is illustrated in (119): adult consistently will answer the matrix *how* question for this sort of sentence, while children will sometimes answer the embedded *what* question.

(119) Medial Wh answers

How did she say what she bought?

Matrix response: She whispered it

Embedded response: A cake

While the exact etiology of this behavior is grounds for debate, I will argue here that it is in some sense due to the syntactic representation children give to Wh questions in particular, and present experimental evidence that shows a distinction in children’s behavior in this regard with FRs as opposed to questions. The latter provides evidence that children make use of syntactic cues to semantically distinguish the different constructions.

3.2.1 Medial answers are a grammatical phenomenon

Grammatical accounts of medial answers often make reference to the form of complex Wh questions in languages like German which use a “Partial Movement" or “Scope Marking" construction as in (120).

(120) German Scope Marking (Müller 1997)

Was	*glaubst*	*du*	*[**wann***	*daß*	*sie*	*gekommen*	*ist]?*
What	think	you	when	that	she	coming	is

‘When do you think she is coming?’

Arguments have been made more recently that the sorts of child behavior shown above are not strictly speaking a grammatical phenomenon, but rather a parsing strategy. That is, children do not go through a “Partial Movement" stage in their grammatical development, where they have German-like rules about long-distance Wh interpretation, but rather do not parse these types of sentences as multi-clausal questions.

The evidence given for this argument partially has to do with the canonical sorts of examples, and the form of certain parentheticals (Omaki et al 2014) in English, or alternately as parataxis of two questions with an answer given to the second (Dabrowska, Rowland, and Theakston 2009). Consider the set of complex Wh questions in (121): (a) is a long distance Wh question, the interpretation of which correlates to a child’s medial answers under a grammatical view; (b) is a simple question with a parenthetical string separating the Wh word from the rest of the question (you could paraphrase this as ‘what was John hiding, according to you?’); (c) is two questions asked together, with one target answer (something like, ‘remind me what you said about what John was hiding’); (d) is the target sentence with an embedded question which isn’t given matrix scope for adults.

(121) a. Long distance Q: What did you say John was hiding?

b. Parenthetical: What, did you say, was John hiding?

c. Parataxis: What did you say, what was John hiding?

d. Embedded Q: Where did you say what John was hiding?

The claim of either non-grammatical account is that children giving medial answers to a question like (121d) are not interpreting it as a German-like way to express the question in (121a), but rather are either essentially ignoring the second Wh word and interpreting 'did you say' as a parenthetical as in (b), or treating the second question as syntactically not being embedded. There are several reasons to be skeptical of this approach, and prefer to analyze this behavior as a difference in adult and child grammatical representations.

One argument for the grammatical view is that the acceptability of the parenthetical and paratactic constructions does not hold for all sentences which produce medial answers. For example, medial answers are not limited to sentences where *say* is the matrix predicate, but also occurs with verbs like *tell* (deVilliers 2016), which do not typically allow this sort of parenthetical (122). Similarly, the parataxis account runs into questions in the cases where the matrix Wh word is not 'what', as question sequences like this do not typically have the function of seeking a single answer (123).

(122) ?What, did you tell me, was John hiding?

(123) #Where did you say? What was John hiding?

Another fact which favors the grammatical view is the occurrence of medial answers where the matrix question is polar (de Villiers et al 2011), as in (124). Note that this is not a case where a parenthetical of this same type is typically possible (125).

(124) Did he say what Jimmy fed the rabbit?

a. Adult: Yes/no

b. Child: An apple

(125) #Did he say, what did Jimmy feed the rabbit?

Like Wh-Scope Marking patterns seen in languages like German and Hindi-Urdu, we see evidence cross-linguistically of a type of Wh-Scope Marking lacking an overt Wh scope marker in the matrix clause, in languages like Malay and French (Oiry 2011 and citations therein).

Given the weaknesses shown here of these extra-grammatical accounts of medial answers, we can lean towards the grammatical account. From this we can make predictions about how children will behave with regards to FRs in similar sentences. Specifically, we predict that, since Scope Marking constructions require a matrix question-embedding predicate, sentences where non-Matrix Wh strings are unambiguously FRs, we should not see medial answers.

(126) a. Did Jimmy say what he fed the rabbit? → medial answer

b. Did Jimmy make what he fed the rabbit? ↛ medial answer

Experiment 1 looked at this prediction and found that unambiguous FRs do not in fact lead to medial answers.

3.2.2 Experiment 1: FRs are not given medial answers

Most adjectives take nominal arguments; as such, sentences like (127) can only be understood as containing an FR, not an embedded Wh question. As such we predict, given the discussion above about children's behavior with regard to embedded Wh words, children will not treat the Wh word in (128) as the matrix question word, despite it being embedded within a question.

(127) [$_{DP}$ What Dora saw] was scary

(128) Was [$_{DP}$ what Dora saw] scary? $\not\rightarrow$ What did Dora see that was scary

Experiment 1 used similar methods to those used in previous studies of medial Wh phenomena and found that indeed children do not give the erroneous interpretation in (128) to sentences of this type.

The participants in the experiment were twelve children ages 4-7, with a mean age of 5;10, at schools in Western Massachusetts. Two children were excluded from the data due to being unable to complete the experiment.

3.2.2.1 Methods

In the experiment, each child was told a series of stories. The child was then instructed to answer a question about some detail after each story asked to them by a puppet character. Some of the questions were of the embedded question form which have previously been found to elicit medial answers; some were of the type in (128). The experiment consisted of 16 questions, with 8 of the form in (128), 4 of the embedded Wh form, and 4 fillers.

If medial answers are subject to sentence-level syntax, as predicted under the grammatical analysis of medial answers, children should give medial answers only to the embedded Wh type of questions, and never to the type in (128). Conversely, if children are parsing material aside from the Wh clause as parenthetical or external, an interpretation of the FR as a question should be possible, of a kind paraphrased as in (129).

(129) What did Dora see? Was it scary?

Absence of this type of answer would be evidence both for a grammatical explanation of medial answers generally, and evidence that children know the relationship between selecting environment and the interpretation of FRs as definite descriptions rather than questions.

3.2.2.2 Results and discussion

Across all children, there were 20 instances of medial answers with embedded Wh questions; only one child gave no medial responses (Table 3.1).

Table 3.1: Experiment 1 results

	FRC	Medial Wh
Non-matrix answer rate	1.25%	50%
Total non-matrix answers	1	20

There was only one instance of a medial answer to an unambiguous FR, to the question shown in (130).

(130) Q. Was what Molly saw big?

A. An elephant

The result seems to be a strong indication that children do not allow the Wh words in questions of this form to have matrix question force.

This provides evidence that for the children as well as adults, the Wh strings are necessarily FRs, and thus the Wh words lack question force, but are rather interpreted as definite descriptions, by way of the type-shifting involved in the derivation of FRs (see Chapter 2). That is, the syntax of these questions requires the semantics of FRs. Evidence that children are not treating the Wh words in FRs in sentences like (128) as question words also comes from their explanations for their answers in this experiment, such as in the following dialogue:

(131) (Story: Ben played the tuba, which was hard, and Molly played the guitar, which was easy)

Cookie Monster: Was what Ben played easy?

Child: No

Experimenter: Why do you think that?

Child: 'Cause he couldn't play it

Absent are responses like "it was a tuba", which would have suggested an interpretation more like 'what did Ben play that was easy'?

This also seems to be evidence against a Usage-Based treatment of medial questions. The relevant sorts of strings (question forms of copular sentences involving Wh strings) are vanishingly rare in the data. In the Adam corpus, the only instances of a string like *was what* are in echo questions or specificational sentences; (132) shows two such instances in Adam's mother's speech (Brown 1974).

(132) a. His sister's name was what? (3;5)

b. This is what it should look like (4;1)

The sort of input in (132) should both be evidence for children to take assume Wh interpretations of the sentences in this experiment by Usage-Based accounts. Given that children see limited instances of the strings associated with the target construction (in this case, Wh strings in the subject position in a copular sentence, ie. predicational pseudo-clefts). Both of the types of sentence in (132) have question-like meanings in some sense: (a) is an echo question, and (b) is specificational, and thus semantically akin to a question-answer pair (Schlenker 2003).

Under Usage-Based theories, if children lack access to the relevant structure in a sentence, they will use a "cut and paste" method of interpretation to try to get at the function being sought by the speaker (Tomasello 2000); children's intake and production is affected by the (assumed) functional goals of their interlocutors, as well as the child's functional goals (Lieven 2010:2547). So in this case, the form-function input they could make use of is either specificational in its function, or interrogative. The latter would produce medial answers in this context; the former would produce predicative sorts of the sorts of responses not seen in (131).

All of these results suggest quite clearly that children do very well at distinguishing, based on factors like selecting predicates, the semantic content of FRs from that of embedded Wh questions. We can minimally say, then, that from quite early on children know that embedded Wh strings can be either DPs or CPs, and that certain syntactic environments force them to be DPs.

The next section looks at another type of context for FRs, and examines whether children can use the *internal* syntax of Wh strings to make this choice - that is, variation of form within the FR rather than variation in selecting environment.

3.3 Children don't know *Wh-NP for FRs

An important thing to note about the results of Experiment 1 is that they provide apparent evidence a case where children are allowing flexibility as far as the interpretation of subsentential strings - that is, Wh strings like *what Ben played* are allowed in both question positions and nominal (ie. FR) positions. Above production data show that children don't have any obvious problems in production of Wh-NP; phrasal Wh expressions occur in production where expected, and typically in an adultlike form. However, here I will show that children's knowledge of Wh-NP effects is incomplete: in a comprehension task, children do not distinguish between questions and FRs based on the presence of Wh-NP.

3.3.1 Disambiguating cues as an acquisition metric

As a central assumption for the present work is that children (and, in general, adults) use broad syntactic constraints and properties not just to rule out ungrammatical utterances, but also to block illicit readings of licit sentences. This is not trivially true; but, there is good evidence to assume it is true. To see this, take for example sentences like (133), which contains a question with two possible types of answers, a "matrix" answer,

where the Wh word *how* is interpreted as an adjunct of *learn*, and an “embedded" answer, where it is an adjunct of *cook*.

(133) How did Ben learn to cook rice?

Matrix: From a cooking show.

Embedded: In broth.

This ambiguity is fragile, and certain syntactic variations on this sentence eliminate it; if the embedded clause is a Wh clause itself, the embedded reading of ‘how’ becomes unavailable.

(134) How did Ben learn what to cook?

Matrix: From a cooking show.

*Embedded: In broth.

While (134) is, like (133), a licit English sentence, it does not share the ambiguity of where *how* is interpreted. This is derivable from a general ban on Wh movement out of Wh Islands (Ross 1967, inter alia). Where there is no plausible matrix interpretation of a Wh word, sentences with embedded Wh clauses are uniformly ungrammatical, as in (135a); the matrix interpretation gives a meaning like ‘John learned who what to give’ which is obviously bad, while the embedded interpretation, which would mean something like ‘Ben learned who to give what’ is unavailable. Thus the sentence crashes on the surface. The ungrammaticality of this sentence connects with the limit of possible meanings of (134), as illustrated in (135b).

(135) Wh Island effects

a. *Who did Ben learn what to give?

→ *Who did Ben learn [what to give] (*t*)

→ *Who did Ben learn [what to give (*t*)]

b. How did Ben learn what to cook?

→ ✓How did Ben learn [what to cook] (*t*)

→ *How did Ben learn [what to cook (*t*)]

Assuming that Wh-islandhood is a primitive of the grammar in some way, we predict that children will not produce sentences like (135a). This seems to be the case.

If full knowledge of a grammatical constraints includes not just non-commission of violations, but also using the constraints to disambiguate potentially ambiguous sentences, we predict that children will be restricted in the interpretation they give to questions like (135b). And in this case, experiments have shown that this is also the case: children will not violate the rule which bans certain sentences in interpreting potentially ambiguous sentences (de Villiers, Roeper, and Vainikka 1990; de Villiers and Roeper 2011); children will never answer 'How did Ben learn what to cook?' with something like 'in broth.' Similar results are seen in Otsu (1981) for Complex NP islands.

So we can move forward with the assumption that, for children acquiring a language, knowledge of the correct set of possible strings in a language can be carried over to this sort of disambiguation. Thus, for the present discussion, we assume that children's knowledge of the internal syntax of English FRs includes knowing the that strings which cannot be FRs could never be given an FR interpretation. Experiment 2 looked at this, with regard to *Wh-NP, finding that children do not seem to make use of the *Wh-NP rule as a way to rule out certain readings, though adults do.

3.3.2 Experiment 2: Children do not use Wh-NP as a disambiguator

Recall from the previous chapter cases where ambiguities of *see what* sequences can be resolved by Wh-NP effects, as in (136).

(136) Situation: Charles sees several books laying open. Among them is the book Sebastian wrote, though he cannot tell which book is which.

a. Charles saw what Sebastian wrote = True

b. Charles saw what book Sebastian wrote = False

Experiment 2 made use of this sort of cue in comprehension to be a crucial test of children's knowledge of the type of syntax for FRs which gives rise to Wh-NP effects.

The adult study was performed with 16 adults, all native English speakers, of various levels of education, recruited from a social media network. There were four items of the type above, counterbalanced, and four filler items of various types. In the child version of the study, the participants were 18 monolingual English-speaking children from Western Massachusetts, ages 5;3 to 6;10 (mean of 6;4).

3.3.2.1 Methods

The experiment compared results from a Truth Value Judgment Task for adults and children using sentences with see-Wh sequences, with situations like the one in (136) as well as situations where both readings are true. This yields a two-by-two experiment with the factors being syntax ('see what (S)' vs. 'see what NP (S)') and story type (whether or not the question reading of the target sentence was true).[4] An example of a minimal

[4]There is a third story type possible, where only the question reading is true. This would be something like (i).

i Cow has a list of who brought what to the party; it says that Goat brought chips, Pig brought hummus, etc. But by the time he gets out to the kitchen where all the snacks are, the chips have all been finished.

In this case, 'Cow saw what (snack) Goat brought' is true under the question reading, but the definite reading is false. While for the purpose of studying semantic development, this sort of story would be

pair of stories is given in (137), with the relevant sentences to be judged by participants, and their expected truth values.

(137) Goat is going to Cow's birthday party. He's bringing Cow a gift in a green box. When he gets to the party, he puts the gift on the table next to the other wrapped presents. Later Cow comes and sees the table full of presents. She gets very excited about all of them. She opens them all. Cow got a hat, a book, and a cake. She likes all the gifts she got.

a. Cow saw what Goat brought (**True**)

b. Cow saw what gift Goat brought (**False**) ⇐

Goat is going to Cow's birthday party. He's bringing Cow a gift in a green box. When he gets to the party, he gives her the gift. She thanks him for it and puts the gift on a table with the other presents. Later she opens all the presents. Goat got her a nice hat! Cow thanked Goat for the gift. She likes all the gifts she got.

c. Cow saw what Goat brought (**True**)

d. Cow saw what gift Goat brought (**True**)

Condition (b) is the critical condition, as it is where the syntax of the sentence to be evaluated (containing Wh-NP) isn't allowed in the semantic environment (the Question reading is false).

The adult version of the experiment consisted of 8 items, including 4 counterbalanced story-sentence sets of the type in (137), and 4 fillers of various types. The child version consisted of the 9 items: the 4 key stories of the type in (137), 4 fillers, and 1

---

of interest, I will leave it aside for now as it doesn't seem to rely on surface syntax and Wh-NP effects, since strings allowing (definite) FR readings are generally a subset of those allowing question readings (Caponigro 2003).

Table 3.2: Adult results for Experiment 2

	Q False	Q True
Wh	68.75% *(12.0)*	93.75% *(6.3)*
Wh-NP	25% *(11.2)*	87.5% *(8.5)*

Table 3.3: Child results for Experiment 2

	Q False	Q True
Wh	62.5% *(12.5)*	81.25% *(10.1)*
Wh-NP	68.75% *(12.0)*	81.25% *(10.1)*

simple training item. Participants were excluded from analysis for failing to provide the target response on the training item or on more than one filler.

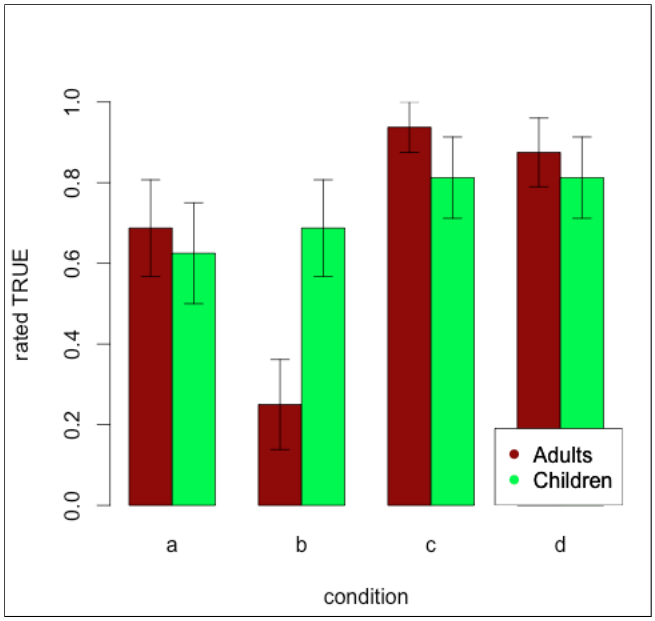
3.3.2.2 Results

No adults were excluded from the analysis. Two children were excluded from the analysis for failing on more than one filler item.

For adults, the stories and prompts were given in text form; for children they were given as series of images accompanied by recorded narratives. The results are presented in Tables 3.2 and 3.3, which show, by condition, rates of “True" responses and standard errors, and shown again in Figure 3.1 comparing both adult and child results side by side.

Table 3.2 shows that in the conditions where the Question (“Q") reading is True, syntax has little to no effect on response, and generally adults will evaluate either sentence type as True; but, where the Q reading is false, the presence of Wh-NP has a considerable effect, with Wh-NP sequences significantly less likely to elicit a “True" response ($t = 2.67$, $p = .01$). Table 3.3 shows that children are equally likely to judge sentences as True regardless of syntax for both conditions. Children were significantly more likely than

Figure 3.1: Adult and child responses by condition, with Standard Errors



adults to give True responses in the Wh-NP condition with Q-False stories (Condition b; $t = 2.67$, $p = .01$).

Two things should be noted about the Q False conditions, particularly the adult results. Where the syntax allows the FR reading, adults are not at ceiling, and indeed aren't far from chance. This should be expected as the surface strings can correlate with either FR syntax or question syntax (and thus can have either semantic value), but only one is true. This seems to show that there is not a strong preference of one reading over the other in this syntactic frame. For the Wh-NP items, adults are not at floor; this is a priori unexpected, but is due to an item effect. The sentence for the relevant item is in (138). Looking at the item in terms of the certain exceptional FRs, including both Wh-ever FRs and, crucially here, those like the one in (138) which allow Wh-NP gives some insight into how this could have been the case: FRs which include a plural NP can be used to give a "limited number" interpretation.

(138) Bill sees what articles Mary writes

This sentence can be given an FR reading on an interpretation of, roughly, “Bill sees such articles as Mary writes." Thus it is in fact expected that adults would accept this, and we can understand this result as not anomalous, but rather confirmation that adults have sensitivity to such fine-grained syntactic cues when choosing between FR and question parses in sentences like this. Chapter 4 (§4.4) discusses this type of FR, and why it is an apparent exception to *Wh-NP.

3.3.2.3 Discussion: This is a representational issue

It appears from these results that children are not making use of a syntactic cue for disambiguation which adults have access to. A question which is thus raised is what the nature and origin of this deviation from adultlike behavior is. Feasible explanations are: children are making non-adultlike parsing choices, being unable to revise their first interpretation of ‘see what’ after encountering an NP (a “Kindergarten Path" effect) or simply do not parse the NP which they encounter; children do not have a clear distinction between the two readings; or, children’s grammatical representations of FRs diverge from those of adults. I argue that the latter is in fact the best explanation.

Children parsing the sentences based on an expectation that *see* should come before a nominal, and so any instance of *see what* should produce an FR reading of the Wh string, immediately runs into problems when compared against the input. To expect an FR interpretation in this environment, children would have to hear many examples of *see what* in environments that favor FR readings. Such an environment would have to look something like the sentences in (139): it would have to involve a condition where a subject clearly sees something but doesn’t know what it is, or where the pragmatics favor treating the referent as an object (rather than a proposition or set of propositions).

(139) a. I see what you made, but I don’t know what it is

b. Do you see what I made? Pick it up.

The Adam corpus (Brown 1974) seems to show no examples like this, or any instances of *see Wh* in an unambiguously FR environment, and overwhelming evidence that children should expect Question semantics given this string. There are 57 instances of see Wh in the input (excluding cases where the Wh word is not the complement of *see*, in situ/echo questions) apparently have question semantics. (140) shows relevant examples, including one (140b), which has a possible, but unlikely, FR reading.

(140) a. Adam, 2;8

Investigator: Do you want to see what it says?

Investigator: Ask your Mommy what it says.

b. Adam, 3;10

Mother: Go backwards and see what happens

The discourse in (a) could have an FR reading only if the child is being asked to look at some writing, and is being asked to look at it only; the following utterance highly disfavors that reading by including *ask ... what it says*. The utterance in (b) has a possible FR reading, but it would have to mean something like 'witness what happens', which is unlikely. There is thus no reason to expect that children would be primed to expect an FR interpretation of 'see what' which they are unable to revise. If anything such an explanation would predict that children trend toward not accepting the sentences where the Q reading is false, regardless of the presence of an NP.

The other possibility, making reference to children's parsing strategies, assumes that the child is simply not parsing the NP into their representation of the input sentence. This would correctly predict the lack of a difference between responses based on the presence/absence of Wh-NP. However, there are some points against this. One is the evidence from production in §3.1.1 above which shows that children, by this age and even a bit earlier, produce Wh-NP sequences in a variety of (adultlike) positions. A second point against this processing-related account comes from the form of the stimuli, which

put focal stress on the NP ('Cow saw what **gift** Goat brought'), making it unlikely that children would simply be ignoring it.[5]

The last possible argument, the argument on conceptual grounds that children simply are not considering both meanings, runs into immediate problems. It would require that children don't perceive of a difference in these contexts between the purely visual and epistemic senses of 'see' (or at least, don't perceive the situations as distinguishing them). One immediate argument against this is that other work on cognitive development suggests that children come to understand the conceptual relationship between seeing and knowing quite early on (O'Neill et al 1992, Schmidt and Pyers 2011), so at the very least the cognitive mechanisms are in place to understand the relevant distinction.

The numerical results in this experiment also weaken the case of such a conceptual explanation, as there is an apparent trend of difference between responses based on the story type. There is, however, no significant main effect of story type observed. A better argument against this challenge comes from the explanations children give for denying sentences in the Q-False story conditions, which indicate knowledge that there is a difference between seeing of a thing and seeing something about it. The following dialogue between the experimenter and a child gives an example of this.

(141) **Prompt**: Goat saw what Cow ate (Q false condition)

Child: No.

Experimenter: Why didn't Goat see?

Child: Because he wasn't there, he was outside.

This is in a case where the syntax doesn't distinguish between the two readings, but the situation does; the child is able to arrive on the question reading of the string, and

[5]Given this stress one might assume that in fact children might be mishearing this as something closer to 'the gift Goat brought'. However, this would still fail to predict the results, and predict much greater acceptance rates for condition B, as the referential reading is the only available reading.

evaluate it as false in the given context. Thus, despite the lack of statistically significant difference between children's responses in the two conditions, it cannot reasonably be assumed that children lack two distinct readings of 'see' in these contexts, but rather do not make use of *Wh-NP to distinguish them.

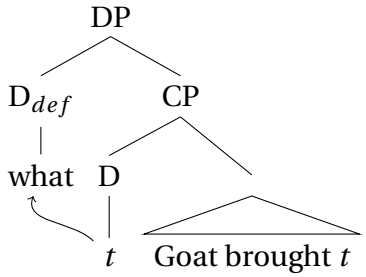
With all these facts in mind, I will continue with the assumption that children's performance on this experiment is based on a grammatical representation of definite FRs which differs from that of adults. But, before discussing the deviation from the target grammar, one must establish how the target grammar represents the construction in question.

3.4 How do children represent FRs?

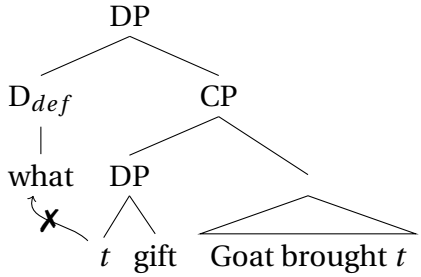
Recall the proposed structure for FRs in the adult grammar which gives rise to *Wh-NP (142): it occurs as a conspiracy between three factors: the derivation of FRs from Wh questions, the requirement of an overt D in standard definite descriptions in English (and the use of a Wh word to fulfill this), and the ban on left branch movement.

(142) Free Relative Clauses with external Definiteness operator

a. Licit structure



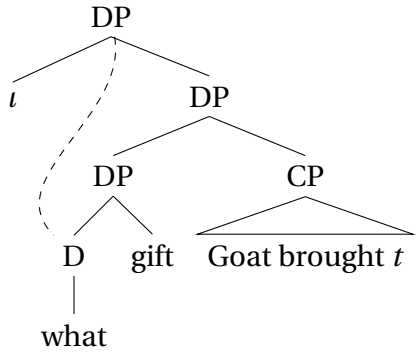
b. Illicit structure (*Wh-NP)



All of the pieces of this analysis are subject to cross-linguistic variation (as shown in the previous chapter), and thus must be learned. A child could consider a number of alternative grammars to English on the way to convergence with the adult grammar, many of which would allow Wh-NP.

Of course, given the discussion in Chapter 2 about the factors which act together in *Wh-NP, there are some more primitive syntactic assumptions made which we could imagine were missing from the child's grammar. In particular, children might not have the (perhaps violable) constraint that semantic type shifters be included in the narrow syntax, or they might lack the constrained labeling alorithm which Cecchetto and Donati (2015) propose which would block labeling of a syntactic object by moving a complex element; that is, the syntax of FRs could be something like (143), where the dashed line indicates the relationship between the root label and its source.

(143) Completely non-adultlike FR



Here the label of the DP is embedded within a complex moved object (a DP including a Wh word), and the ι operator which allows the structure to have an FR interpretation

is syntactically contentless. As such, to give this structure its proper interpretation and distribution, the label must come from a lower D^0, the closest of which is in a position which could not typically label the structure.

While the constraint I proposed in the previous chapter that semantic type-shifters carry syntactic categorial labels wherever possible may either not apply in general or not apply to children's grammars, an assumption that children (particularly five year olds) do not follow a labeling algorithm that should be a basic building-block of the syntax, I suggest that this assumption is not optimal and will not consider it further, barring any compelling evidence for it.

Having abandoned this more radical structure, and having argued above that the explanation for this behavior still must be grammatical in nature, the source of children's divergence must be one of the possibilities in (144).

(144) i. Delay in acquiring requirements on overt D^0

ii. Delay in acquiring ban on left branch movement

iii. Early misanalysis of FRs as RCs

I argue that (i) is the best option, though (ii-iii) must be considered.

Recall that experimental evidence from Gavruseva and Thornton (2001) showed children producing non-Pied Piping of possessive Wh expressions, but just in the case of long distance movement, and primarily with children producing the *'s* morpheme in the lower position (*who do you think Ø's hat...*). Given this, there is reason to think that this behavior as more akin to illicit pronominal resumption than of left branch movement. Resumption within pronouns in children's speech as they acquire a variety of languages, including English, has been observed at fairly significant, not dissimilar to these numbers, in children's production of non-subject Relative Clauses (Labelle 1990, Pérez-Leroux 1995). Labelle (1990) in particular argues that this is a evidence of chil-

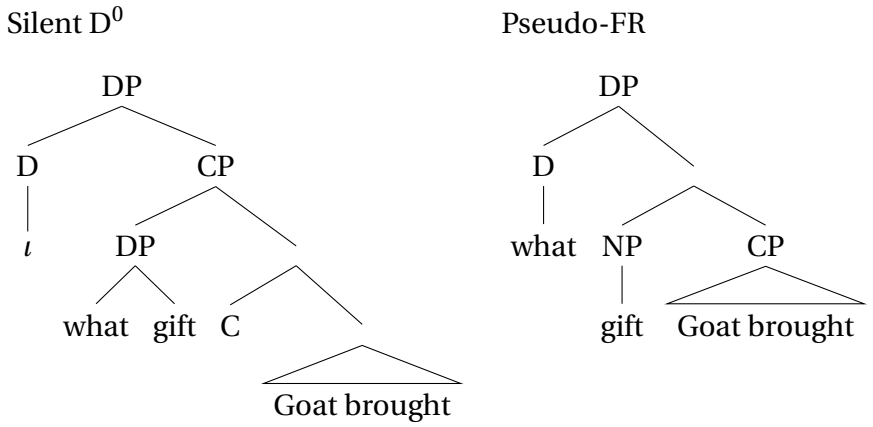
dren giving a non-movement analysis to certain Wh constructions (though Guasti and Shlonsky (1995) argue against this).

The sentences observed by Gavruseva and Thornton (2001) have in common with pronominal resumption that they involve a long distance dependency and production of an additional element in the position from which movement took (or would have taken) place. It is also worth noting that the similar type of Wh sub-extraction Gavruseva and Thornton discuss in French as an analog to this type of Left Branch movement, pronominal resumption in relatives is allowed in at least some varieties of French (as Labelle notes). So to the extent that children producing this sort of movement in English are using a French-like grammar, it is more likely that they are using the French-like Resumption syntax rather than Left Branch movement.

Considering this similarity, as well as the fact that there does not seem to be a parallel sort of non-target production phenomenon for short distance pied piping or pied-piping of non-possessive Wh-NP, analyzing children as going through a stage which allows Left Branch Movement, and so (ii) from (144) is not the best explanation of the Wh-NP phenomenon observed here. So, we move to choosing between options (i) and (iii).

The representations of FRs resulting from one of these are shown in (145). In the first, the "Silent D^0" analysis, either Wh-NP moves as a unit and does not project, or the CP is dominated by a silent definite determiner which labels the entire structure. This is essentially a variant of (143) where the type shifter carries a categorial feature which is allowed to label the structure given a restrictive labeling algorithm. In the second, Wh-NP does not form a constituent, but rather NP is the sister of the gapped clause, and *what* is externally merged as Head of a DP, and which takes the NP+CP structure as a complement. This structure more closely a headed RC than an FR. To this end, I will call this a "Pseudo-FR", after the terminology of Battye (1989).

(145) Possible misanalyses of FRs for children



The latter analysis is roughly the analysis of Wh-ever type relatives (from which the term “Pseudo-FR" is borrowed) used to account for the lack of Wh-NP effects, as well as several other properties, of Wh-ever constructions discussed by Battye (1989) and Cecchetto and Donati (2010).

(146) Wh-ever FRs

a. Cow will eat what*-(ever) dish Goat cooks.

b. Cow will eat what*-(ever).

c. Cow will eat how*-(ever) much Goat cooks.

The former analysis, where Wh-NP forms a constituent and is moved on its own, essentially analyzes FRs as instances of Wh movement alike to others such as Wh questions, in that a DP with a Wh item in it is moved to the specifier of CP, and the syntactic difference between FRs and Wh questions relies on a silent definite D head which projects its label on the entire construction. That this should be an allowable derivation is predicted by the semantic analysis of FRs of Chierchia and Caponigro (2013), who analyze FRs as plural definite descriptions denoting the entity which has the property denoted by the gapped clause (that is, an FR like ‘what Goat brought’ denotes the maximal set of things that Goat brought). The difference between adult and child syntax thus

becomes whether or not the Wh word must be the phonological exponent of the definiteness operator. Whether or not this is the ideal analysis for Wh-ever (see the following chapter) it is in any case a priori a feasible analysis for children to take.

Other studies of Relative Clauses in acquisition show some issues for the Pseudo-Free Relative analysis. As discussed above, FRs are produced earlier than other relatives. Thus this analysis would involve children learning FRs first as if they were headed relatives, but still not producing Headed RCs early on.[6]

Of course, the children in the experiment described above are 5-6, an age where children have begun producing headed relatives, so it is possible that at this age children have one unified representation for relatives of all kinds. This would involve a sort of “U-shaped" acquisition path for FRs, where their first representation is apparently adultlike, then they learn a generalized form for all relatives (allowing Wh-NP), and then they eventually learn the adultlike form (disallowing Wh-NP).

Table 3.4 shows the sort of developmental paths predicted by both analyses: The Silent D^0 hypothesis predicts that at the earliest stage, before Headed relatives emerge, children will have Silent D^0 syntax for FRs and nothing for Headed Relatives (HR), and go through a stage where the Silent D^0 syntax remains for FRs, but HR syntax has emerged for Headed Relatives, before converging on the adultlike representation for both. The Pseudo-FR analysis predicts that children will have the adultlike analysis for FRs - where the Wh word moves to D^0 - at the beginning and end of the acquisition path, but apply HR syntax to all relatives once it emerges, before dividing the two syntaxes again.

---

[6]This assumes that children entertain one grammatical representation for a construction at a time, eventually changing them to a different one if given sufficient input; a different model would allow both analyses to be “competitors" (in the sense of Yang 2003) for the best representation of FRs until the input convinces them to allow one over the other. Under this assumption the acquisition path involving the Pseudo-Free Relative analysis becomes more plausible. However, given the absence of RCs in children's production, we could assume that the child's grammatical state gives greater weight to an analysis more like Silent D^0 at this stage.

Table 3.4: Paths predicted by different analyses

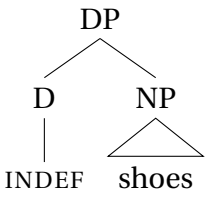
	Stage 1		Stage 2		Stage 3	
	FR	RC	FR	RC	FR	RC
Silent D^0	Silent D^0	–	Silent D^0	RC syntax	Wh-to-D^0	RC syntax
Pseudo-FR	Wh-to-D^0	–	RC syntax	RC syntax	Wh-to-D^0	RC syntax

While this sort of acquisition path is possible, the acquisition path predicted by the Partial Derivation analysis seems more likely: this would involve making a generalization about all Wh movement (that Wh expressions move to a specifier above C^0), and later learning a particular feature of a single Wh construction (Head movement of the Wh word).[7]

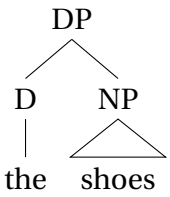
Another advantage of the Silent D^0 analysis is that it aligns with some other facts about the acquisition of syntax-semantics in the nominal domain: children have difficulty consistently requiring that morpho-syntactically indefinite DPs have indefinite readings (Perez-Leroux et al 2004a, Gavarró et al 2006), particularly in “NP/DP" languages like English, where there is a correlation between definiteness and overt D heads (Perez-Leroux, Gavarró, and Roeper 2011); that is, for children learning English in particular, indefinite DPs such as (147a) are over-assigned definite semantics expected for something like (b). Thus, while something like (147a) can only include a silent existential/indefinite D head for adults, children allow a silent definiteness operator here.

[7]A similar line of reasoning is pursued in Guasti and Shlonsky (1995) in order to explain the early emergence of FRs compared to other relatives: FRs allow for obvious, overt Wh movement, and do not require the more complex sort of null operator (or, equivalently, copy-deletion movement) syntax associated with headed relatives, and thus are able to be acquired early.

(147) a. The girl needs shoes



b. The girl needs the shoes



In view of the Silent D^0 analysis of children's performance on Experiment 2, this is akin to saying children at preschool ages acquiring English do not consistently apply the rule that definite semantics must be accompanied by a phonologically overt D head. Perez-Leroux, Gavarró, and Roeper (2011) suggest that this effect with bare NPs correlates with the parameter which allows mixed NP/DP phenomena in English. If the form of English FRs is a type of extension of this correlation between overtness and definiteness, then a delay in adultlike performance on Wh-NP effects is predicted to occur as long as in interpretation of bare NPs as indefinites, or longer.

3.4.1 Production-comprehension asymmetries

A crucial assumption about acquisition being made here, following the discussion in Chapter 1, is that knowledge of constraints on form corresponds not just to production matching adultlike grammars (absence of mistakes of commission), but also to the use of syntactic constraints to disambiguate where possible. That is, if children have adultlike knowledge of the distribution of Wh-NP, they will not only fail to over-produce Wh-NP, but also understand the disambiguation that occurs in (136) above.

These asymmetries have been observed in two different directions - in addition to cases like Wh-NP, where a grammar that apparently allows Wh-NP in FRs does not give

rise to non-adultlike production, there are cases where comprehension suggests adult-like knowledge of things not produced. Specifically, despite the evidence cited above that children are delayed in producing headed RCs relative to FRs, there is also evidence that they comprehend headed RCs earlier than they produce them, closer to the time they begin producing FRs (Adani 2011). Thus, an account of the development of relatives broadly must allow both headed RCs and FRs into children's grammars early, but account for the discrepancy in production.

Children's speech can be described as "Conservative" in the sense that they tend not to produce utterances if they do not have enough evidence to set all of the relevant parameters (Snyder 2007). A strict sense of Conservatism would be that children will not produce a construction until they have set the relevant parameters. However, FRs appear early in children's production, despite the evidence presented here that not all parameterization of the nominal domain is complete. One explanation of this sort of fact is Maturational - there are possible operations and representations in adult grammars which children do not have access to until some later point in development (Borer and Wexler 1987).

In the case of FRs, a sort of Conservatism seems to be exhibited by the fact that, based on the above experimental data, children do not have a grammar which gives rise to a *Wh-NP rule, and yet children do not produce illicit Wh-NP in FRs. We could describe this in Maturational terms if we assume that adjunction of a Wh word to D^0 is unavailable until later in a child's development. However, unless we were to limit this maturation to this very narrow instance, this would be an untenable position, as other types of head adjunction seem to be available to children - for example, children produce sentences with Auxiliary head movement (T-to-C movement, as in *Will Charles leave?* etc.).

If we adopt a weaker definition of Conservatism, we could have a better explanation of this behavior: Parameters are set stochastically and over time (Yang 2003), both an

amount of uncertainty in setting and the number of relevant Parameters unset corresponds to less production. In the case of FRs, most parameters regarding Wh movement are set by an early age; only exact conditions on overtly realizing D^0 (and ways in which it may be realized) are delayed.

Further, allowing for children to have access to adultlike representations of RCs can help us understand the asymmetry between their early comprehension and delayed production. RCs are delayed further because there is more in the grammar that goes into their production, including the heavily under-determined choice between movement and matching analyses.

The source of this production-comprehension asymmetry may not necessarily be entirely grammatical, however; children show a difficulty in comprehending Object relatives which seems to relate to cases where the relativized NP must move past a subject which is grammatically similar, in animacy or Φ features (Friedmann, Belletti, and Rizzi 2009), a fact which is perhaps best understood as a grammatical constraint which complicates online comprehension. A distinction between FRs and RCs in this way is that Wh words are necessarily of a different syntactic class from referential DPs, and so movement of Wh words over subjects in this case does not give rise to the same sorts of issues for comprehension. Friedmann, Belletti, and Rizzi give some experimental evidence supporting this idea.

An issue with this hypothesis, however, is that object Wh questions seem to show a similar subject-object asymmetry, and involve the same sort of movement over a subject as object FRs (O'Grady 1997). So again, the difficult grammatical choice of how to analyze an RC must play some part in the relative delay in their production.

Recalling the discussion in Chapter 2 about the nature of the feature [+Overt], the fact of this production-comprehension asymmetry regarding *Wh-NP may be revelatory about the nature of this feature. One of the options I posit above as a mechanism for evaluating success or failure of a derivation with regard to [+Overt] that representations

with or without Wh-to-D movement can be generated by the syntax, and it is only at PF that those without are blocked or judged as ungrammatical. With this assumption, we could guess that children generally construct FRs (and other DPs) with and without overt definite D, to the extent that Wh-NP is rarely seen in children's speech; it just isn't the case that, upon hearing what should be an illicit FR, they do not **enforce** the restrictions imposed by [+Overt], and so do not judge the offending structure as ungrammatical.

This may speak as well to the gradual and stochastic nature of language learning and Parameter Setting (per Yang 2003): as children's grammars approach Adult English, they are more and more likely to generate the [+Overt] version of definite D^0, and eventually come to a state where they would never do without it. This is a fairly comfortable theoretical position to be in, as it allows analogy with other aspects of lexical variation (such as generalization of particular morphological rules; see discussion in Chapter 1). In the next section I will discuss what the manner of Parameter Setting relevant to the Overtness feature should look like.

3.5 Parameters of the Nominal domain

The Silent D^0 analysis of children's FRs allows for a nuanced picture of the acquisition of (necessarily) filled positions within the DP: part of the acquisition of nominal syntax involves knowing which functional projections require overt material and which don't (as well as determining the precise sequence of functional heads available), which allows the sort of cross-linguistic variation in FRs and other aspects of the nominal domain described in the previous chapter.

3.5.1 The scope of the problem

Cross-linguistic variation in the form of nominals shows a diverse typology of determiner systems in terms both of what functional categories are overtly expressed and, perhaps, which are present in underlying representations. In this section I will describe

some observed variation in both what must be expressed overtly in the nominal domain, and what may fulfill this obligation, as a way of defining acquisition goals.

A learner of any language must identify (i) what overt determiners are available in their grammar, (ii) what conditions (if any) allow for no determiner to be present, and (iii) whether or not there is ever an unpronounced functional head in the absence of an overt determiner. An examination of cross-linguistic variation of what conditions allow bare nominals (in the surface sense) reveals a good amount of variation. Some of this variation is sketched in (148), for the moment using a broad notion of the distinction between definite and indefinite.

(148) Typology of overtness-definiteness interactions.

Hindi-Urdu: nominals may be bare

Welsh: Indefinites are bare and definites are marked

English: Plural indefinites may be bare; other nominals must be marked

French: Definites must be marked; non-predicative indefinites must be marked

Cross-linguistically there seems to be something implicational here: there don't seem to be any languages where indefinites but not definites have a mandatory overt determiner.

(149) Anti-Welsh nominals

a. Definite: Sebastian likes book

b. Indefinite: Sebastian likes *(some/a) book

This may point to something important for learners; if they encounter a sentence like *Sebastian likes book* with a definite reading, this is evidence for a Hindi-Urdu-type language, as it is not available in any of the other types in (148).

Based on observing form-meaning pairings like (148), a learner must hypothesize about the syntactic form of their nominals: what functional material is going unpronounced? There are different ways to approach this question. One might be to assume that learners will only project functional heads they have evidence for (ie, don't project DP if there is no D^0). Some languages seem to show consistent behavior as to whether DP is consistently projected even where overt D^0 is absent (Bošković 2005), for others it's been argued to be the case that D^0 is sometimes entirely absent (Chierchia 1998, Cheng and Sybesma 1999).

Given this variation, the child's path in establishing the exact structure of their nominal domain, and the number of functional heads which may be unpronounced where present, is a non-trivial one, though certain implications may make certain paths easier than others. Below I will show some specific steps which children have been observed taking.

3.5.2 Some observable stages in the path

There is a good deal of extant research on the path children take to recognizing the necessary meanings of bare nominals. A summary by Pérez-Leroux, Gavarró, and Roeper (2011) gives evidence that children in general tend to converge on the target grammar's treatment of (surface) bare nominals, specifically with reference to restricting bare nominals to indefinite interpretations in English vs. in Romance languages; the latter are more constrained regarding the presence (and especially absence) of determiners.[8]

Various studies looking at this acquisition problem make reference to the Nominal Mapping Parameter (NMP) of Chierchia (1998). This is, to be more specific, a *pair* of

[8]Much of this work, starting with Chierchia (1998), points out that Bare Objects seem to be special, both in the acquisition path and typologically; bare subjects not allowed in Romance. The experiment discussed here involves object FRs. We might predict some distinction for children's understanding of *Wh-NP in subject positions, based on this.

binary parameters which determine the possible distribution of bare nominals. The two parameters concern whether bare nominals may be interpreted as predicates (±Pred) or arguments (±Arg). This is summarized by Guasti et al (2008) the possible permutations of these parameters, and examples of each, as in (150), a simplified version of the typology given in Chierchia (1998).

(150) Languages in terms of NMP

a. [+arg, -pred] → Chinese-type languages (classifier languages)

b. [+arg, +pred] → Germanic

c. [-arg, +pred] → Romance

d. [-arg, -pred] → Impossible language

As pointed out in Pérez-Leroux, Gavarró, and Roeper (2011), the setting of parameters like these may involve inference based on surface-unrelated patterns, especially as it deals with a lot of surface ambiguities. For example, A child learning French who encounters a sentence like *Charles est peintre* 'Charles is (a) painter' will not instantly be able to choose between a Chinese-type or Romance-type setting (a distinction which, very inconveniently for the child, contrasts on both parameters), and will have to rely on relating this sentence to sentences with classifiers.

The conceptual prediction that there should be a delay in certain languages on the setting of these parameters is borne out by specific data from specific experimental data: Children learning English show higher rates of errors in interpreting bare nominals compared to children learning languages like Catalan (Pérez-Leroux, Gavarró, and Roeper 2011) or Spanish (Pérez-Leroux et al 2004b).[9] Since bare nominals are so restricted in these languages, the "triggers" for their parameter settings are more abundant than in

[9]See also Pérez-Leroux et al (2004a, Gavarró, Pérez-Leroux, and Roeper 2006, and others.

English, where bare nominals appear in a wider variety of positions, and thus the relevant parameter setting lacks a singular trigger.

The problem for learning the form of FRs, and how they fit into the setting of these parameters for English nominals, there is an issue that stems from their semantic value as *plural* definite descriptions (Caponigro et al 2012); bare nouns are available in English argument positions, but only if they are plural. The learning of FR syntax comes from mapping the syntax of plurality to a particular type of determiner, exactly what children have been shown to be delayed in in the above-cited studies. Since children learning English do not encounter sentences like *Molly is doctor* or *Ben ate sandwich*, the setting of the relevant parameters for the heads involved in singular nominals is relatively easy. But, for plurals, the data are more confusing, and children are more delayed.

While I have asserted here that the delay in *Wh-NP arises due to the fact that the learning of the parameter which requires definite D^0 to be overtly realized, I have not yet explained why it seems to be slightly more delayed than the proper interpretation of bare plurals. We can see data that would be particularly confusing for the child here, where strings that at least on the surface are very similar to FRs which seem to have a singular interpretation:

(151) This sandwich is what Molly made

This is involves a Wh construction in a nominal position (predicative position in copula construction), so a child would assume its properties are those of FRs. But, are two issues with this sentence that would lead to a problem for children learning the representation of FRs: there is no apparent determiner, despite apparently being a definite description, and the reference is clearly singular.

3.5.3 Parameters, micro- and macro-

The discussion here has concerned the setting of Parameters within the nominal domain. As such we must now return to the discussion in Chapter 1 examining what the

nature of parameters is, both within and without this domain. Do parameters represent a fixed set of “switches" which come pre-installed as part of the acquisition device, and which may be set to “on or off" to derive any particular target grammar, or are they an epiphenomenon emerging from some other property of language learning or of the mind? In Chapter 1 I discussed in some detail the idea from Minimalist and proto-Minimalist work (Borer 1987, Chomsky 1995) suggesting that parametric variation in general is a property of lexical items: parametric variation does not necessarily need to make reference to the “switchboard" of innate parameters envisaged in work like Chomsky (1981), but can be described in terms of differences between the functional lexicons of individual languages, and thus parameter **setting** is subsumed in the acquisition of the lexicon rather than being a separate task within acquisition.

A specific formulation of this which can be useful for discussions of parameter setting - that is, the acquisition of a particular grammar - comes from Biberauer and Roberts (2015). They discuss the distribution of formal features at different levels of a hierarchy across a language, to account for parametric phenomena which seem to apply to languages as a whole (“Macroparametric" values) vs. those which are restricted to smaller portions of a grammar (“Microparametric" values). Specifically, they propose a hierarchy of parametric types with four levels.

(152) **Parameter hierarchy** from Biberauer and Roberts (2012):

For a given value v_i of a parametrically variant feature F:

a. **Macroparameters**: all heads of the relevant type, for example, all probes, all phase heads, and so forth, share v_i;

b. **Mesoparameters**: all heads of a given natural class, for example, [+V] or a core functional category, share v_i;

c. **Microparameters**: a small, lexically definable subclass of functional heads, for example, modal auxiliaries, subject clitics, etc. shows v_i;

d. **Nanoparameters**: one or more individual lexical items is/are specified for v_i

Unlike most traditional versions of the theory of Parameters, especially those which focus mostly on Macroparameters (Chomsky 1981, Borer and Wexler 1987), this not only states that parameters may have different scopes from each other, but the same parameters may in principle apply at any possible scope in any given language. There might be a language where all probes have the feature value v_i, where all heads in the verbal domain do, where some subset of functional heads do, or where one idiosyncratic head does.

To spell out how this could apply to the present case, we can posit that the syntactic property which requires overt phonological material is a parameter [±Overt]. In English we could describe this as a Microparameter - it is a feature of the class of functional heads which are definite determiners (*the* and D^0_{rel}); in Romance, it is Mesoparametric as it is the appropriate parametric value for any D^0 (just about). We could also imagine potential languages along this hierarchy where the setting is Macroparametric (any functional head must be pronounced) or Nanoparametric (there is one particular determiner which consistently must be pronounced).

Put this way we could reframe (152) as covariance between features: a Mesoparametric value involves covariance between a setting of some core functional feature like [+N] or [+V] with some other feature like [+Overt], whereas a Microparametric setting involves covariance between some feature like [+D] with such a feature. That is, a parameter setting is an *if-then* statement like (153): any lexical item with a named set of features must have another feature.

(153) For some sets of features F_1 and F_2, for all lexical items L:

if L = [+F_1, ...] then L = [+F_2]

Defining the terms in (153) as **sets** of features is necessary to get narrower parameter settings that might refer, for example, to only modal auxiliaries, or only plural determiners.

As a concrete example, the Nominal Mapping Parameter setting in Germanic could be described as covariance of [+Arg], [+Pred], and [+N] - thus, a Mesoparametric setting.

This model is useful in terms of acquisition as we could describe parameter setting in terms of the child postulating a distribution of features across heads at different levels on the hierarchy. To start with we could assume, for any parametric value v_i, a child either begins conservatively, assuming first that any parametric value is nanoparametric until evidence provides a larger level across which they can generalize v_i; or the child begins liberally, assuming first that v_i is macroparametric until there are exceptions which lead them to narrow the scope of v_i.[10]

The case of the feature [±Overt] in English, as defined in Chapter 2, seems to be a case where a narrow generalization becomes generalized: evidence seems to suggest that children learn that the definite D^0 which selects NPs is necessarily pronounced (that is, *the*) earlier than they learn *Wh-NP. We could then describe this as children learning a Nano-Parametric setting where just the particular lexical item *the* is [+Overt], which is generalized to a Micro-Parametric setting which we could describe as something like in (154).

(154) For any lexical item X:

If X = [+D, +Def], then X = [+Overt]

3.5.4 Applying the Minimalist toolbox

The works cited above on the acquisition of different types of nominals rely on the principles of semantic type proposed in Chierchia (1998), in particular the parameteri-

[10] In principle we could also assume some medial starting point, but it is unclear what sort of learning model could pressure change in both directions at once.

zation of different nominal types. Another important part of Chierchia's proposal is his use of semantic Type-Shifters, and his Blocking Principle, a constraint which limits the distribution of type shifters which lack syntactic content. In the previous chapter I proposed the reformulation of the Blocking Principle in (155):

(155) Wherever possible, any semantic type shift operator must carry a syntactic category label.

Aside from being a tool for understanding the form and interpretation of different types of nominals, we can think of shifters, and the constraint in (155), as a tool for acquisition: knowing the lexical semantics of different predicates necessarily informs the semantics of their arguments.

We can use this rule, and some others, to start putting together a formal representation of the steps in the acquisition path, and the choices that get a child there. First, recall some of the ways in which Merge constrains possible representations in general, discussed in Chapter 1. Among these are that movement requires there must be an Agree relationship between a moved object and its target, and that the categorial label of the output of Merge must project from some locally accessible position.

The first constraint tells the learner that a string with Wh movement must include a head which licenses it - an X^0 with a [Wh] feature. If a child knows the basic form of Wh questions, they will know that this head should be a C^0, and so posit that a Wh word in a Wh string, either an embedded question or an FR, will have a C^0 which attracts the Wh word. This is formalized in (156).

(156) If [what ... t_{wh} ...] then [what C^0_{wh} ... t_{wh} ...]

Or, to put this in a way more closely following the template in (153):

(157) If [what X^0 ... t_{wh} ...] then X = [+C, +wh]

The second constraint, in concert with (155), allows the learner to determine the nature of the D^0 in an FR: when a Wh string appears in a nominal position, the child will have to find a way to label it as a DP *and* interpret it properly. Since the Wh string is first labeled a CP, the constraint on the locality of labeling requires that there be a D^0 merged with it to label it properly. Since (155) wants semantic information to be realized in the syntax, the D^0 should also be a syntactic representation of the semantic type shifter. This is formalized in (158), as what the child assumes when encountering a Wh string after a DP selecting Verb.

(158) If [V_D [what ...]] then [V_D [$_{DP}$ D^0 [$_{CP}$ what C^0 ...]]]

The outcome of (158) is the representation proposed above for children's FRs: the Silent D^0 analysis in (145), with a Wh-CP relabeled and type-shifted by a definite D^0 without phonetic content. The learner has gotten to this point simply by applying two basic properties of the structure building process of the narrow syntax to Wh strings in nominal positions. The above discussion shows that there is good reason to think that getting this far is quite easy for children. This formalizes the notion from Guasti and Shlonsky (1995) that Wh FRs are acquired early because the role of Wh movement in their derivation is clear, without necessitating the appeal to the Maturation of particular null operators as they do.

The last step - learning how to block Wh-NP in an FR - is more difficult, as it involves processes which are not basic to the syntax and are subject to crosslinguistic variation; and indeed, we have seen evidence that this step is the more difficult one, a step taken much later. But, the innate constraints on representations are still required to get the target representation, specifically the constraint requiring only heads to adjoin to other heads. When the child learns that Fill D^0_{def} must be satisfied for FRs, they may only posit that a **Head** is satisfying that. Since the child already has evidence that there is not a determiner like *the* used for this, they must assumed the Wh expression is doing so. So, they must assume that the Wh word moves as a head to adjoin to D^0, as shown in (159).

(159) If [$_{DP}$ [$_{D}$ what] [t_{wh} ...]] then t_{wh} is a Head trace.

This constraint blocks a representation where a Wh phrase moves to adjoin to D^0; this (again, along with the rule against left branch movement in English) blocks Wh-NP in an FR, the desired outcome.

We can thus see that the Minimalist approach here is quite fruitful in generating, consistent with the observed steps, the acquisition of FRs. We have also seen in this chapter evidence that the Minimalist approach has specific advantages over Usage Based approaches. In fact, it would in general be easy to assume, under such an approach, that FRs should be difficult to acquire in the first place, as they require an apparent structural polysemy - distinct from structural **ambiguity** in that there wouldn't be underlying syntactic structure disambiguating (like, say, a type-shifting D^0).

Proponents of Usage Based approaches often postulate that acquisition should begin with narrow generalizations of form-meaning mapping: constructions should be formulaic at first (Dabrowska, Rowland, and Theakston 2009), and "the scope of [their] productivity" should raise over time (Lieven 2010:2547). Generalizing the form of Wh strings to both questions and FRs relies on a more productive set of generalizations than headed RCs: if one is starting with a tighter form-meaning mapping, one should assume that there is greater utility early on in having only one semantic/communicative function for Wh strings, and assigning the function of FRs to a unique construction.

To put this clearly, a non-nativist learner relies on keeping things either **simple** in a broad, surface sense and/or **reliable**, in the sense of a close mapping between form and communicative function. In terms of simplicity, FRs and Headed RCs both rely on a similar dependency between phonological material and a "gap", and these are sensitive to similar constraints and difficulties typologically and in acquisition (O'Grady 1997, 2005). Thus, this sort of simplicity does not choose between the two forms of relativization, and does not predict their relative order of acquisition.

Thus, the other metric, reliability, ought to be what is at work. If it were, we would predict a favoring of Headed RCs over FRs; the naive generalization that a learner would make to get to an FR interpretation would be something like “interpret Wh strings as questions unless selected-for by a nominal position", which relies on a polysemy for Wh strings which is governed by a formal syntactic dependency (ie., selection). A generalization which is narrower, and thus more reliable and preferable under this system, would be something like “Assign Wh-gap dependencies question interpretations, and assign NP-gap dependencies nominal/relative interpretations." So by either metric, the learner should prefer RCs earlier on, which children cannot be said to do.

Conversely, a Minimalist learner (in the sense discussed here) prefers constructions which require the fewest analytic choices, can be most easily generalized across, and does not require the invention of new syntactic tools (like those used in RCs). Given the sort of learner proposed here, who uses the necessary constraints on Minimalist derivations to deduce targets of movement and null heads, we **do** expect FRs to be acquired first, rather than having the learner make guesses about NP head movement or head matching/deletion.

3.5.5 The right cues

Given the variation observed above, and the delays we observe in acquisition of the somewhat idiosyncratic determiner system of a language like English, we must determine what cues may exist in the input to signal the child to an English-like grammar.

A possible type of evidence as to the exact distribution of overt Determiners in English comes from *there* existential constructions. These, as has often been observed, show a “definiteness effect": definite descriptions and universal quantifiers cannot occur in these constructions (160).[11]

[11] See also Pérez-Leroux and Roeper’s (1999) discussion, with reference to bare nominals and scope ambiguities.

(160) a. There are dogs here

b. There are some dogs here

c. There are three dogs here

d. *There are the dogs here

e. *There are all dogs here

Given this, every instance of a sentence like (160a) will constitute evidence against a grammar where the lack of overt D allows definite interpretations. However, this does not get the child all the way there; we can conceive of (at least) three possible grammar types relevant to this: grammars which require all nominals to have an overt D (*[$_D$ Ø]), grammars which require all definite nominals to have an overt D (*[$_{D+def}$ Ø]), and grammars which lack overt D (*[$_D$ X]). Sentences like (160a) - and indeed any sentence with a bare NP - is evidence against the first type, but does not choose between the other two.

Conversely, sentences like (160b) are allowed in grammars where D^0 always must be overt (D→[+Overt]), or English-like languages where only some nominals require overt D (D$_{def}$ →[+Overt]), but not true “NP languages" (D→[-Overt]). So for this to be a helpful cue, the child must hear both of these to learn that *some* and a lack of overt D are both available for indefinites, and thus to arrive at an English like grammar. This is illustrated in Table 3.5.

Table 3.5: Existentials and determiners

	D→[+Overt]	D$_{def}$ →[+Overt]	D→[-Overt]
There are dogs	*	✓	✓
There are some dogs	✓	✓	*

Data from the MacWhinney (2000) corpus seems to show a relative paucity of existentials with no overt quantifier. Out of 49 plural existentials, there are 8 instances like

(161). It is unclear if this is enough input to give the child a clear choice between the options in Table 3.5.

(161) We just pretend there are witches.

This alternation could also be important in helping children to learn that *some* has a specific indefinite/new information reading rather than a simple weak quantificational reading.

In general, the cues for learning the correct distribution of null quantifiers (at least in English and Romance type languages) must be largely semantic; even in the case of existential constructions which are on the surface syntactic, the reasoning involves semantic knowledge. Thus, this is a case where learning involves attributing syntactic features based on semantic generalizations. That is, the feature value [+Pronounce] is associated in the adult grammar with the features [+D, +Definite].

Regarding FRs, a child must learn two different lexical items with these features: *the* and D^0_{rel}. In the adult grammar (by the analysis here), these are both [+D, +Def, +Overt]. But, they differ in respect to their selectional features: *the* selects an NP and D^0_{rel} selects a CP. So, full acquisition of the English determiner system involves choosing a grammar which generalizes [+Overt] to any lexical item which is [+D, +Def] (a Micro-Parameterization) over one in which only *the* is [+Overt] (a Nano-Parametrization).

The next chapter investigates another piece in the lexicon of English determiners, the Wh-ever morpheme, seen in Wh-ever Free Relatives, which is syntactically and semantically distinct from simple FRs. I will show that this semantic distinction, understood in terms of the system of parameterization of formal features discussed here, can give rise to at least some of the syntactic distinctions, particularly the availability of Wh-NP.

(162) Sebastian read [whatever **book** Charles gave him]

This discussion will seek to construct more clearly a particular path children may be taking in acquiring the adultlike grammar for English nominals and FRs by making syntactic generalizations based on Formal Features associated with semantic distinctions.

CHAPTER 4

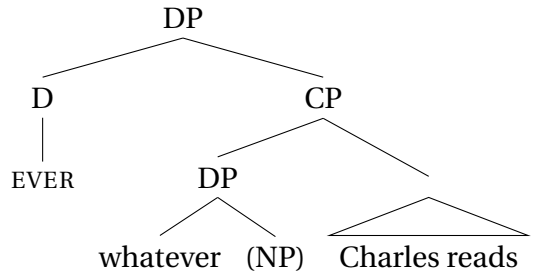
FREE RELATIVES OR WHATEVER

The discussion in the previous chapters has principally concerned the canonical, “simple" Free Relatives, which involve plain Wh words. As much of this has been focused on syntactic restrictions on simple FRs like *Wh-NP, it is necessary to discuss apparent exceptions to it. Another type of FR, those with “Wh-ever" morphology, allows a much wider variety of Wh expressions than the simple type, including Wh-NP.

(163) Sebastian will read whatever (book) Sebastian gives him.

This sort of exception, as well as other syntactic properties of “Wh-ever," raises the question of what representational differences there are between these and the “simple" FRs discussed previously. I will examine two sides of this: a semantic side and a syntactic side. However, the semantic aspects of Wh-ever will serve primarily as a diagnostic as to the extent to which different uses of Wh-ever can be morpho-syntactically united. I will argue that the syntactic exceptionality in (163) arises because the Wh word does not fill D^0 in Wh-ever FRs as it does in simple FRs, and that this is allowed because of semantic properties of the former - specifically that they are not canonical definite descriptions. The resulting syntax is in (164), with the Wh expression remaining in the specifier of CP and a null EVER head in D^0 (but see discussion below for some viable alternatives).

(164) *Whatever* stays in [Spec,CP]



The chapter is structured as follows: The first section discusses the basic distributional properties of Wh-ever FRs, their surface syntactic differences from simple FRs, and their semantics, focusing on the distinction between the Free Choice reading and the Ignorant reading of Wh-ever; I will identify semantic properties which may serve as diagnostics of the syntax of Wh-ever in different semantic contexts. The second discusses the proposition, originating in Battye (1989), that Wh-ever FRs are in fact “Pseudo-Free Relatives" and have the syntax of headed RCs rather than FRs, with special attention to Italian. The third section will argue against the Pseudo-FR analysis, and that at least for English Wh-ever, the syntax of Wh-ever can be unified with that of simple FRs, but I will find spaces where the Pseudo-FR analysis must **also** be available - what could be called an “Italian corner" of the English lexicon. The fourth section presents data on the acquisition of Wh-ever, and discusses the relationship of the findings to the present syntactic analysis; I find that children’s interpretations of Wh-ever seem more like universal quantifiers than definite descriptions, and that similar patterns can be seen in adult behavior.

4.1 Wh-ever: the facts

The most striking difference between Wh-ever and simple FRs, and the most important one for the current discussion, is the much wider variety of Wh expressions which appear in the former. In fact, almost every type of Wh expression blocked in FRs is avail-

able for Wh-ever (165). Exceptions to these are that PP pied piping, Heavy pied piping, and FRs headed by *why* are unavailable for Wh-ever FRs as for simple FRs (166).

(165) a. Sebastian will talk to [whoever Charles brings to the party]

b. Sebastian will drink [however many bottles of wine Charles has]

c. %Charles will visit [whoever's house Sebastian is staying in]

(166) a. *Sebastian will greet [to who(m)ever Charles talked]

b. *Sebastian will leave [whyever Charles is crying]

c. *Sebastian will read [whichever book the author of which Charles knows]

The badness of the sentences in (166a-b) can be assumed to be derived the same way as in their equivalents, for the reasons discussed in Chapter 2: requirements on case connectivity blocking PP pied piping in FRs and the selectional properties of rationale clauses blocking *why* FRs. The fact that (166c) is bad seems prima facie like it should fall out from the fact that there is a Wh word within the FR which is not heading it - a fact which will become important in the discussion below. The facts in (165) must be understood in some sense by similarities and differences between simple FRs and Wh-ever FRs.

Another difference is that, unlike simple FRs, Wh-ever does not require a clausal element.

(167) a. *Sebastian will eat what

b. Sebastian will eat whatever

Lastly, and something that makes Wh-ever FRs really special among English relatives of all types, is that they allow relativization out of It-Clefts, something impossible in any other type of English relativization. Like many other properties discussed here, this creates a distinction between other relative clauses and Wh questions (169).

(168) Clefts in relatives

a. *Sebastian will eat the thing it is that Charles is making

b. *Sebastian will eat what it is that Charles is making

c. Sebastian will eat whatever it is that Charles is making

(169) Clefts in questions

a. What is it that Charles is making?

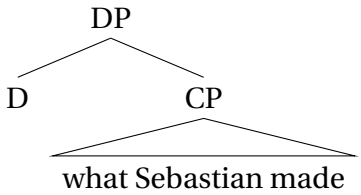
b. Sebastian wonders what it is that Charles is making

While the goal of this chapter will be primarily to derive the facts in (165-167), the availability of clefts just in these relatives is worth noting particularly in how it interacts with the semantics of Wh-ever, as discussed below.

4.1.1 The two readings of Wh-ever

Recall the assumptions from the previous chapters that simple FRs have the semantics of definite descriptions, which are derived syntactically (and perhaps semantically) from Wh clauses of the type seen in questions. The analysis of FRs Caponigro (2003) and Chierchia and Caponigro (2013) states this as meaning that an FR denotes the entity which is an answer to the question, or the entity which has the “Topic Property" of the question. The denotations in (170) show this; it can be assumed here that the Movement of *what* to D is either post-syntactic (as is often assumed of Head Movement, and discussed in Chapter 2) or reconstructs at LF; it does not seem to be the case that *what* denotes the definite operator, but rather is only a morpho-phonological exponent of it.

(170) LF Composition of simple FR



a. ⟦CP⟧= λP$\exists$x[P = λw[Sebastian made x in w]]

b. ⟦D⟧= λQ[ιx[TP(Q)(x)(w)]]

c. ⟦DP⟧= ιx[TP(what Sebastian made)(x)(w)]

= ιx[Sebastian made x in w]

There are two different readings for Wh-ever FRs: it can be read either as a type of Free Choice item or to refer to a specific entity with the property denoted by the clausal element.

(171) Charles will read whatever Sebastian is writing

a *Ignorance*: I don't know what Sebastian gave him, but Charles will read it

b *Free Choice*: If Sebastian is writing something, Charles will read it

Various attempts have been made to unite these under a uniform semantics; Dayal (1997) asserts that they can and must be.

(172) ⟦whatever$_i$ [$_{IP}$ ... t$_i$ ...] ⟧= λQ $\forall$i-alternatives $\in$ f(w)(f) [Q(i)(ιx[P(i)(x)]]

Essentially, this says that while Wh-ever FRs are definite descriptions of a sort, *ever* introduces a modal element - a quantification over alternatives. The referent of Whatever(P) is the thing x for which P(x) is true in each alternative world in the speaker's belief state. When the alternatives are possible choices, we get a Free Choice reading; when they are simply entities, an Ignorance reading.

A similar analysis from von Fintel (2000) puts ignorance into the denotation as a presupposition.[1] His denotation is in (173); its most important distinction from (172)

[1]He also has a similar analysis for a "indifferent" reading of Wh-ever. However, his examples don't seem to form a class which is distinct from both the Free Choice reading and the Ignorant reading:

(i) I grabbed whatever tool was handy

(ii) Zack simply voted for whoever was at the top of the ballot

comes from the fact that the quantification across worlds comes as part of the presupposition of Wh-ever, and is not asserted.

(173) ⟦whatever⟧= λP: [$\exists$w',w" $\in$ F: ιx. P(w')(x) $\neq$ ιx. P(w")(x)] . ιx. P(w)(x)

That is, it is presupposed that in at least some worlds in the modal base F, the x for which P(x) is true at that world is distinct from the one in every other world.

4.1.2 Properties of the two readings

We can examine the differences between Free Choice and Ignorance Wh-ever by looking at their different properties in environments that do not license FCIs - specifically *episodic* environments (Giannikidou 2001, Dayal 2013). This is demonstrated in (174): a perfective sentence blocks FCI *any*, but allows Wh-ever. However, the possible meanings of Wh-ever are restricted.

(174) a. *Charles made anything Sebastian is eating

b. Charles made whatever Sebastian is eating

The sentence in (174a), with *anything,* is surface ungrammatical as *anything* has no licit reading in this context. The sentence in (b) is grammatical, but a Free Choice reading is blocked for the same reason that *anything* is blocked in (a). Thus, all that remains here is the Ignorance reading. Here I will use this distinction to look at some properties of each of the two readings.

Dayal (2013) defines three types of Free Choice *any* which are licensed in different environments: partitive, unmodified, and subtrigged (that is, where *any* heads a Relative Clause with a modal element). Each of these is illustrated in (175), from Dayal (2013).

While indifference is implicated in these examples, it does not seem necessary to derive it from a separate denotation. I will work under the assumption that indifference in these cases can either arise from an implicature of a sort (which seems to be the case in (ii)) or can be accounted for by the implications of indifference that come from typical uses of FCIs.

(175) a. **Partitive**:

Bill may read any of these books.

b. **Unmodified**:

Any student can attend this event.

c. **Subtrigged**:

Bill may read any book he finds.

We can see the ways in which FCI Wh-ever is similar to other FCIs by looking at Dayal's typology of environments, shown in Table 4.1.

Table 4.1: FCI licensing environments

	Generic	Possibility	Necessity	Episodic
Partitive *any*	*	✓	*	*
Unmodified *any*	✓	✓	✓/*	*
Subtrigged *any*	✓	✓	✓	✓

On the surface, it appears that Wh-ever FRs have the distribution of Subtrigged *any*, occurring in all four environments.

(176) a. Whoever goes to school here works hard

b. Sebastian may read whatever books he finds/you have

c. Sebastian must read whatever books you have

d. Sebastian has read whatever books you have

However, not all of these make both interpretations of Wh-ever available; specifically, only the first two allow the Free Choice reading. The only possible reading of (176c) is that there are some particular books you have, the speaker is not aware of what they are, and Sebastian has to read them, and (d) has to mean that there is some particular

set of books you have and Sebastian read those books. On the other hand, (176a) can have either the ignorant reading (I don't know who goes to school here exactly, but they work hard) or the Free Choice reading (if someone goes to school here, they work hard), and similar for (b).

So we see that non-Free Choice environments force the ignorance reading for Wh-ever. Another context which seems to do so is Wh-ever FRs with an It-cleft, a fact alluded to but not discussed in any detail in von Fintel (2000). this is seen by the infelicity of (177a) compared with (b).

(177) a. #Every day, Charles talks to whoever it is that is outside his door that day

b. Charles is talking to whoever it is that is outside his door right now

The use of *every day ... that day* forces a Free Choice interpretation, but this appears to be blocked by the use of the cleft. Conversely, the anti-Free Choice, episodic context in (b) allows the cleft. The reason for this is unclear, but may have something to do with the semantic contribution of the cleft.

4.1.3 Non-argument Wh-ever

Unlike simple FRs, Wh-ever FRs' distribution is not limited to argument positions; they have a sentence-initial, conditional-like use (cf. Rawlins 2008). Note the contrast in (178).

(178) a. Whatever Sebastian makes, Charles will be happy

b. *What Sebastian makes, Charles will be happy

This more closely resembles a correlative in some ways than a typical English FR. It differs from correlatives in a language like Hindi-Urdu in that it does not typically have a corresponding pronoun in the main clause which is restricted by the adjoined correlative, though it may have something like this.

(179) a. Hindi-Urdu Correlative

*[jo Sebastian banātā hai] Charles *(us-se) khush hogā*

REL Sebastian makes PRS Charles it-with happy will.be

'What Sebastian makes, Charles will be happy with (it)'

b. Whatever Sebastian makes, Charles will be happy (with it)

FCI *any* has a somewhat similar use, though it is ungrammatical without *if* (similar to "Unconditionals" discussed in Rawlins 2008). Giannikidou and Cheng (2006) point out that this also seems to give a somewhat different reading from conditional Wh-ever, as the latter may have a presupposition of existence.

(180) Conditional FCIs (Giannikidou and Cheng 2006:157)

a. If any student calls, I am not here

b. Whichever student calls, I am not here

Notably, the example in (180) could just be a Negative Polarity reading of *any*, which would of course answer the question of why it is different than the Free Choice reading in (b).

4.2 Simple FRs and Pseudo-FRs

A lot of early literature on the topic takes Wh-ever relatives to be a variant of FRs; however, a number of analyses of them propose a syntax that has more in common with headed relatives than with the analysis presented in previous chapters of simple FRs. Here I will examine a recent approach to Wh-ever FRs which gives this division.

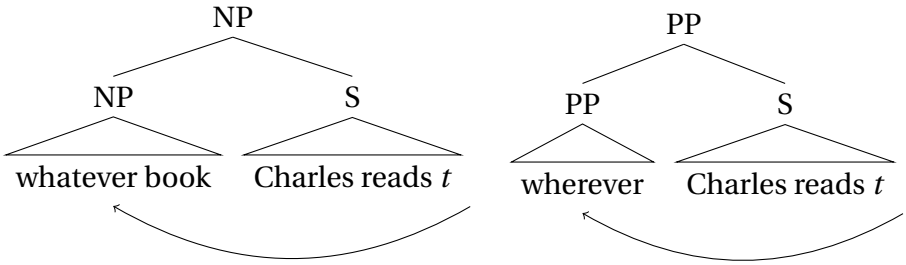
4.2.1 The loci of Wh-ever Words

The analysis of Battye (1989) claiming that the Italian equivalent of Wh-ever relatives have the syntactic properties of Headed RCs rather than of simple FRs is a depar-

ture from earlier literature which generally takes Wh-ever relatives to be FRs, and in fact perhaps the canonical FRs. It is thus best understood in context of these analyses.

Older accounts which represent simple and Wh-ever FRs the same way were divided into two broad camps, sometimes called the Head Analysis and the Comp Analysis (Groos and van Riemsdijk 1981). The Head analysis has the Wh expression occupying the position whose category projects onto the entire FR; this originates in the analysis of FRs in Bresnan and Grimshaw (1978), who propose that a **phrasal** Wh expression may project after moving in this way (181; this is done to account for categorial "matching" effects, where the category of the Wh expression seems to match that of the FR (however, see §2.2.1.2 in chapter 2 for apparent exceptions to this).

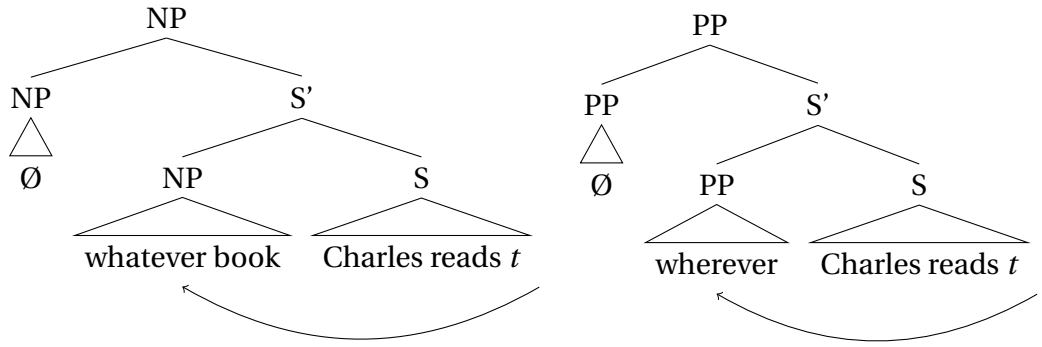
(181) Head Analysis of FRs



This has in common with the analysis of simple FRs here as well as those in Donati (2006) and Cecchetto and Donati (2015) that the Wh expression moves to a position where it may project its categorial label onto the entire FR. However, it lacks the assumption that only simplex Wh expressions may project (and thus cannot account for the restrictions on the form of FRs discussed in Chapter 2, such as *Wh-NP).

The Comp analysis, as given in Groos and van Riemsdijk (1981), asserts that the Wh expression sits in the position within a clause where Wh expressions otherwise sit (COMP in earlier terms, [Spec,CP] in more current terms), and this clause is sister to a null category. Groos and van Riemsdijk account for **variation** in whether languages require categorial matching based on a parameter as to whether, essentially, this position may project; in English it cannot.

(182) Comp analysis of FRs



Without delving into the details of Groos and van Riemsdijk's argument, the crucial point of this analysis is that the Wh expression occupies the specifier position of a clause, and a null syntactic object projects its category on the FR. This is most similar to the syntax of children's FRs proposed in Chapter 3. For the current discussion, we could take the crucial distinction between the two to be whether the Wh expression is in the position which projects onto the entire FR; the Head Analysis claims it is, whereas the Comp analysis claims it is not.

These older accounts assume the same representations for both simple and Wh-ever FRs: either both have a Head analysis or both have a Comp analysis. However, some surface syntactic differences between these types have led some scholars to posit that there are in fact (at least) two types of syntactic representations for what we traditionally call FRs, distinguishing Wh-ever from simple FRs.

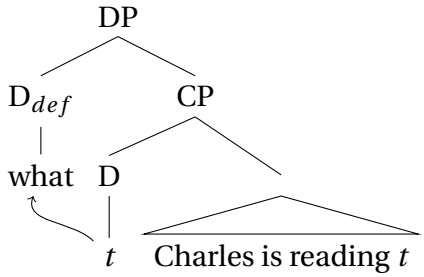
One way this has been done has been to divide analyses between simple FRs, which we might call FRs proper, and "Pseudo" Free Relatives, alluded to in the previous chapter. This idea was introduced by Battye (1989), and revisited by Cecchetto and Donati (2011, 2015).

This analysis assumes that syntactic properties of -ever type FRs are due to these having the representation of headed relatives, with a potentially-null nominal head with a CP complement, under a quantifier which only on the surface resembles a Wh word-qua-Wh word, but does not itself undergo movement; definite FRs on the other hand,

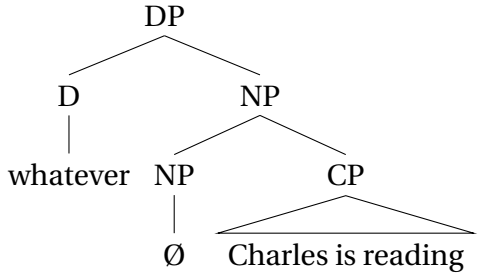
are headed by a regular Wh word which has moved. The trees in (183) show the two representations, based on the analyses in the previous chapters.[2]

(183) Representations of *what-(ever) Charles is reading*

a Simple FR



b Pseudo-FR (Headed RC)



In a sense this is a compromise between two sides of an older debate on what the head of any FR is: whether the Wh word which seems to move is the head of the relevant entire FR, or there is a null nominal, which acts as the notional head (see discussion in Groos and Van Riemsdijk 1981). By this analysis, standard, definite FRs are headed by a moved Wh word, and lack an external NP head, whereas Wh-ever relatives are headed by an external NP head which may be null.

4.2.2 Pseudo-FRs in Italian

Italian possesses a series of Free Choice Items which, when used in relative clauses, resemble FRs due to their Wh-like morphology. These are *qualunque*, *qualsiasi* (which

[2]It should be noted that there is still movement within the CP, according to this analysis; it is whatever type of movement usually occurs in headed RCs. The particular analysis of this movement does not particularly affect this; see the discussion in Appendix A.

are both typically glossed as 'whatever'), and *chiunque* ('whoever'). Battye (1989), in arguing that these require the Pseudo-FR analysis, points to several syntactic patterns exhibited by these which are not possible with Italian simple FRs headed by *quanto* ('what) or *chi* ('who'). For simplicity, I will refer to these Italian FCIs as QQC.

(184) QQC Relatives (Battye 1989:219-220)

a. *Viaggerò con [chiunque tu voglia invitare]*
I.will.travel with whoever you want invite.INF
'I'll travel with whoever you want to invite'

b. *Ti spiegherò [qualunque problema tu non capisca]*
to.you I.will.explain whatever problem you not understand
'I'll explain whatever problem you don't understand to you'

One of the ways Battye discusses in which QQC relatives differ from simple FRs is that QQC can occur with out a clausal element, which is not possible for simple FRs.

(185) Bare QQC (Battye 1989:226)

a. *Lei parlerebbe dei suoi problemi amorosi con [chiunque Ø]*
she speak.COND of her problems amorous with whoever
'She would speak about her love problems with anyone

b. *A questo punto mi piacerebbe leggere qualunque libro Ø*
at this point me.DAT please.COND read.INF [whatever book
di fantascienza
of science fiction]
'At this point I would be happy to read any science fiction book'

c. **Lui avrebbe parlato con [chi Ø]*
He have.COND spoken with who
Intended: He would have spoken with the person

d. **[Quanto Ø] è vero*

what is true

Intended: That is true

In (185a-b), there is no clausal element following the QQC quantifier, and the sentences are acceptable; in (c-d), no clausal element follows the Wh word and the sentences are ungrammatical. Thus, we could say that QQC can be "bare" in this sense, whereas the Wh words used in simple FRs cannot. This is a feature which QQC shares with other determiners, or other nominals in general, but not with FRs: Any given quantifier is not restricted to occurring in a relative structure, but the Wh determiners used in FRs are so restricted.

A second difference is that overt complementizers and relative pronouns can occur with QQC relatives, but not simple FRs. Similarly, Wh Relative Pronouns are available for only the former.

(186) Complementizers with QQC (Battye 1989:229-230)

a. *[Qualsiasi confezione* ***che*** *non dia soddisfazione] sarà sostituita*

Whichever product **that** not give satisfaction be.FUT replaced

'Any product that is unsatisfactory will be replaced'

b. *Mi sarei incontrato con [qualsiasi rappresentante* ***da cui***

Me would've met with whichever representative **from whom**

tu l'-avessi comprato]

you it-had bought

'Me, I would have met with any representative you had bought it from'

(187) Complementizers with simple FRs (Battye 1989:230)

a. **[Chi* ***che*** *viene] troverà la porta chiusa*

Who **that** comes find.FUT the door closed

Intended: The person that comes will find the door closed

b. **Bado a quanto* ***a cui*** *badi tu*

pay.attn-1st to [what **to which** pay.attention-2nd you]

Intended: I pay attention to what you pay attention to

Again, a feature shared by QQC relatives and regular headed relatives (the availability of complementizers and relative pronouns) is unavailable for simple FRs.

A third distinction drawn concerns the availability of certain infinitival clauses in simple FRs but not with QQC.

(188) Infinitival (free) relatives (Battye 1989:228-229)

a. *Cerco [chi mandare al mio posto]*

I.seek who send.INF in my place'

'I am looking for someone to send in my place'

b, **Cerco [qualsiasi libro comprare]*

I.seek whatever book buy.INF

Intended: I am looking for whatever book to buy

These data all show properties which are quite different between QQC relatives and simple FRs in Italian, supporting Battye's analysis of these as syntactically akin to headed RCs rather than FRs ("Pseudo-FRs"). The analogy made by Cecchetto and Donati (2011, 2015) to English Wh-ever relies on similar facts holding in English. But, this point-by-point scrutiny of this analogy raises some issues with this account.

4.2.3 English and Italian

The extension of the Pseudo-FRs to English proposed by Cecchetto and Donati (2011, 2015), used to explain the absence of *Wh-NP effects with Wh-ever, relies on divergence in forms between simple FRs and QQC relatives in Italian being seen in English.While we can clearly see some parallels between English and Italian in this regard, a closer

look at differences between the two languages, and a critical reexamination of the data presented by C&D, raises questions about this extension.

As with QQC, a piece of evidence presented in favor of the Pseudo-FR analysis in English is the availability of Wh-ever expressions without an overt clausal restrictor in English and Italian. Call this “Bare Whatever", though this may involve any Wh-ever words as well as Wh-ever NP sequences (189), first observed for English by Bresnan and Grimshaw (1978).[3] This, of course, contrasts with simple FRCs, which do not allow bare Wh words under any circumstances (190).

(189) a Sebastian will eat whatever (fruit)

b Sebastian can leave whenever

(190) a *Sebastian will eat what (fruit)

b *Sebastian can leave when

While the sentences (189) are syntactically well-formed and have a reading similar to other Wh-ever sentences which include a clausal element, the sentences in (190) could be grammatical only as questions; (190a) could not mean *Sebastian will eat that (fruit)*.

Other comparisons between English and Italian are murkier. The Infinitival construction seen in Italian discussed by Battye (1989; example 188 above) does not have a direct equivalent in English, but there is a similar infinitive reduced relative which seems to make a distinction between FRs and Wh-ever relatives. However, this similarity with Italian becomes more complicated: The construction is marginal with simple FRs, and wholly ungrammatical with Wh-ever, but perfectly good with headed RCs.

[3]However, see below a discussion of the example they give, which is somewhat distinct from those given here.

(191) Infinitival FRs

a. ?Bordeaux is what to drink at a dinner party

b. *Wines from Bordeaux are whatever to drink at a dinner party

c. This Bordeaux is the wine to drink at a dinner party

The pseudo-FR hypothesis itself would predict that, given that (191c) is grammatical, (a) should be worse than (b) - but the opposite is true. The reasons for this are complicated: (191a) seems to be most akin to a Specificational Pseudo-Cleft, which on the surface resembles an FR in a copular sentence, but may underlyingly more similar to an embedded question (Schlenker 2003). This accounts for the fact that the relevant string in (191a) is acceptable in an embedded question but not in another FR frame (192).

(192) a. Sebastian wondered what to drink

b. *Sebastian brought what to drink

Since Wh-ever words are ungrammatical in most embedded questions, and in particular infinitival questions, it follows that (191b) is ungrammatical. As headed RCs are in any case syntactically distinct from simple FRs and pseudo-Clefts, the representation of the infinitival RC in (191c) is an unrelated issue.

Compared with Italian, where QQC freely occurs with complementizers and relative pronouns, in English, Wh-ever can marginally occur with a complementizer but not with a Wh relative pronoun. This is in stark contrast to QQC in Italian.[4]

[4]Cecchetto and Donati (2015:52) claim that these are good for English, and in fact suggest that a Wh relative expression can be included, as in Italian (i). I personally find this to be completely ungrammatical.

(i) ???Whichever boy to whom I speak keeps telling me the same thing

I will argue based on my own grammar for the moment that this is not a general fact about English, and also maintain that even the complementizer in the above example is marked-at-best for English, unlike Italian.

(193) a. %Charles wrote whatever that Sebastian is reading

b. *Charles wrote whatever book which Sebastian is reading

c. *Charles brought whichever student who Sebastian liked

Upon closer inspection, then, it seems that the argument for Pseudo-FRs in English based on comparison with Italian is not entirely perfect. We are left with basically two compelling pieces of data which equate Wh-ever more with headed RCs than with simple FRs: The wider array of Wh expressions allowed with Wh-ever, and Bare Whatever. The former is in fact the problem which Cecchetto and Donati (2011, 2015) are attempting to solve with this analysis, rather than a solution to it. The latter will require some further examination.

4.3 The syntaxes of Wh-Ever

I will argue here that Wh-Ever FRs can indeed be analyzed as a type of FR rather than a pseudo-FR, and that the relationship between the Wh expression and the determiner which carries the -ever morpheme can explain the absence of *Wh-NP. However, I will argue that this syntax doesn't do quite enough, and that the pseudo-FR analysis for English is not entirely wrong: The grammar must allow for a pseudo-FR analysis of Wh-ever as a sort of last resort to repair specific constructions, including Bare Whatever and some other options. This analysis accounts for restrictions on the readings of Bare Whatever, and other cases where one reading of Wh-ever is blocked.

4.3.1 Properties of Bare Whatever

The most striking fact about Bare Whatever compared to other uses is that only the Free Choice reading is possible. Recall that the different readings can be identified partly on whether or not they occur in an environment that licenses Free Choice Items generally. In FCI environments, Wh-ever FRs are typically ambiguous between the two readings (194). In these same environments, Bare Whatever is available, but can only take

an FCI reading, not an ignorance reading (195). In non-FCI environments it is ungrammatical on either reading, in contrast to standard definite descriptions (196). This shows that Bare Whatever is strictly limited to the FCI reading.

(194) Sebastian read whatever Charles gave him

a. → In the past, Sebastian generally read anything Charles gave him

b. → Sebastian read the thing Charles gave him, whatever it is.

(195) Sebastian will read whatever

a. → Sebastian will read anything, in general

b. ↛ Sebastian will read whatever that thing is

(196) a. *Sebastian has read whatever

b. Sebastian has read this thing (though I don't know what it is)

4.3.1.1 Bare Whatever and Free Choice environments

Aside from the fact that Bare Whatever bars the Ignorant reading of Wh-ever, Free Choice Wh-ever and Bare Whatever also have different distributions, in terms of the typology of Dayal (2012). Recall from above that, Wh-Ever has a similar distribution, to Subtrigged *any* FCIs. Conversely, Bare Whatever has the distribution of unmodified *any*; It is freely available in generic and possibility modal environments (197a-b), it is unavailable in episodic contexts (197d), and in the case of necessity modals (197) is limited to readings where the necessity is interpreted as referring to a generic set of events ('It must generally be the case that Bill will work with whatever student').

(197) a. Bill works/will work with whatever student

b. Bill can work with whatever student

c. Bill must work with whatever student ?[at any given time]

d. #Bill worked with whatever student

Thus we see that Bare Wh-ever is barred in the same cases as unmodified *any* FCIs, namely in episodic contexts. It should be noted that not all speakers allow Bare Whatever productively in this way, but for those that do this is the most natural distribution.

The third type of FCI Dayal discusses is Partitive (*any of the...*). This construction is marked with *whatever*, but fine with *whichever*. It is somewhat marked without the clausal element.

(198) a. Bill will work with whichever of his students ?(comes to class first).

b. Bill can work with whichever of his students ?(comes to class first).

c. Bill must work with whichever of his students *(comes to class first).

d. Bill worked with whichever of his students *(came to class first).

To the extent partitive Wh-ever is available without the clausal component, it seems to have a distribution similar to Bare Whatever; with necessity modals and in episodic contexts, it is marked in its "bare" form. With its clausal component, it has the same distribution as FR Wh-ever.

This gives us an updated typology of Free Choice environments, shown in Table 4.2.

Table 4.2: Wh-Ever and FC environments

	Generic	Possibility	Necessity	Episodic
Partitive *any*	*	✓	*	*
Unmodified *any*	✓	✓	✓/*	*
Subtrigged *any*	✓	✓	✓	✓
Partitive Wh-ever	✓	✓	*	*
Bare Whatever	✓	✓	✓/*	*
Wh-ever FR	✓	✓	✓	✓

Looking only at these semantic facts, we might conclude that this pattern can be attributed to the simple fact that Bare Whatever is equivalent to *any* FCIs without a clausal restrictor, which is what the Pseudo-FR analysis of Wh-ever in English would predict. However, as this analysis is intended to capture the syntactic distribution of Wh-ever, it is important to note that in the case of Wh-ever FRs, the occurrence of complex Wh in Wh-ever is not reliant on semantic subtrigging by embedded mood/aspect in the same was as subtrigged *any*. The contrast in (199) shows that, under a necessity modal, *any* is unavailable when it embeds an episodic clause, in the same context Wh-ever is available embedding the same clause, though it is restricted to the Ignorant reading of Wh-ever.

(199) a. Sebastian must read whatever book Charles gave him

b. #Sebastian must read any book Charles gave him

Thus we see that the *syntactic* distinction between Wh-ever and standard FRs cannot simply be derived by taking all instances of Wh-ever as analogous to relative clauses headed by *any*.

4.3.1.2 Bare Whatever and implicit restrictors

A second type of evidence that Wh-ever and Free Choice *any* are distinct from each other comes from the availability of covert, anaphoric restrictors for different types of DPs. A typical DP can be be interpreted as restricted by some clausal material from previous discourse. Simple FRs bar this.

(200) Standard DPs

Sebastian brought many of the drinks at the party ...

Charles tried all of the red wines (that Sebastian brought).

(201) Simple FRs

Sebastian brought a lot of the guests that were at the party, as well as a lot of the drinks ...

Charles enjoyed all of what *(Sebastian brought).

Wh-ever expressions show mixed effects in this regard, and it again correlates with the available readings, shown by the contrast in (202). As with the sort of Bare Whatever seen above, the case of Wh-ever with an elided clausal restrictor does not allow for the Ignorant reading, only the Free Choice reading.

(202) a. Charles will love whatever apples I bring, but hate whatever bananas

b. *Charles ate whatever apples I brought, but hated whatever bananas

This may be because the Free Choice reading allows Bare Whatever, but the Ignorance reading - which is forced in (202) - does not. This suggests that the sentence in (202) doesn't involve elision of a clausal restrictor, but Bare Whatever with an implicit restrictor of the type in (201).

Another peculiarity of this type of conjunction is that it cannot be conjoined with a sentence including a QP with a covert restrictor and a *different* quantifier.

(203) a. Any book Charles writes will be successful, but {any/*whatever} article will be unpopular

b. Every book Charles writes will be successful, but {all, some} of the articles will be unpopular

This is in fact similar to a pattern in similar conjoined sentences using adjunct FRs, which are marked-to-ungrammatical when contrasted with a locative PP (204).

(204) ?Charles vacationed near the place Sebastian was born, not where Julia was

These facts about conjunctions, implicit restrictors, and quantifiers, all show ways in which Wh-ever is distinguished from other quantifiers, including Free Choice *any* - in particular when the Ignorant reading is forced. It seems to show that, where it is not syntactically “bare", Wh-ever strongly needs to be associated with a clausal element - a property of Wh words in simple FRs.

4.3.2 Extraction and Complementizers

Wh-ever FRs are distinct from other Relatives in that extraposition of the clause from its head, and intervening material like ‘in the world’, are marginal with Wh-ever but not with headed RCs with *any*. Intervening material like ‘in the world’ is particularly bad in the absence of a head noun.

(205) a. Charles will give any book to Sebastian that he reads.

b. ?Charles will give whatever book to Sebastian that he reads.

(206) a. Charles will read any book in the world that Sebastian gives him.

b. ?Charles will read whatever book in the world Sebastian gives him.

c. *Charles will read whatever in the world that Sebastian gives him.

Something we find is that, to the extent that extraposition is acceptable with Wh-ever relatives, it is not good without a complementizer, and it is restricted to FCI readings.

(207) a. Charles will give whatever book to Sebastian *(?that) he reads.

b. *Charles gave whatever book to Sebastian (that) he read.

As with Bare Whatever, we see from these examples a type of Wh-ever relative which looks like an FR, and a type which looks like a Headed RC.

4.3.3 Possessives in Wh-ever FRs

Looking at possession in Wh-ever FRs shows us more cases where the analogy with headed RCs becomes difficult. Specifically.

(208) Possession in Wh-ever FRs

a. Sebastian will talk to whoever's book he admires.

b. *Sebastian will read whoever's book he admires.

c. *Sebastian will talk to whichever author whose book he admires.

(209) Possession in *Any* RCs

a. *Sebastian will talk to anyone's book he admires.

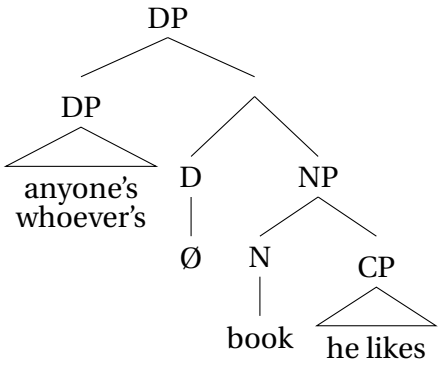
b. Sebastian will read anyone's book he admires.

c. Sebastian will talk to any author whose book he admires.

While (208a) is somewhat marginal, the FR clearly can only refer to the book, as shown by the comparison with (b). The Pseudo-FR analysis requires that the Wh-ever word be in the position of the D^0 above the clause; comparison with (208c) shows that putting *whichever* in that higher position blocks the lower possessive Wh expression.

When comparing these with RCs with *any*, which in the Pseudo-FR analysis are equivalent, we see that none of these facts do not hold. An expression like *anyone's book* can only refer to the book, not the author, unlike with *whoever's book* (209a-b), and the clausal element may freely include a Wh expression.

(210) Possessive RCs in Pseudo-FR analysis



This representation gives the proper interpretation for *anyone's book*: the nominal head is *book*, and the expression can be the complement of *read*, as in (209b). However, if the same representation is given to *whoever's book*, we get the wrong interpretation, as the resulting string can only refer to the possessor and not the book (208a-b).

Another case where possession gives clues about the structure of Wh-ever is cases of Heavy Pied Piping. In cases like (211), it appears that heavy pied piping only allows the FCI reading, as shown by the infelicity/ungrammaticality of (b).

(211) Heavy Pied Piping in Wh-ever Relatives

a. Charles will read whatever book the author of which is good-looking.

b. *Charles read whatever book the author of which was Sebastian.

The meaning of (211a) can only be that Charles will read any book as long as it was written by a good-looking person, a Free Choice reading. The aspect of (b) doesn't allow for an FCI, and so there is no available reading for the sentence and it is ungrammatical. This is another case, like Bare Whatever, where a syntactic form blocked in simple FRs is available with Wh-ever; simple FRs cannot have heavy pied piping, as seen in (212).

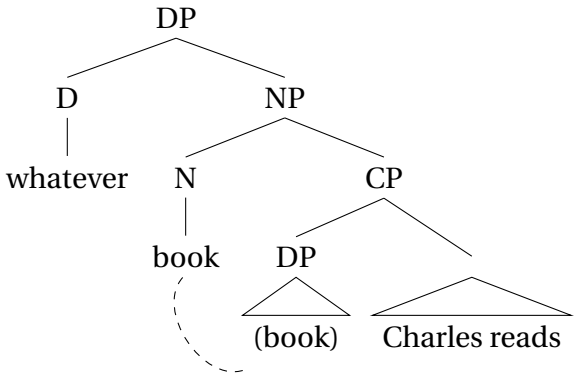
(212) *Charles read what the author of which was good-looking.

Like Bare Whatever, then, Heavy Pied Piping is seemingly syntactic evidence in favor of the Pseudo-FR analysis, but also a case where only the Free Choice reading is allowed, and not the Ignorant reading.

Syntax similar to that of a simple FR blocks heavy pied piping, as both the Wh word and the pied-piped expression would be targeting the same position, the specifier of CP. So, the availability of a sentence like (211a) would seem to need a Pseudo-FR representation. Given these data, as well as the Bare Whatever data, a first guess about the representation of Wh-ever FRs might be that Free Choice Wh-ever has Pseudo FR syntax (that is, headed RC syntax) and ignorance Wh-ever has a more FR-like syntax.

However, the sentences in (208) seem to show that Free Choice Wh-ever cannot always have Pseudo-FR syntax. So a revised guess might be that Pseudo-FR syntax is a sort of last resort: Wh-ever typically has a syntax like a simple FR, but in certain cases it can be forced to have a representation more like a headed RC (213) to accommodate things like a lack of clausal restrictor, heavy pied-piping, or extraposition, and in this case Wh-ever must be the Free Choice variety.[5]

(213) Pseudo FR = RC



4.3.4 A Grammar that can do both

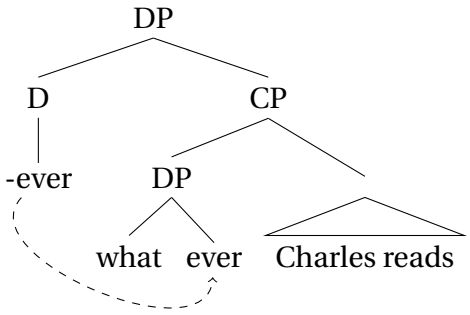
To account for all of the data presented here, I propose that there are in fact two different syntaxes associated with Wh-ever: an FR syntax, where the basic properties of simple FRs are carried over, and a Pseudo-FR syntax, where the Wh-ever word is simply a determiner akin to *anything*.

[5] See Appendix A for an argument for this specific RC syntax.

The FR syntax for Wh-ever modifies the simple FR syntax slightly. It shares the property that D^0 selects a Question CP and semantically changes it into a referential expression, but as the D head is not a canonical Definite Description, it does not require head movement of the Wh word, and thus does not give rise to *Wh-NP or other effects.

As with the Pseudo-FR analysis, The framework here requires that the Wh expression not occupy the D^0 head of a Wh-ever FR. There are two places it could be: in the specifier of the CP which is sister to D^0, or in the specifier of the entire DP. The D^0 must in some sense contain or select for Wh-ever morphology; one way to capture this is to say that the *-ever* morpheme originates in D and is expressed on the Wh word by a type of morphological lowering. This is shown in (214). A second option is that Wh-ever originates on the Wh expression, and Agrees with a D^0 with a feature [EVER], and the Wh expression moves to the specifier of the larger DP (216).

(214) *-Ever* moves downward



This analysis is similar to one mentioned, but not elaborated upon, by Donati (2006). She alludes to this as a version of the Pseudo-FR analysis, essentially, stating that the clausal element is “generated as the complement of an external determiner" *-ever* (Donati 2006:41, note 10), though she posits that the morphology involves *what* moving, not *ever*, as illustrated in (215).

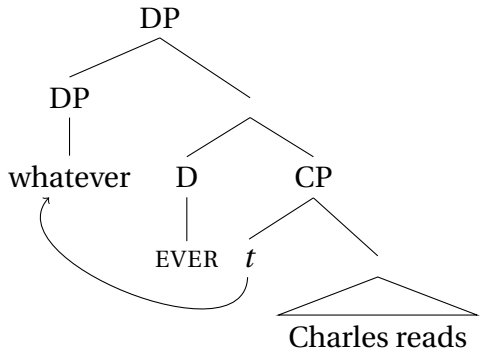
(215) Charles will read [$_{DP}$ [$_{D}$ ever] [$_{CP}$ [$_{DP}$ what book] ...]]

While Donati describes this as a version of the Pseudo-FR analysis, it is more akin to FRs as defined here in that the Wh clause is relabeled and semantically type-shifted

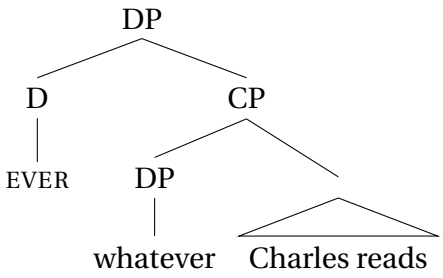
by an external D^0 which enters into a morpho-syntactic relationship with the Wh word. This is distinct from the idea of a Pseudo-FR as described in Battye (1989) and Cecchetto and Donati (2015), where the Wh morphology on the determiner is unrelated to any Wh movement.

A second option is that D^0 Agrees with the Wh expression, and Wh-ever morphology is the exponent of this Agreement. With this agreement, the Wh-ever expression might or might not move, and thus could occur in either the specifier of DP or of CP. The choice between these two is very difficult, as there is no plausible material that could occur between either between Wh-ever and its clause or before it (*long whatever book Sebastian is reading* or *Whatever book long Sebastian is reading*).

(216) a. *Whatever* moves to [Spec,DP]



b. *Whatever* stays in [Spec,CP]



If we assume the syntactico-semantic argument analysis here that derives them semantically from Wh questions, the latter syntax would have to involve a derivation from questions with -ever morphology. Recall that, at least for some speakers, these are available (217). In fact, they seem to have semantic properties which seem related to the in-

terpretation of Wh-ever FRs: they seem to have a modal aspect, or at least prefer modal environments, as shown by the apparent degradedness of (217b) relative to (a); and, they resist being D-linked, shown by the fact that reference to a more specific set in the context makes the Wh-ever question quite bad (217c).

(217) a. Sebastian wonders whatever Charles could be reading

b. ?Sebastian wonders wherever Charles is

c. *Sebastian wonders whatever book by Tolstoy Charles could be reading

Bresnan and Grimshaw (1978) give several pieces of evidence for a reason why the Wh-ever in questions of the type seen in (217) is not the same as the Wh-ever in FRs. We can add to their evidence the fact that Wh-ever questions do not allow the full range of Wh expressions that Wh-ever FRs do,[6] and the fact that, the latter *do* allow D-linking.

(218) Limits on Wh expressions in -Ever Questions

a. Sebastian wants to visit however many stars there are in the sky

b. *Sebastian wonders however many stars there could be in the sky

(219) *D-linking in -Ever Questions

a. Sebastian gave Charles whichever book he is reading

b. *Sebastian wonders whichever book Charles could be reading

Given this then, the Lowering analysis might be best. However, nothing crucial seems to hinge on this choice; all of these analyses share the property of being FRs where the Wh expression does not adjoin to D^0.[7]

---

[6]In fact, if Wh-ever FRs were derived from Wh-ever questions, this would be a counterexample to Caponigro's Generalization that Wh expressions allowed in FRs are a subset of those allowed in questions.

[7]Evidence complicating questions of the relationship between FR Wh-ever and question Wh-ever comes from the distinctions in (i-iii): in the placement of the genitive morpheme, the whose-ever order seems better than the whoever's order, and seems available for FRs as well.

For certain cases, especially Bare Whatever, extraposition, and pied-piping, the grammar makes available the Pseudo FR analysis: Wh-ever is generated in D^0 and may or may not have NP and CP restrictors. Where these cases are not present, the grammar assumes the FR analysis is best, and thus complementizers and relative pronouns are blocked. However, given the fact that some speakers accept complementizers without extraposition slightly more than others, it seems that the **strength** of this preference varies.

Only FC Wh-ever is allowed in the Pseudo-FR analysis; we could thus say that this is a place within the grammar of English which looks like Italian: in just these cases, English Wh-ever has properties which led to the Pseudo-FR analysis for QQC, while in other cases, that analysis is insufficient. These include cases of Ignorant Wh-ever, as well as in cases of conjoined sentences including Wh-ever.

4.3.5 Some other issues

One thing that is not immediately explained by this syntax is the exceptionality of (ignorant) Wh-ever among all other relatives in allowing extraction out of a cleft (220). This includes both simple FRs and headed FRs, including those with free choice *any* and, as stated above, Wh-ever FRs with a free choice meaning, as observed previously in Grosu (1996). This contrasts with questions, both embedded and matrix (221).

(220) Clefts in relatives

a. Sebastian is eating whatever it is that Charles brought

b. *Sebastian is eating what it is that Charles brought

(i) ?I wonder whoever's book this could be

(ii) ??I wonder whose-ever book this could be

(iii) ?I bought whosever book I most admired

However, conclusions based on judgments as marginal as this are not sufficient to choose the best analysis in this case.

c. *Sebastian is eating the food it is that Charles brought

d. *Sebastian will eat any food it is that Charles brings

(221) Clefts in questions

a. What is it that Charles brought?

b. Charles asked what it was that Charles brought

In fact, this generalization is fairly cross-linguistically robust; Hindi-Urdu relatives, whose surface properties are quite different from those of English, also shows this distinction between clefts in relatives and questions. This is shown in (222), where the cleft is allowed in the complement of the verb *pūchhnā* 'ask' in (b), but not within the relative clause in (b).

(222) Hindi-Urdu embedded clefts

a. *kisān-ne puchhā [voh kaun hai jis-ne ghās khayī]*
farmer-ERG asked it who is REL-ERG grass ate
'The farmer asked who it is that ate the grass'

b. **kisān-ne voh gāy pālī [jo voh hai jis-ne ghās khāyı]*
farmer-ERG that cow raised REL that is REL-ne grass ate
Intended: The farmer raised the cow that (it is) that ate the grass

This fact has been alluded to in work on the semantics of certain Degree relatives in Carlson (1977) and Grosu and Landman (1998), that some relatives (such as Wh-ever FRs) allow *there* clefts while others don't, as in (223).

(223) a. #The only sailor that there was on the island drowned. (Carlson 1977)

b. I took away whatever books there were on the table. (Grosu and Landman 1998)

So what could be special about Wh-ever FRs in this regard? On the one hand, this could be explained by the fact that (according to the syntax assume above) Wh-ever FRs, like other FRs, are derived directly out of Wh questions rather than being a type of headed RC, and generally questions but not RCs allow embedded clefts. However, this does not explain the unavailability of clefts within simple FRs; nor does it explain the semantic restriction against Free Choice Wh-ever with a cleft, a fact not discussed in these previous discussions of Free Choice Wh-ever and Amount or Degree Relatives (but alluded to in von Fintel 2000).[8]

Recall as well that clefts are restricted to Ignorance Wh-ever. If we continue to assume, as I have argued here, that Ignorance Wh-ever and Free Choice Wh-ever FRs typically have the same syntax, then this distinction in the availability of clefts is necessarily semantic. That said, it is outside the scope of this work to establish a semantics for clefts that rules out all forms of relativization except for Wh-ever FRs with an ignorance reading.

An explanation for this may be something about the semantic and pragmatic facts about Clefts and Wh-ever. The simpler part of this may be the distinction between the two readings: A clefted XP must be focal, and an answer to a Question Under Discussion; Free Choice semantics necessarily involve specific referents being not-at-issue. Thus, the only reading available is Ignorant. The second issue is more complicated. It may be the case that referents of RCs, or at least heads of RCs, be necessarily old-information, but a matrix question should be able to ameliorate this. But we find that matrix questions do not make clefts within relatives any better, as the lack of contrast in (224) shows.

---

[8] Grosu and Landman (1998:161) propose that the explanation for Clefts being available just in the case of certain Free Relatives is due to the Wh expressions being “CP internal" in Wh-ever FRs, much like the syntax proposed in (216b) above; we could treat this as evidence in support of that analysis as superior to the alternative where Wh-ever moves above D^0; but again, this does not say anything about the distinction between Ignorant and Free Choice Wh-ever, so it does not necessarily make the choice for us.

(224) a. *That man is the professor it is that Sebastian hates.

b. *Who is the professor it is that Sebastian hates?

So it might have to be the case that semantic requirements on clefts must be met within a spellout domain (such as a CP, or perhaps within the DP in the case of RCs).

Another issue which I have not dealt with here is a special case of Bare Whatever. This case comes from the original example given in Bresnan and Grimshaw (1978:339), shown in (225). This variety of Bare Whatever is notable because the Wh-ever word is followed by *possible*, something which is not available to other simple DPs.

(225) a. She writes whenever possible

b. *She writes any time possible

Thus, while this is an example of Wh-ever without an apparent clausal element, it seems to be of a different sort from the Bare Whatever discussed above. Other evidence that this is the case is that it seems to be limited to adjunct Wh-ever, as shown by the contrast between (225) and (226). We also see in (227) that this seems to be restricted to certain adjectives

(226) a. *She reads whatever books possible

b. *She talks to whoever possible

c. She writes wherever possible

(227) a. She writes whenever convenient

b. ?She writes whenever relaxed

c. *She writes whenever witty

Bresnan and Grimshaw analyze this as involving deletion of *it is* (228). While this does not necessarily derive the restriction to certain Wh words, this can say something

about the restriction to certain adjectives. Since *whenever it is possible* already involves more deleted material (a complement clause of *possible*), the deletion would require reconstruction of an illicit sentence. This analysis still does not derive the facts in (226).

(228) Deletion and Bare Whatever

a. She writes whenever ~~it is~~ possible ~~to write~~

b. *She writes whenever ~~it is~~ witty ~~to write~~

c. *She writes whatever ~~it is~~ possible ~~to write~~

All of these examples are reasons for further examinations of properties particular to Wh-ever FRs; however, since none of these are properties shared with simple DPs (such as definite descriptions with *the*), they all fit into the FR analysis better than with the Pseudo-FR analysis.

4.4 A special case: Limited Number FRs

A type of FR-like pattern similar in some respects to Wh-ever, specifically in that Wh-NP is allowed is that in (229); these involve an overt NP that is necessarily plural and is interpreted as having a low cardinality (Grosu 1996). Note that unlike Wh-ever, however, it is restricted to only *what* and does not allow the wider range of Wh expressions associated with the former. This allows both plural count nouns and mass nouns, both having the Limited Amount meaning.

(229) “Limited Amount" FRs

a. Sebastian read what book-*(s) Charles had

b. *Sebastian read which books Charles read

c. Sebastian ate what food Charles made

Since these do seem to be a type of definite description (at least on the surface), this is a matter that complicates the present analysis of other types of FRs; here I will examine a potential way to accommodate them into the syntax of FRs I have proposed here.

4.4.1 These have special properties

Several facts about these seem to block a Pseudo-FR analysis of these right away: Extraposition is entirely impossible, an equivalent of Bare Wh-ever is not allowed, and the "head" cannot be an NP (as it would be in a headed RC) as it can contain the determiner *few*.

(230) a. *Sebastian gave what books to Charles that he had handy

b. *Sebastian gives what (few) books

c. Sebastian gave what few books he had handy to Charles

Note the similarity of (230) to RCs with *the few* (231), which, although they involve a regular definite determiner and thus seem like they should be regular headed RCs, are quite strange without the relative component. This has a corresponding form for mass nouns, which uses *little* in place of *few*.

(231) a. Sebastian gave the few books *(that he had handy) to Charles

b. Sebastian gave the little food *(that he had handy) to Charles.

A semantic property which unites these two is that, while they have general properties (and apparent form) of definite descriptions, and are not limited to Free Choice environments, there is some idea of uncertainty in them: cases where the exact number or nature of the referents is part of the common ground in a conversation make these somewhat degraded. The sentences in (232) and the discourse in (233) show this: when the referent is common knowledge, like the set of planets in the solar system, *what NPs* or *the few* constructions are infelicitous.

(232) a. #The sun exerts force on what planets there are in our solar system.

b. The sun exerts force on the few planets (there are) in our solar system.

c. The sun exerts force on the planets in our solar system.

(233) There were a couple books on the floor and on the table;

a. #Sebastian took what (few) books there were on the table.

b. Sebastian took the books (that were) on the table.

In his description of this phenomenon, Grosu (1996:261) suggests that this is due to a lexical semantic property of *what,* namely that, just in these constructions, *what* carries the requirement that it can only combine "with plural or mass nouns ... so long as no cardinality specifications that are either precise or high relative to some applicable scale... ." While Grosu grants that there is no guarantee that there isn't a more principled explanation (particularly considering *what* in other constructions carries no such semantic requirement), this is at least on the surface idiosyncratic, and cannot be said to emerge from semantic identity with definite descriptions as definite descriptions with *the* carries no such requirements.

4.4.2 Agreement

A fact that still must be dealt with in the account I propose here is the correlation between lack of overt NP restrictors and singular agreement. This holds for all FRs: definite, generic, and Wh-ever.

(234) a [What Sebastian wrote] is/*are very good

b [What books Sebastian writes] *is/are very good

c [Whatever Sebastian wrote] is/*are very good

d [Whatever book Sebastian wrote] is/*are very good

e [Whatever books Sebastian wrote] *is/are very good

These facts fall out naturally if we adopt the Pseudo-FR analysis; the NP restrictors in (234b,d-e) act as the “head" of the FR in the same way as in regular headed relatives, and any other DPs, and their agreement features are expressed on the verb. However, under the account proposed above something else must be said.

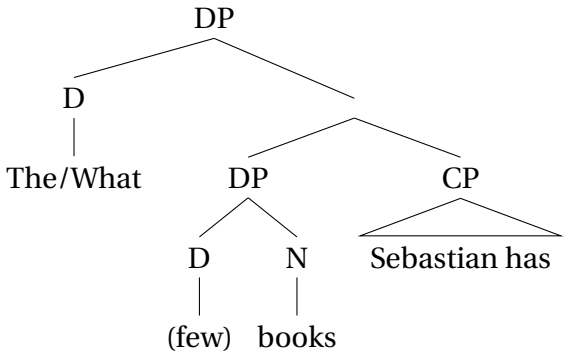
One of the remarkable things about these data is that, despite the fact that definite FRs are sometimes (and maybe always) *plural* definite descriptions, the FR in (234a) can *only* have singular agreement.

I will appeal again to the generalization made in §2.2.3 about type shifters, and their correlation with overtness in English.

4.4.3 Pseudo-Pseudo FRs

The data here seem to show that Limited-Number relatives are distinct from both standard FRs and Wh-ever FRs, particularly that it doesn't seem plausible that the Wh word moves in the way that forms other FRs, but also that they cannot be fully described in terms of typical RCs. Given this, I suggest the following syntax, which would be shared by both Limited-Number FRs and the parallel structures with *the few NPs* in (231), as in (235).

(235) Structure of Limited Number Relatives



Here, *few*-NP forms a constituent apart from the determiner *the* or the Wh word; we could call this a “Pseudo-Pseudo-Free Relative", in the sense that it takes the analysis of Battye (1989) that this construction that looks like a Wh-FR does not involve the Wh

word moving in the same way as in a simple FR, but rather resembles more closely a headed RC.

One difference between the version with *the* and the version with *what* is that the *the* version allows high cardinality, but does not allow mass nouns.

(236) a. Charles read the many books Sebastian had.

b. *Charles drank the much wine Sebastian had.

While Grosu (1996) does not mention this distinction, it seems to lend credence to his hunch that this has to do with idiosyncrasies of the construction and the lexical items involved rather than something more principled or “interesting" (Grosu 1996:262).

4.5 Wh-ever in acquisition and in adult comprehension

A common theme in the semantic literature on Wh-ever concerns the relationship between the (universal) quantificational properties vs. the definite properties of these expressions. The semantic analyses proposed in the above sections find ways to unite both of these in the denotation of the -ever morpheme. Particularly since the syntactic analysis proposed in this dissertation, and the analysis of the acquisition path for FRs, relies on the relationship between the semantic class of simple FRs - namely, definite descriptions - and the realization of definite D^0, this question is of interest.

Here I will look at three empirical domains related to Wh-ever. Observations of children’s production of Wh-ever in different semantic contexts show that children produce both Wh-ever and Free Choice *any* relatively early and in similar age ranges, but that Wh-ever seems to be used in Ignorance contexts before it is used in Free Choice contexts, and that FC *any* emerges before FC Wh-ever. Experimental data from adults show that Wh-ever of both semantic types are interpreted more like universal quantifiers than definite descriptions with *the*, but that in general there are semantic distinctions to be observed between Free Choice and Ignorance contexts. Experimental data from chil-

dren show a similar patterning of Wh-ever with universal quantifiers, but some distinctions from adults with regards to the interpretation of Free Choice contexts.

4.5.1 Wh-ever in children's production

Observation of naturalistic data shows children producing Wh-ever in a variety of syntactic and semantic contexts. Table 4.3 summarizes the production of Wh-ever, comparing production of the two semantic types of Wh-ever for 10 children in the CHILDES database (MacWhinney 2000), from corpora where a significant number of instances of any Free Choice Items could be found: Adam (Brown 1973); Ross and Mark (MacWhinney 2000); Barbara, John, and Stuart (Henry 1995); and Emily, Emma, and Matt (Weist and Zevenbergen 2008). The search included all Wh-ever words, but excluded uses of *however* in exceptive/sentence-level contexts (*however, John is tall*, etc.)

Table 4.3: Instances of Wh-ever

Child	Ignorant	FC	First instance	First type
Adam	1	2	4;0	FC
Ross	9	9	4;1	Ignorant
Mark	15	6	3;1	Ignorant
Barbara	1	0	3;11	Ignorant
John	1	1	3;6	Ignorant
Stuart	2	0	3;10	Ignorant
Emily	1	3	2;9	FC
Emma	5	0	3;8	Ignorant
Matt	0	3	4;2	FC

Across all the data, while both types are observed, there are slightly more instances of Ignorant Wh-ever; 35 total instances of Ignorant uses and 24 of FC across the 10 corpora.

There are three children who show only Ignorant uses, and none who show only FC uses. Instances of each semantic type are shown in (237).

(237) a. Ross, 4;1

Whatever number you get you have to circle it

b. Matt, 4;2

This one goes here and wherever he wants

Comparing Free Choice Wh-ever with Free Choice *any*, the latter is more productive in general; among the five children who produce both, four produce *any* first.[9] The exception, Emily, begins to produce them both at the same age.

Table 4.4: Production of FC Wh-ever and *any*

	Wh-ever		*Any*	
Child	N of instances	Earliest	N of instances	Earliest
Adam	2	4;0	4	3;5
Ross	9	4;4	12	4;3
Mark	6	4;2	14	3;9
Barbara	0	–	1	3;6
John	1	3;9	0	–
Stuart	0	–	0	–
Emily	3	2;9	3	2;9
Emma	0	–	4	4;4
Matt	3	4;2	7	3;10

[9]Barbara's only use of Wh-ever is actually ambiguous between FC and Ignorant: Among these five children, none produce Wh-ever more frequently than *any*.

(i) It's when you can put that whatever way you want

The environment (under *can*) allows for FCIs generally, but the context also suggests a specific way is being talked about. I will assume that the intent here is an Ignorant reading.

An inference we could make based on comparing Tables 4.3 and 4.4 is that the apparent imbalance of Ignorant and FC Wh-ever in early production is not due to a delay in knowledge of the semantics of Free Choice, as free choice *any* is produced quite early in general.

4.5.2 Experiment 3: The quantificational force of Wh-ever

Previous work on the acquisition of different quantifiers has shown that children show mixed behavior with regard to interpreting expressions as exhaustive. For these purposes, I'll define exhaustivity as in (238).

(238) An expression Δ is **Exhaustive** iff for any predicate P and entity x, P(Δx) = T only if there is no subset s of x for which P(s) = F.

That is, a determiner is exhaustive if it must refer to a maximal set. Exhaustive determiners include universal quantifiers, Wh expressions, and definite determiners. Children have generally been shown to interpret universal quantifiers as exhaustive from an early age (Philip 1995 inter alia). With Wh questions, children show a delay in requiring exhaustive interpretations, especially exhaustive pairing with multiple Wh questions (de Villiers and Roeper 2011), as well as for definite descriptions including prototypical definite determiners (Karmiloff-Smith 1979, Maratsos 1974, Wexler 2005) as well as simple FRs (Modyanova and Wexler 2007, Caponigro et al 2012). The delays also differ qualitatively; see the discussion in Chapter 3.

However, in this regard, Wh-Ever expressions of either semantic type have not been investigated. In fact, no study I know of has investigated whether children interpret FCIs as exhaustive or not. Studies of Free Choice interpretations have examined Free Choice inferences for disjunction (Crain 2012), but there is no literature on the quantificational force children assign to FCIs like *any* or Wh-ever.

Since FCI semantics seems to include both a universal quantificational element (ie., quantification over items in possible worlds) and a definite element (the maximal set in

each possible world), this becomes an interesting question. Further, since the argument I have presented here explains certain syntactic properties of Wh-ever vs. simple FRs in terms of properties of the syntactic/semantic class of definite determiners, examining whether children interpret FCIs as a type of definite description or a type of universal quantifier is a potential proof-of-concept for this syntactic theory.

4.5.2.1 Design

The experiment is based on the study of the exhaustiveness of simple FRs in Caponigro et al (2012). Their study involved two experiments: a truth value judgment task[10] and an act out task. The TVJT involved presentation of pictures of foods on a plate, some with only cookies, some with only onions, and some with both. For each picture, the child was asked a question like *Does Cookie Monster like what's on the plate?* with different determiners. In the act-out task, children were presented with food on plates and in buckets; for each trial, the child was asked to take out *what's in the bucket*, etc.

The distinction between the ignorant and FC uses of Wh-ever had to be taken into account in designing the experiment. To this end, there were two parts of the TVJT: a Free Choice part and an ignorant part. These involved both two different semantic frames and two different (but similar) tasks.

The Free Choice portion involved a simple TVJT: the participant was introduced to a character, Elmo, who likes to eat only some things. Specifically, Elmo will eat cakes but not soups. The participant is then asked to help a puppet, Minnie, to find food for Elmo. Minnie is shown a series of pictures with restaurants and different foods made there. For each picture, Minnie says something like *Elmo will eat whatever they make there.* The child is asked to judge this as right or wrong. This particular frame was necessary to

[10]Since the prompts were in the form of polar questions, rather than evaluations of a sentence as right or wrong, this was really a "pseudo-TVJT."

support an FCI reading: the restaurants (typically) make different things, and Elmo *will* eat these things (on a *willing to* reading of *will*).

The Ignorant portion involved a guessing game. Another character, Abby, can make some but not all kinds of food; she makes salads, but not cookies. The participant was told that they would be looking at pictures of food, and Minnie would look away and guess if Abby made the things in the picture (*Abby made whatever you have there*); the participant was instructed to tell Minnie if she guessed right or wrong. This frame forced an ignorant reading by being episodic, and the context allowed for ignorance because the guesser does not see what is on the screen.

For both parts, each picture showed an array of three foods, with four visual conditions: 0, 1, 2, or 3 of the foods are things Elmo will eat, or things that Abby made. There were four syntactic conditions, with four different determiners: *the things*, *some of the things*, *everything*, or *whatever*. Each portion of the task had eight images. The order of the images was constant for all instances, and the order of prompts varied across four pseudo-randomized lists. So, each participant hears each frame with each visual condition once, half of each being Free Choice and half being Ignorant.

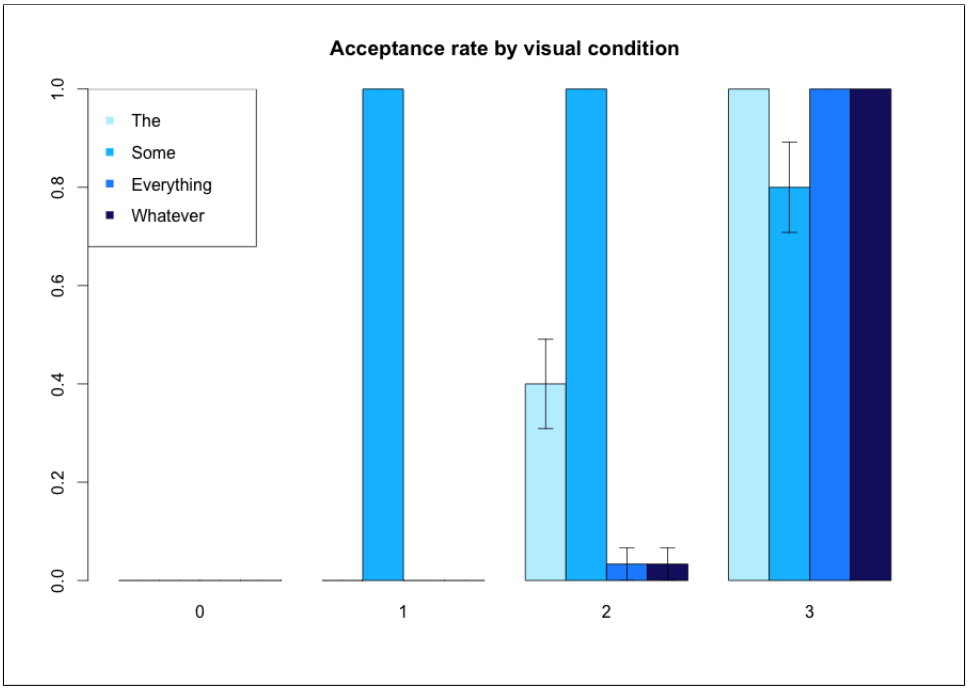
The participants of the experiment were 20 adult native speakers of English, all undergraduate students at the University of Massachusetts, and 12 children at schools in Western Massachusetts, a mean age of of 6;2 (ranging from 5;8 to 6;6). Participants were excluded from analysis if they gave affirmative responses to any item where the visual condition was 0, indicating they were either not comprehending the task or not comprehending the criteria for Elmo eating something/Abby having made something.

4.5.2.2 Adult results and discussion

For the adults, a 3-way ANOVA showed no main effect of semantic condition (F_1 = 2.2, p = .18), significant main effects of visual condition (F_3 = 335.1, p < .001) and determiner (F_3 = 126.8, p < .001), a significant interaction between visual condition and

determiner ($F_9 = 52.7$, $p < .001$) and a significant three-way interaction between semantic condition, visual condition, and determiner ($F_9 = 2.7$, $p < .01$). Figure 4.1 shows the results by determiner and visual condition.

Figure 4.1: Adult responses by visual condition, with Standard Errors



The figure shows results which fit predictions: *everything* and *whatever* are only accepted with a maximal visual condition; *some* is accepted in all non-zero conditions, with some rejections in the maximal condition; and *the* is accepted in the non-maximal condition some of the time.

The visual-semantic-determiner interaction effect comes from different rates of acceptance of *the* in the 2/3 visual condition for Free Choice versus Ignorance contexts. In the FC context, plural non-maximal *the* was accepted 67.7% of the time, vs. 13.3% in the Ignorance context. This difference is significant ($t = 4$, $p < .01$). Acceptance of non-maximal *the* seems to be restricted to **plural** non-maximal cases, never singular non-maximal. Since *some* is freely accepted in the 1/3 condition, the result is clearly not caused by *the* being interpreted as *some of the*.

The crucial result here is that, with regards to requirements of maximality, *whatever* and *everything* pattern together, and neither patterns with *the*, in both semantic contexts.

These results show that adults distinguish the quantificational force of *wh-ever* from that of definite descriptions, as evidenced by the fact that only the latter gives rise to acceptance of non-maximal conditions. This suggests that quantifier-like denotations for *wh-ever* (and perhaps all FCIs) like that of Dayal (1997, 2012) are preferable to definite denotations like those of Giannikidou and Cheng (2006).

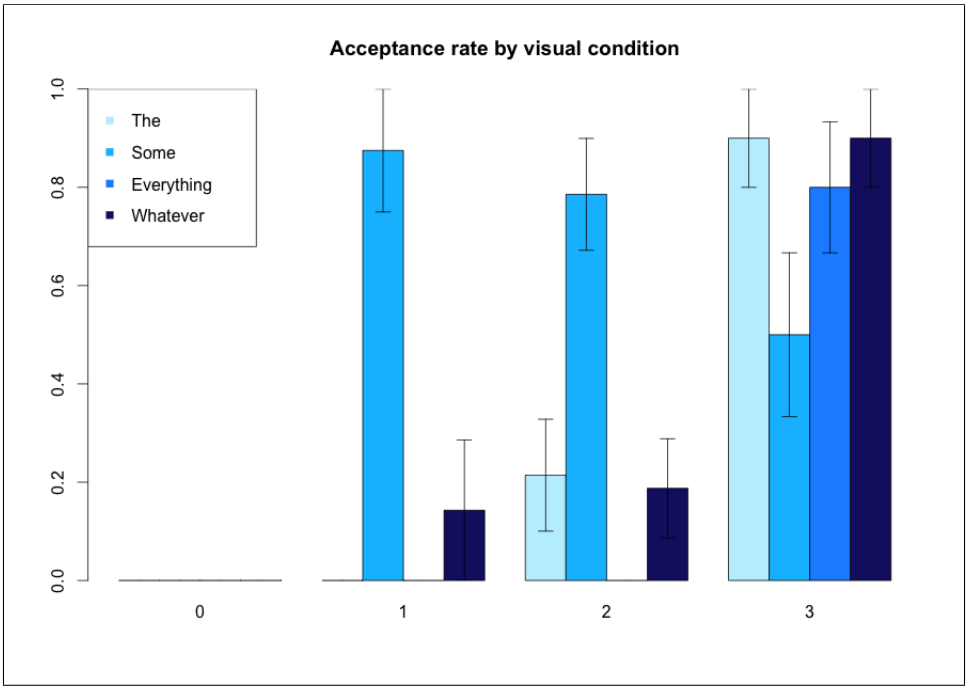
This, in turn, provides a type of evidence for the syntactic theory of FRs and Wh-ever I have proposed here: the *Wh-NP effect in simple FRs is derived from the generalization that all definite descriptions in English must include an overt D^0, that movement of the Wh word in a simple FR satisfies this, and that children's acquisition of *Wh-NP comes along with a full generalization of the overt D^0 requirement. The exceptionality of Wh-ever in allowing Wh-NP and other complex Wh expressions arises because simple FRs and typical definite descriptions are in the same syntactico-semantic class, to the exclusion of Wh-ever. These experimental data show that there is a semantic reflex of this distinction for adults.

These results cannot on their own rule out the Pseudo-FR hypothesis for Wh-ever. However, taken in the context of the above syntactic evidence distinguishing Wh-ever from Headed RCs, it is a necessary result for the hypothesis defended herein to be viable.

4.5.2.3 Child results and discussion

Two children were excluded from the analysis for giving affirmative responses where the visual condition was 0. The results by determiner and visual condition are shown in Figure 4.2.

Figure 4.2: Child responses by visual condition, with Standard Errors



The child results showed significant main effects of visual condition (F_3 = 58.1, p < .001) and determiner (F_3 = 6.7, p < .001), with no main effects of semantic condition or age and no interactions.

The adult and child patterns were largely similar, except in that the children were not sensitive to the semantic distinction in terms of their interpretation of *the*. The children were slightly less likely to accept non-maximal *the* than adults, though this difference was not significant (t = -2.3, p = .21).

So, while differences exist between children's and adults' performance on these tasks, there is no sign that these are differences in how children vs. adults interpret Wh-ever, but rather in the nature of non-maximal interpretations of definite descriptions. A theory of why adults are willing to accept non-maximal *the* in general will need to account in some way for why the FCI condition here, but not the Ignorant condition, and a the-

ory of children's non-maximal interpretation of *the* will have to account for there **not** being a difference in behavior by semantic condition.

However, the requirements on maximality of Wh-ever are not subject to this same sort of variation for children or for adults. For both groups, Wh-ever behaves more like a universal quantifier than a definite description in this task.

Recall from the discussion of naturalistic data above that while children produce both Wh-ever and Free Choice *any* quite early, Free Choice Wh-ever is slightly delayed relative to both. This experiment provides evidence that this is not necessarily because there is any problem with assigning the correct semantic representations to Wh-ever. It may rather be due to the fact that Free Choice Wh-ever has two available syntaxes, as proposed above.

This would be similar to the phenomenon discussed in Chapter 3 where headed RCs are produced later than simple FRs, possibly for the reason that there are two possible representations for them available in the target grammar; in this case, the child does not have a clear choice between the FR and Pseudo-FR analysis of Wh-ever, but only for Free Choice Wh-ever. This leads Free Choice Wh-ever in particular being delayed in production relative to Ignorant FR, which has only one target representation.

4.5.3 Wh-ever and and nominal parameter setting

Here I have provided experimental evidence that, while Wh-ever shares some semantic properties with Free Choice *any* and with definite descriptions, it does not fit into a syntactic class with *any* or *the* nor a semantic class with simple FRs. A consequence of this is that, for the learner, Wh-ever must be learned as a special kind of FR which does not share the properties of simple FRs which give rise to *Wh-NP.

In the context of the theory of acquisition adopted in the previous chapters, and the evidence that children show a delay in acquiring *Wh-NP for simple FRs, this is understood as a case where the "rule" arises in development as children make proper general-

izations about the distribution of the formal features responsible for it. That is, in terms of the theory of parameter setting discussed in Chapter 3 as language-specific generalizations about the co-occurrence of different features, it emerges as children make the proper generalization about where the feature [+Overt] occurs, which requires certain morphemes to be pronounced.

As the delay of *Wh-NP seems to extend beyond the point where English-speaking children start requiring *the* to be overtly pronounced, we need an articulated description of the formal features associated with the different determiners to describe this. The primary difference between the definite D^0 in an FR and *the* is essentially selectional: the latter selects an NP, the former a CP (to be denoted as uninterpretable selectional features *u*N and *u*C). Aside from this, they are both D heads and both definite. The D head associated with Wh-ever is not definite (but still has a *u*C feature). These Feature sets are shown in (239).

(239) *the* = [+D, +def, *u*N, ...]

D_{FR} = [+D, +def, *u*C, ...]

D_{ever} = [+D, -def, *u*C, ...]

To move through the English acquisition path, a child must first learn that *the* has a [+Overt] feature, then make a generalization that any functional head which is both +D and definite must also have [+Overt] (240-241). This involves going from a narrower Parameter setting (a nano-parameter in the hierarchy discussed in Chapter 3, after Biberauer and Roberts 2012) to a broader one (a micro-parameter).

(240) Overtness Parameter in Child English: *the* = [+Overt, ...]

(241) Overtness Parameter in Adult English: If [+D, +def] then [+Overt]

A conception of the acquisition path in these terms is a move towards the goals discussed in Chapter 1 of a Minimalist theory of acquisition and Parameter Setting: rather

than a child needing to make broad Parametric choices or appeal to articulated transformational rules or specific representational Principles, associations between formal features of natively-available functional heads are learned, and these interact with simple and general properties of labeling and head movement to create surface phenomena like *Wh-NP. Further studies on this and other topics could make use of this framework in attempts to model the acquisition of other language- and construction-specific phenomena as interactions between specific Parameter Settings of this type and general properties of the Minimalist derivational system.

4.6 General conclusions

In this chapter I have argued for ways to unite the syntax of Wh-ever FRs with that of simple FRs in English. This has relied on the same general principles that have been discussed throughout this dissertation: that in languages which derive FRs from Wh constructions, the emergent properties of FRs originate from the interaction of Wh movement with the general syntax of the nominal domain.

The resultant representations for FRs illustrate the necessary path of acquisition as involving learning the syntax of FRs by applying knowledge of the properties of different types of nominals to Wh strings to make them “fit" into specific positions for them to be licensed syntactically and for them to compose semantically. That is, specifically, a string corresponding to an FR is given Wh-clause syntax and type-shifted and relabeled as a nominal. In Chapter 1 I proposed a possible - and perhaps necessary - mechanism for positing this sort of structure whereby a constituent is labeled based on the environment in which it appears, and requires the existence of a head which can act as the label of that constituent. I formalize this as in (242).

(242) **Head Positing Formula**

If [α [β ...] ...] whereα selects XP, then [α [$_{XP}$ ...] ...]

If [α [$_{XP}$ ...] ...], then [α [$_{XP}$ X^0 ...] ...]

The learner's goal after positing this head is to determine whether the head has a phonological exponent, and what that exponent is. For the representation of FRs proposed in this dissertation, the specific goal is to determine that the Wh word in a simple FR is the labeling head of the DP, and that it has moved to this position. This process can be modeled as in (243): a Wh string is given typical Wh-CP syntax by default (i); when appearing in a position which is necessarily nominal (such as the complement of a DP-selecting verb), it is given a type-shifting D head so the selectional requirements can be met, resulting in a CP embedded within a DP (ii); finally, since this D is definite, it carries a feature which requires it have a phonological exponent, and so the Wh word adjoins to it, moving out of the CP and effectively blocking certain strings (like Wh-NP) from being in FRs (iii).

(243) Representing an FR:

(i) *what Sebastian read* → [$_{CP}$ what Sebastian read]

(ii) *write what Sebastian read* → [$_{VP}$ write$_D$ [$_{DP}$ D [$_{CP}$ what Sebastian read]]]

(iii) [$_{DP}$ D [$_{CP}$ what Sebastian read]] → [$_{DP}$ [$_D$ what] [$_{CP}$ t_D Sebastian read]]

In Chapter 2 I gave arguments for why the output in (iii) best represents simple FRs in English by generating only the FRs we observe in English. Further, it does so in a way corresponding to the intuition behind (243), that the form of FRs must be learned from combining aspects of Wh movement (what can move, and to where) and the nominal system (definite D^0 must be overtly realized). This makes use of the tools available to a Minimalist learner as described in Chapter 1 (such as the Head Positing Formula), and

delivers a representation of FRs which produces effects like *Wh-NP, and the other phenomena discussed in Chapter 2. Further, it does so specifically by making generalizations about the distribution of formal features and the relationship between observable forms and functional structure.

In Chapter 3 I gave evidence that the steps (i-ii) in (243) are acquired early and (iii) is acquired late, and this corresponds with the early development of both the syntactic properties of Wh movement and the categorial distinction between FRs and questions, and the later development of the requirements on overt D^0. That is, there is a stage in acquisition where (ii) is the representation of simple FRs, and (iii) does not apply. Aside from providing validation of this theory of FRs, these results also contribute to a general picture of the developmental path of representations in the nominal domain, specifically what sort of Parameters exist which are relevant to the determiner system in different languages, how these Parameters are set, and how these settings interact to produce different developmental phenomena

All of these facts fit within the ideas of an acquisition path following Minimalist guidelines sketched in Chapter 1: surface-visible properties, such as the presence of movement and the correlates of selection, are part of earlier grammars, but properties which are opaque on the surface (such as the destination of movement, and the set of functional heads which are or aren't pronounced in a given sentence type) are acquired later; and, when changes to the grammar are made during acquisition, they will always satisfy Minimalist criteria. So for example, in the acquisition of simple FRs, children will not solve the requirement of overt definiteness head by moving a phrasal Wh expression to adjoin to D^0.

It also fits into the idea discussed throughout this dissertation, originating in Borer (1984) and Chomsky (1995), that the locus of language specific "rules" of syntax is variation in formal properties of particular lexical items, here in particular specific D^0 heads; the rules which distinguish English FRs from any other construction come from formal

features such as [+Def] and [+Overt] being associated with each other in the lexicon; the set of operations allowed in the narrow syntax are constant across grammars, and only the input to the syntax results in visibly different patterns.

The results presented here largely validate the intuition followed in Construction-based accounts of syntax and acquisition that the grammar must be structured in a way that exceptions are not treated as anomalous, but rather that they are built into the system. However, formal properties of derivation and representation, which are derived from the present definitions of Merge and the lexicon, provide a guideline for both canonical and “exceptional" syntactic patterns. This is thus more compatible with this Minimalist-Nativist model of language acquisition rather than an account more closely derived from Usage- or Input-based accounts such as those of Tomasello (2000) or Lieven (2010); an acquisition mechanism which is entirely functional, or is strongly Input-driven, is naive to formal structures and the sorts of abstract representations discussed here. A learner of such a system would not be able to propose a connection between an abstract D head and Wh movement which produces adultlike English FRs, nor would the sorts of parallelisms between the development of FRs and other definite descriptions be observable.

In addition to broadly providing support to a formal, Nativist model of language acquisition, the results discussed here provide a framework in which future acquisition studies can be based. For example, the developmental paths of FRs in languages with different functional lexicons than English, or of a wider array of nominals and English, could take the results here as a starting point. The role of other properties necessary for a Minimalist system in acquisition can also be examined: where are places where the strict binarity of Merge guides children through acquisition problems, for example? In general, the result of this work is to provide hope that Minimalism, despite its bare-bones notion of the Language Acquisition Device, can still guide children learning lan-

guages through grammatical principles, rather than having to rely solely on properties of the input or Third Factor constraints.

APPENDIX

PIED-PIPING AND HEADED RELATIVES

Above I showed how particular restrictions on complex Wh expressions which differ between FRs and Wh questions. Similar sorts of data show a different, but related, set of patterns for Full, or (Headed) Relative Clauses (here, RCs). Here I will discuss the differences between Wh expressions which occur in RCs, and what sort of differences in their derivations can account for this.

A.1 Headed relatives

Headed RCs are syntactically similar to FRs in that they have nominal distribution, embed a gapped clause derived by $\bar{A}$ movement, and denote an entity or set of entities. They are less restricted than FRs semantically in that they can combine with any sort of determiner or quantifier which can be associated with a simple DP, and perhaps most obviously, contain an overt NP which acts as the notional “head" (but on the assumption that nominal expressions are in general DPs and not NPs, not the categorial head) of the expression - a fact which immediately seems to contrast with the *Wh-NP effect seen in FRs.

(244) Headed RCs

a. Sebastian {borrowed, *wondered} the book that Charles was reading.

b. Sebastian borrowed the/all (the)/some/two/no the books Charles was reading.

c. *Sebastian borrowed the/all (the)/some/two/no what Charles was reading.

In terms of the constraints on FRs seen above, we can see properties of Headed RCs which are and aren't shared by FRs. Specifically, certain types of complex Wh expressions are allowed in headed RCs - they allow pied piping of prepositions and possessed NPs, and a broader range of Wh words, including *which* and *who* (in this case, unanimously for all speakers of English). They do not, however, uniformly allow complex Wh, as Wh-NP sequences which don't involve a pied-piped preposition or possessum are wholly barred in full RCs (246).

(245) Complex Wh in RCs

a. Charles remembered the party **to which** Sebastian had brought the wine.[1]

b. Charles met the boy **whose wine** Sebastian brought.

c. Charles drank the wine **which** Sebastian brought.

d. Charles met the boy **who** Sebastian liked.

(246) *Wh-NP in RCs

a. *Charles attended the party which disaster Sebastian hosted.[2]

b. *Sebastian drank the bottles how much wine made him sick.

While (246) has a potentially recoverable meaning - there were a number of parties, Charles attended the one which was a disaster and which Sebastian hosted - the sentence is an impossible way to express this.

---

[1]It is worth noting, however, that this sort of RC depends on a speaker accepting PP pied piping in general. Many speakers of Modern American English, especially those who do not rigidly follow prescriptive rules, rarely-if-ever produce this sort of pattern in any sort of Wh movement construction, including RCs and main clause and embedded Wh questions. That said, for those who do, the contrast between full relatives and embedded Wh on the one hand and FRs on the other is quite strong.

[2]There is a variety of this seen in legal documents and some literary registers, as in the following passage:

(i) Rosa Fast issued a brief statement inviting Flora Hill to "come on down and meet this little sweetheart," **which invitation** Hill did not take up.
Salman Rushdie, *Two Years, Eight Months, and Twenty-Eight Nights* p81

This is, however, not typically produced or accepted in natural spoken English.

Headed RCs also allow overt complementizers, which in Modern English vary freely with Wh words (or neither) in relatives, which are barred in FRs entirely (247b-d) and are strongly dispreferred in embedded Wh questions (247). Both of these are instances of what is generally called the "Doubly Filled COMP" filter (Chomsky and Lasnik 1977).

(247) The distribution of *that*

a. Sebastian brought [the wine that Charles served at the party].

b. Sebastian brought [what (*that) Charles served at the party].

c. *Sebastian baked [that Charles served at the party].

d. *Sebastian baked [Ben served at the party].

e. ?Sebastian wondered [what that Charles served at the party].

I will not endeavor to explain all the peculiar facts about the distribution of *that* in different RCs (as in 247), but they do serve to illustrate another distinction between the surface syntax of RCs with and without NP heads.

A.2 To pied-pipe or not

Noting some of the facts in (245), Donati (2006) suggests that they involve is a constraint on movement, which she formulates as the generalization in (248). I suggest a revision of this generalization (249), based on Clauss (2013). There were originally several reasons for this revision, but an important one for this discussion was that it distinguished pied piping of prepositions and possessed NPs (245a-b) from Wh-NP per se, as we have already seen that these pattern differently in that only the latter is barred in standard RCs. Both of these generalizations capture the fact that a variety of complex Wh expressions cannot occur in standard FRs.

(248) **Constraint on Pied Piping** (Donati 2006)

A simple Wh structure excludes pied-piping exactly in those cases in which it occurs in a nominal position

(249) **Constraint on Complex Wh** (Clauss 2013)

Wh Phrases may not move to the specifier position of D

Another generalization made about complex Wh comes from (Caponigro 2003): that Wh expressions allowed in Wh-FRs in any given language are a subset of those allowed in Wh questions in that language. Adding Headed RCs to the picture, in English we still see a subset relationship between complex Wh in Questions and all Relatives, but we do not see a subset relationship between Headed and Free Relatives. Instead we see the more complicated correspondences in Table A.1.

Table A.1: Wh expressions allowed in Relatives cross-linguistically

	What	Who	How	How much	why	Wh-NP	PP	Whose-NP
Q	✓	✓	✓	✓	✓	✓	✓	✓
Headed RC	*	✓	*	*	?	*	✓	✓
English FR	✓	%	*	*	*	*	*	*

The reason this becomes an intersecting rather than subset relationship is in particular because of *what*, which cannot appear in full RCs in Mainstream American English, though there are varieties of English where it can. Still, this is enough to show that the same sort of implication between Wh words in FRs and Wh questions does not entail an implication between Wh words in FRs and Headed RCs. Instead, the implications are more complex; (250) gives a second pass at a generalization of these facts.

(250) **Entailment relations between Wh expressions in different constructions**

i Wh expressions available to FRs are a subset of those available to Wh questions

ii Wh expressions available to Headed RCs are a subset of those available to Wh questions

iii Constructions which allow embedded pied piping are a subset of those which allow Wh-NP

However, there is another type of pied piping which complicates this slightly: the availability of a more complex type of pied piping in Headed RCs but not in questions (or, unsurprisingly, FRs), where the Wh word is embedded more deeply in a DP, as in (251), “Heavy" pied piping by Safir (1986).

(251) Heavy pied piping

a. Sebastian wrote the book **a passage from which** Julia read to Charles.

b. *Sebastian wrote **a passage from what** Julia read to Charles.

c. %**A passage from which book** did Julia read to Charles?

Depending on the interpretation of “Wh expressions" in (250), this isn't necessarily an exception. If “Wh expression" is taken to mean “Wh word", then the generalization goes through as is. However, if the aim of this generalization is to capture the full distribution of simple and complex Wh phrases in different constructions (as Caponigro's Generalization aims to do), then something more must be said. A revision of (250) is needed to capture the fact that while in general types of Wh expressions in questions (specific Wh words, or embedded pied piping, Wh-NP) are always a superset of those in FRs and in Headed RCs, there is a type of embedded pied piping allowable only in Headed RCs.

(252) **Entailment relations between Wh expressions in different constructions, revised**

i Types of Wh expressions available to FRs are a subset of those available to Wh questions

ii Types of Wh expressions available to Headed RCs are a subset of those available to Wh questions

iii Constructions which allow embedded pied piping are a subset of those which allow Wh-NP

iv Constructions which allow Heavy pied piping are a superset of those that allow embedded pied piping

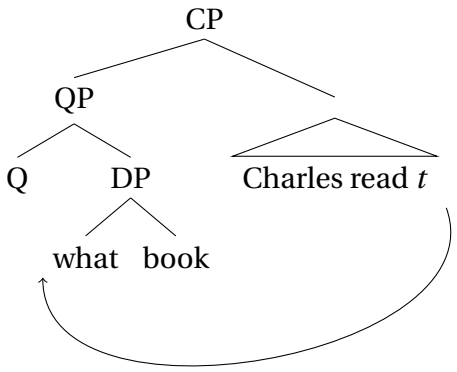
While these facts may be in some sense innate, determining the set of pied piping properties available in a target language remains an acquisition problem. Derivations for most of these generalizations are outside of the scope of the current project; rather they serve here as a background for the discussion of patterns seen in FRs in particular, and what sort theories of pied piping might best explain children's production and understanding of FRs and of *Wh-NP in particular.

A.3 How to Pied-pipe

An ideal outcome of the present discussion would be that the analysis in Chapter 2 would extend to deriving the particular set of constraints seen for RCs as well as those of RCs. Here I will lay out some possible representations of RCs and of pied piping structures which may achieve this at least somewhat.

The analysis of Wh-question formation in Cable (2010) assumes that all instances of pied-piping cross-linguistically occur because the Wh word must be dominated by a QP, a phrase containing a Q^0 Head (which only has a morphological expontent in some languages) which gives the question interpretation, and this constituent is what moves.

(253) QP movement, after Cable (2010)



A first thing to ask about this would be whether it affects the present analysis of FRs. It does not seem necessary that it should; movement of the Wh word to D would still have to take place, and movement away from an NP complement would still constitute illicit left branch movement. The movement would just be coming out of a DP within the QP.[3]

For the purposes of explaining specific pied-piping phenomena in English, Cable does not discuss specific implementations, but does suggest that the optionality of PP pied piping in English arises from the fact that P is a lexical category in English; thus QP may dominate either just the DP or the PP which contains it, and this choice determines whether or not the PP is pied-piped with the Wh expression. Possessive pied-piping would be forced since to do otherwise would still be left-branch movement.

A consequence of that is that PP pied-piping can be understood as a form of Heavy Pied-Piping. In both cases, the QP dominates a bigger constituent. We can ask then how constrained either of these is, and how we might derive that.

There are cases where a constituent smaller than the “Heavy" constituent but bigger than the Wh constituent is blocked, as in (254). This is a case where the barring of the

[3]One might worry that this would violate the Head Movement Constraint if it involves *what* moving past the Q head. However, since Q is silent in English, there is no way of knowing that such movement is occurring. I will not take a strong stance on the Head Movement Constraint here.

smaller pied-piped constituent follows from a more general island constraint, as it is moving out a Subject Island (255).

(254) a. The man [a picture of whom] is hanging on my wall.

b. *The man [of whom] a picture is hanging.

(255) a. *Who is [a picture of *t*] hanging on your wall?

b. *Of whom is [a picture *t*] hanging on your wall?

We might describe this as simply as, pied-piping cannot move a constituent which cannot otherwise be moved. Or, in terms of the QP theory, we could make a generalization like (256).

(256) **Constraint on Heavy Pied-piping**

QP may dominate either:

(i) A constituent headed by a Wh word, or

(ii) A constituent containing a Wh word which is capable of Ā movement.

Further evidence given by de Vries (2005) presents evidence from Dutch and German that ability of Ā-Movement conditions the size of pied piping, as in this contrast: preposition stranding is barred in Dutch except in the case of a specific set of locative pronouns; in colloquial Dutch these may take the place of animate Wh words and permit preposition stranding in a relative.

(257) *Preposition Stranding (de Vries 2005:41)

a. *De man over wie we spraken*

The man of whom we spoke

b. **De man wie we over spraken*

The man who we about spoke

(258) ✓Preposition Stranding (de Vries 2005:43)

a. *De man waar hij over sprak*

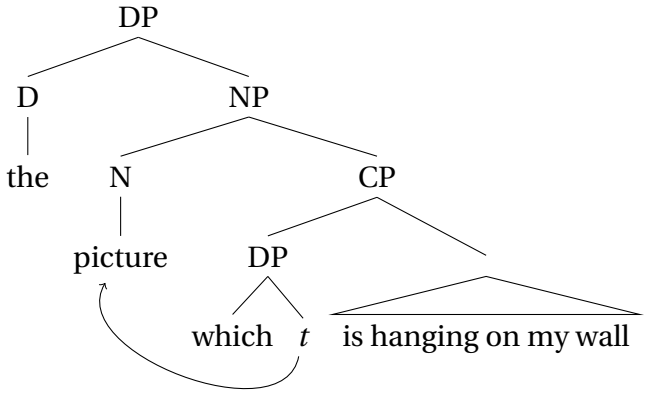
The man where he about spoke

b. *De bron waar hij uit putte*

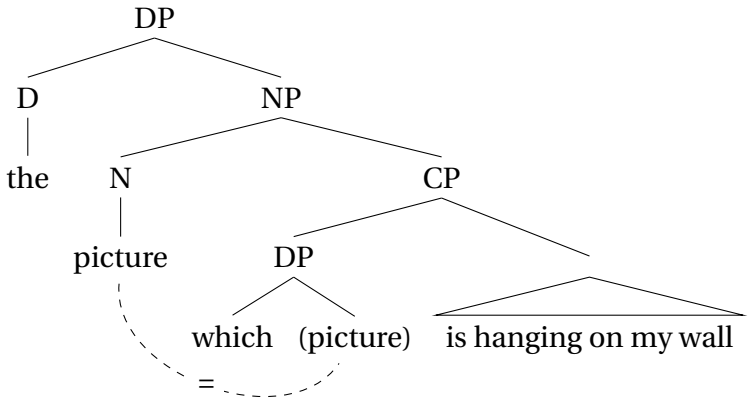
The spring where he from drew

A last question we might ask is, how is Heavy Pied-piping derived? Work by Bhatt (2002) and Hulsey and Sauerland (2006) has suggested that there must be two types of RC analyses within English: One where the head NP itself is moved out of the RC, and one in which the Head is generated external to the RC, and an identical NP is moved within the RC and goes unpronounced. Following Bhatt (2002) I will call these the Raising and Matching analyses.

(259) Raising analysis



(260) Matching analysis



Evidence presented in Bhatt (2002) and Hulsey and Sauerland (2006) suggests that the interpretation of certain potentially-ambiguous adjectival modifiers is dependent on the choice between the two syntaxes. The relevant contrast is given in (261); the two readings are called “high" and “low" based on whether the adjective *first* is interpreted relative to the saying event (the matrix predicate) or the writing event (the embedded predicate).

(261) I read the first book that John said that Tolstoy had written (Hulsey and Sauerland 2006:116, after Bhatt)

High reading: The first book about which John said that Tolstoy had written it

Low reading: The x such that John said that the first book Tolstoy had written was x

Since Heavy pied piping involves a complex nominal which would normally be an island, we might predict that, accepting the analysis that only the high reading allows matching and the low reading must involved raising, Heavy pied piping in a similar sense should only allow the high reading. We see in (262) that this seems to be the case.

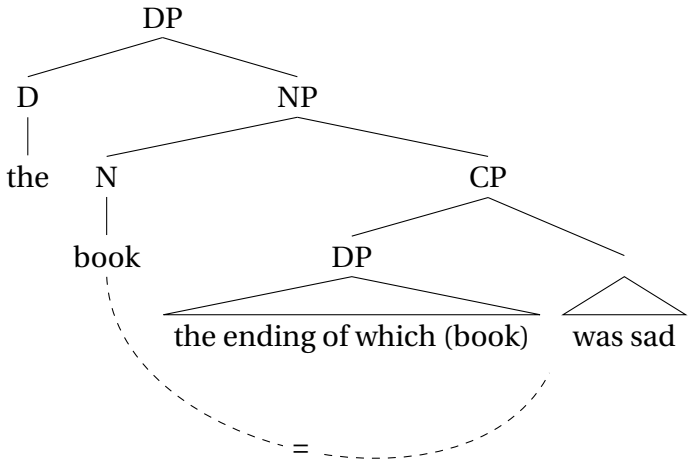
(262) I read the first Tolstoy book the ending of which John said was sad

high reading: The first Tolstoy book about which John said that the ending was sad

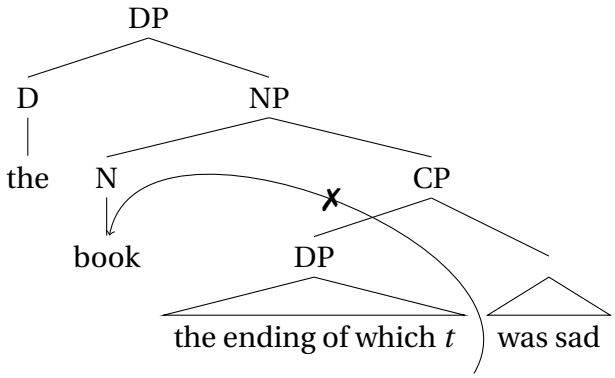
low reading: #The x such that John said that the first sad book Tolstoy wrote with a sad ending had written was x

Since it seems possible to get the high reading but not the low reading here, it follows from Hulsey and Sauerland's argument that this sentence necessarily has the matching analysis. This gives us a plausible syntax for the Heavy pied-piped Wh expression in (262): it must involve head matching, since raising would involve movement out of a complex DP, and the interpretation associated with matching is the only one available.

(263) Heavy pied-piping with Matching



(264) *Heavy pied-piping with Movement



This conforms to the generalization in (256): pied-piping is constrained by general constraints on movement, and this plays out in a number of syntactic and semantic ways. This still leaves unsolved the issue of why Heavy Pied Piping is not allowed in questions, but it still follows that it should be barred in Free Relatives just as with other Pied-Piping.

BIBLIOGRAPHY

Adani, F. (2011). Rethinking the acquisition of relative clauses in italian: towards a grammatically based account. *Journal of Child Language 38*, 141–165.

Bates, E., I. Bretherton, and L. Snyder (1988). *From First Words to Grammar*. Cambridge: Cambridge University Press.

Battye, A. (1989). Free relatives, pseudo-free relatives and the syntax of cp in italian. *Rivista di Linguistica 1*, 219–250.

Béjar, S. and M. Rezac (2009). Cyclic agree. *Linguistic Inquiry 40*, 35–73.

Belletti, A. (2009). *Structures and Strategies*. New York: Routledge.

Belletti, A., E. Bennati, , and A. Sorace (2007). Theoretical and developmental issues in the syntax of subjects: Evidence from near-native italian. *Natural Language and Linguistic Theory 25*, 657–689.

Belletti, A. and C. Hamann (2004). On the l2/bilingual acquisition of french by two young children with different source languages. In P. Pr'evost and J. Paradis (Eds.), *The Acquisition of French in Different Contexts*, pp. 147–174. Amsterdam: John Benjamins.

Belletti, A. and C. Leonini (2004). Subject inversion in l2 italian. In S. Foster-Cohen, M. Sharwood Smith, A. Sorace, and M. Ota (Eds.), *EUROSLA Yearbook*, pp. 95–118. Amsterdam: John Benjamins.

Berwick, R., P. Pietroski, B. Yankama, and N. Chomsky (2011). Poverty of the stimulus revisited. *Cognitive Science 35*, 1207–1242.

Bhatt, R. (2002). The raising analysis of relative clauses: Evidence from adjectival modification. *Natural Language Semantics 10*, 43–90.

Borer, H. (1984). *Parametric Syntax*. Dordrecht: Foris.

Borer, H. and K. Wexler (1987). The maturation of syntax. In T. Roeper and E. Williams (Eds.), *Parameter Setting*, pp. 123–172. Dordrecht: Reidel.

Bošković, Ž. (2005). Left branch extraction, structure of np, and scrambling. In M. Saito and J. Sabel (Eds.), *The Free Word Order Phenomenon*, pp. 13–73. Berlin: Mouton de Gruyter.

Bresnan, J. and J. Grimshaw (1978). The syntax of free relatives in english. *Linguistic Inquiry 9*, 331–391.

Brown, R. (1973). *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.

Cable, S. (2010). *The Grammar of Q: Q-Particles, Wh-Movement, and Pied-Piping*. New York: Oxford University Press.

Caponigro, I. (2003). *Free Not to Ask: On the Semantics of Free Relatives and Wh-Words Cross-Linguistically*. Ph. D. thesis, UCLA.

Caponigro, I., L. Pearl, N. Brooks, and D. Barner (2012). Acquiring the meaning of free relative clauses and plural definite descriptions. *Journal of Semantics 29*, 261–293.

Carlson, G. (1977). Amount relatives. *Language 53*, 520–542.

Cecchetto, C. and C. Donati (2010). On labeling: Principle c and head movement. *Syntax 13*, 241–278.

Cecchetto, C. and C. Donati (2015). *(Re)Labeling*. Cambridge, MA: MIT Press.

Cheng, L. L.-S. (2000). Moving just the feature. In U. Lutz, G. Müller, and A. van Stechow (Eds.), *Wh Scope Marking*. Philadelphia, PA: John Benjamins.

Cheng, L. L.-S., C. Heycock, and R. Zamparelli (2017). Two levels for definiteness. In *Talk given at GLOW Asia*.

Cheng, L. L.-S. and R. Sybesma (1999). Bare and not-so-bare nouns and the structure of np. *Linguistic Inquiry 30*, 509–542.

Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics 6*, 339–405.

Chierchia, G. and I. Caponigro (2013). Questions on questions and free relatives. In *Sinn und Bedeutung 18*.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton de Gruyter.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1968). *Language and Mind*. New York: Harcourt, Brace, Jovanovich.

Chomsky, N. (1977). On wh-movement. In P. Culicover, T. Wasow, and A. Akmajian (Eds.), *Formal Syntax*, pp. 71–132. New York: Academic Press.

Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale and S. Keyser (Eds.), *The View from Building 20*, pp. 1–52. Cambridge, MA: MIT Press.

Chomsky, N. (1995). Bare phrase structure. In H. Campos and P. Kempchinsky (Eds.), *Evolution and Revolution in Linguistic Theory*, pp. 51–109. Washington, DC: Georgetown University Press.

Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry 36*, 1–22.

Chomsky, N. (2013). Problems of projection. *Lingua 130*, 33–49.

Chomsky, N. and H. Lasnik (1977). Filters and control. *Linguistic Inquiry 8*, 425–504.

Cinque, G. (2013). Cognition, universal grammar, and typological generalizations. *Lingua 130*, 50–65.

Collins, C. (2002). Eliminating labels. In S. D. Epstein and T. D. Seely (Eds.), *Derivation and Explanation in the Minimalist Program*, pp. 42–64. Blackwell.

Crain, S., L. Koring, and R. Thornton (Forthcoming). Language acquisition from a bilinguistic perspective. *Neuroscience and Biobehavioral Reviews*.

Dabrowska, E., C. Rowland, and A. Theakston (2009). Medial wh relies on a sort of adjunction of two question forms. *Cognitive Linguistics 20*, 571–598.

Dayal, V. (1996). *Locality in Wh Quantification: Questions and Relative Clauses in Hindi*. Dordrecht: Kluwer.

Dayal, V. (1997). Free relatives and *ever*: Identity and free choice readings. In A. Lawson (Ed.), *SALT VII*, pp. 99–116.

Dayal, V. (2013). A viability constraint on alternatives for free choice. In A. Falaus (Ed.), *Alternatives in Semantics*, pp. 88–122. Basingstoke: Palgrave MacMillan.

de Villiers, J. (2007). The interface of language and theory of mind. *Lingua 117*, 1858–1878.

de Villiers, J., T. Roeper, E. Harrington, and E. Gadilauskas (2011). Tense and truth in children's question answering. In *Boston University Conference on Language Development 36*, Somerville, MA. Cascadilla Press.

de Villiers, J., T. Roeper, and A. Vainikka (1990). The acquisition of long distance rules. In Frazier and de Villiers (Eds.), *Langage Processing and Language Acquisition*, pp. 257–297. Boston: Kluwer.

de Vries, M. (2005). Possessive relatives and (heavy) pied-piping. *The Journal of Comparative Germanic Linguistics 9*, 1–52.

Déchaine, R. and M. Wiltschko (2002). Decomposing pronouns. *Linguistic Inquiry 33*, 409–442.

Deen, K. and N. Timyam (forthcoming). Condition c in adult and child thai. *Language*.

Donati, C. (2006). On wh head movement. In N. Corver and L. Cheng (Eds.), *Wh Movement: Moving On*, pp. 21–46. Cambridge, MA: MIT Press.

Flynn, S. and B. Lust (1980). Acquisition of relative clauses: Developmental changes in their heads. *Cornell Working Papers in Linguistics 1*, 33–45.

Friedmann, N., A. Belletti, and L. Rizzi (2009). Relativized relatives: Types of intervention in the acquisition of a-bar dependencies. *Lingua 119*, 67–88.

Gavarró, A., A. T. Pérez-Leroux, and T. Roeper (2006). Definite and bare noun contrasts in child catalan. In V. Torrens and L. Escobar (Eds.), *The Acquisition of Syntax in Romance Languages*, pp. 51–68. Amsterdam: John Benjamins.

Gavruseva, E. and R. Thornton (2001). Getting it right: Acquisition of whose-questions in child english. *Language Acquisition 9*, 229–267.

Giannikidou, A. (2001). The meaning of free choice. *Linguistics and Philosophy 24*, 659–735.

Giannikidou, A. and L. L.-S. Cheng (2006). (in)definiteness, polarity, and the role of wh-morphology. *Journal of Semantics 23*, 135–183.

Groos, A. and H. van Riemsdijk (1981). Matching effects in free relatives: A parameter of core grammar. In A. Belletti, L. Brandi, and L. Rizzi (Eds.), *Theory of Markedness in Generative Grammar: Proceedings of the 1979 GLOW Conference*, pp. 171–216. Pisa: Scuola Normale Superiore di Pisa.

Grosu, A. (1996). The proper analysis of “missing-p" free relative constructions: A replay to larson. *Linguistic Inquiry 27*, 257–293.

Grosu, A. and F. Landman (1998). Strange relatives of the third kind. *Natural Language Semantics 6*, 125–170.

Gualmini, A. and S. Crain (2005). The structure of children's linguistic knowledge. *Linguistic Inquiry 36*, 463–474.

Guasti, M. T., A. Gavarró, J. de Lange, and C. Caprin (2008). Article omission across languages. *Language Acquisition 15*, 89–119.

Guasti, M. T. and U. Shlonsky (1995). The acquisition of french relative clauses reconsidered. *Language Acquisition 4*, 257–276.

Hall, W. S., W. E. Nagy, and R. Linn (1984). *Spoken Words: Effects of Situation and Social Group on Oral Word Usage and Frequency*. Hillsdale, NJ: Erlbaum.

Hamblin, C. L. (1973). Questions in montague english. *Foundations of Language 10*, 41–53.

Han, C.-H., J. Lidz, and J. Musolino (2007). V-raising and grammar competition in korean: Evidence from negation and quantifier scope. *Linguistic Inquiry 38*, 1–47.

Harley, H. and E. Ritter (2002). Person and number in pronouns: A feature-geometric analysis. *Language 78*, 482–526.

Hauser, M., N. Chomsky, and W. T. Fitch (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science 298*, 1569–1579.

Hawkins, J. (2004). *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.

Henry, A. (1995). *Belfast English and Standard English: Dialect Variation and Parameter Setting*. New York: Oxford University Press.

Hulsey, S. and U. Sauerland (2006). Sorting out relative clauses. *Natural Language Semantics 14*, 111–137.

Izvorski, R. P. (2000). *Free Relatives and Related Matters*. Ph. D. thesis, University of Pennsylvania.

Karmiloff-Smith, A. (1979). *A Functional Approach to Child Language*. Cambridge University Press.

Karttunen, L. (1977). Syntax and semantics of questions. *Language and Philosophy 1*, 3–44.

Kidwai, A. (2000). *XP-Adjunction in Universal Grammar: Scrambling and Binding in Hindi-Urdu*. New York: Oxford University Press.

Kratzer, A. (1996). Severing the external argument from its verb. In J. Rooryck and L. Zaring (Eds.), *Phrase Structure and the Lexicon*, pp. 109–137. Dordrecht: Kluwer.

Labelle, M. (1990). Predication, wh-movement, and the development of relative clauses. *Language Acquisition 1*, 95–119.

Labelle, M. (1996). The acquisition of relative clauses: Movement or no movement? *Language Acquisition 5*, 65–82.

Larson, R. (1987). Missing prepositions and the analysis of english free relative clauses. *Linguistic Inquiry 18*, 239–266.

Lieven, E. (2010). Input and first language acquisition: Evaluating the role of frequency. *Lingua 120*, 2546–2556.

Longobardi, G. (1994). Reference and proper names: A theory of n-movement in syntax and logical form. *Linguistic Inquiry 25*, 609–665.

Lukyanenko, C., A. Conroy, and J. Lidz (2014). Is she patting katie? constraints on pronominal reference in 30-month olds. *Language Learning and Development 10*, 328–344.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. (3rd Edition ed.). Lawrence Erlbaum Associates.

Maratsos, M. (1974). Preschool children's use of definite and indefinite articles. *Child Development 45*, 446–455.

McDaniel, D. (1989). Partial and multiple wh-movement. *Natural Language and Linguistic Theory 7*, 565–604.

Modyanova, N. and K. Wexler (2007). Maximal trouble in free relatives. In *Boston University Conference on Language Development 32*.

Newmeyer, F. (2004). Against a parameter-setting approach to typological variation. *Linguistic Variation Yearbook 4*, 181–234.

O'Grady, W. (1997). *Syntactic Development*. London: University of Chicago Press.

O'Grady, W. (2005). *Syntactic Carpentry: An Emergentist approach to syntax*. Mahwah, NJ: Erlbaum.

Oiry, M. (2011). A case of true optionality: Wh-in situ patterns like long movement in french. *Linguistic Analysis 37*, 1–28.

Omaki, A., I. D. White, T. Goro, J. Lidz, and C. Phillips (2014). No fear of commitment: Children's incremental interpretation in english and japanese wh-questions. *Language Learning and Development 10*, 206–233.

O'Neill, D., J. W. Astington, and J. Flavell (1992). Young children's understanding of the role that sensory experiences play in knowledge acquistion. *Child Development 63*, 474–490.

Onishi, K. and R. Baillargeon (2005). Do 15-month-old infants understand false beliefs? *Science 308*, 255–258.

Otsu, Y. (1981). *Universal Grammar and Syntactic Development in Children: Toward a Theory of Syntactic Development*. Ph. D. thesis, MIT.

Patterson, G. and I. Capongro (2016). The puzzling degraded status of *who* free relative clauses in english. *English Language and Linguistics 20*, 341–352.

Pearl, L. (2016). Integrating conceptual and structural cues: Theories for syntactic acquisition. Talk given at SynLinks, University of Connecticut.

Pérez-Leroux, A. T. (1995). Resumptives in the acquisition of relative clauses. *Language Acquisition 4*, 105–138.

Pérez-Leroux, A. T., A. Gavarró, and T. Roeper (2011). Language typology and the acquisition of bare noun/dp contrasts. *Canadian Journal of Linguistics 56*(3), 413–422.

Pérez-Leroux, A. T., A. Munn, C. Schmitt, and M. Delrish (2004a). Bare nominals and the morphosyntax of number. In E. Cresti, T. Satterfield, and C. Tortora (Eds.), *Current Issues in Romance Linguistics*, pp. 217–231. Amsterdam: John Benjamins.

Pérez-Leroux, A. T., A. Munn, C. Schmitt, and M. Delrish (2004b). Learning definite determiners. In *Supplementary Proceedings of the 28th Boston University Conference on Language Development*.

Pérez-Leroux, A. T. and T. Roeper (1999). Scope and the structure of bare nominals: Evidence from child language. *Linguistics 37*, 927–960.

Rawlins, K. (2008). *(Un)Conditionals: An Investigation in the Syntax and Semantics of Conditional Structures*. Ph. D. thesis, UC Santa Cruz.

Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (Ed.), *Elements of Grammar*, pp. 281–337. Amsterdam: Kluwer.

Rizzi, L. (2005). On the grammatical basis of language development. In G. Cinque (Ed.), *The Oxford Handbook of Comparative Syntax*, pp. 70–109. New York: Oxford University Press.

Rizzi, L. (2013). Locality. *Lingua 130*, 169–186.

Roeper, T. (2011). The acquisition of recursion: How formalism articulates the child's path. *Biolinguistics 5*.

Roeper, T. and J. de Villiers (2011). The acquisition path for wh-questions. In J. de Villiers and T. Roeper (Eds.), *The Handbook of Generative Approaches to Language Acquisition*, pp. 189–246. Dordrecht: Springer.

Roeper, T. and E. Williams (Eds.) (1987). *Parameter Setting*. Dordrecht: Reidel.

Ross, J. (1967). *Constraints on Variables in Syntax*. Ph. D. thesis, MIT.

Safir, K. (1986). Relative clauses in a theory of binding and levels. *Linguistic Inquiry 17*, 663–689.

Schlenker, P. (2003). Clausal equations (a note on the connectivity problem). *Natural Language and Linguistic Theory 21*, 157–214.

Schmidt, E. and J. Pyers (2011). Children's understanding of the link between sensory perception and knowledge. In *33rd Annual Conference of the Cognitive Science Society*, Austin, TX.

Snyder, W. (2007). *Child Language: The Parametric Approach*. New York: Oxford University Press.

Thornton, R. (1991). *Adventues in Long Distance Moving: The Acquisition of Wh-Questions*. Ph. D. thesis, University of Connecticut.

Tomasello, M. (2000). First steps towards a usage based account of language acquisition. *Cognitive Linguistics 11*, 61–82.

Travis, L. (1984). *Parameters and Effects of Word Order Variation*. Ph. D. thesis, MIT.

Vikner, S. (1995). *Verb Movement and Expletive Subjects in the Germanic Languages*. New York: Oxford University Press.

von Fintel, K. (2000). Whatever. In *SALT 10*, pp. 27–40.

Warneken, F. and M. Tomasello (2009). Varieties of altruism in children and chimpanzees. *Trends in Cognitive Sciences 13*, 397–402.

Weir, A. (2014). *Fragments and Clausal Ellipsis*. Ph. D. thesis, University of Massachusetts Amherst.

Weist, R. and A. Zevenbergen (2008). Autobiographical memory and past tense time reference. *Language Learning and Development 4*, 291–308.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua 106*, 23–79.

Wexler, K. (2005). Cues don't explain learning: Maximal trouble in the determiner system. In E. Gibson and N. Pearlmutter (Eds.), *Processing and Acquisition of Reference*. Cambridge, MA: MIT Press.

Yang, C. (2003). *Knowledge and Learning in Natural Language*. Oxford University Press.