Explaining the Resultative Parameter

by

Daniel A. Milway

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Linguistics
University of Toronto

Abstract

Explaining the Resultative Parameter

Daniel A. Milway
Doctor of Philosophy
Graduate Department of Linguistics
University of Toronto
2019

This thesis proposes an explanatory account for the fact that some, but not all languages exhibit adjectival resultatives. It does so by synthesizing a number of earlier proposals. From Snyder (1995, 2016), I adopt the proposal that the setting of the resultative parameter must be acquired indirectly, being inferred from the presence or absence of bare stem compounding in the primary linguistic data. From Kratzer (2005), I adopt a small clause analysis of resultatives and the proposal that Snyder's parameter is related to inflectional features on adjectives. Finally, from Son and Svenonius (2008), I adopt the proposal that adjectival resultatives involve a single syntactic argument being shared by the primary and secondary predicates.

These proposals, combined with minimalist assumptions, yield: *(i)* a structural analysis of resultatives, in which they are represented by a resP adjoined to a VP, with a DP argument undergoing sideward movement between them, and *(ii)* an analysis of the parameter, according to which languages generate adjectival resultatives only if their lexicon contains uninflected categorizing heads. I then show that, under Chomsky's (2013, 2015) label theory, with some modifications, an adjectival resultative can be derived only if the result adjective is categorized by an uninflected *adj* head (*e.g.*, $\{adj_{\emptyset}, \sqrt{flat}\}$). This demonstration concludes the theoretical explanation of the resultative parameter.

The remainder of the thesis addresses some consequences of the explanation. First I address the appearance that my theory seems to wrongly rule out copular clauses and depictives in, for example, French. I argue that this stems from our conception of Agree, and present a theory of postsyntactic Agree that gives the right empirical results. Next, I address the previously unrecognized generalization that sideward movement from adjuncts is always

necessary when possible. Using Cinque's (1996) analysis for ACC-ing clauses, I show that this can be explained by a novel theory of adjunction synthesizing the pair-merge (Chomsky 2004) and late adjunction (Stepanov 2001) theories. Finally, I discuss the semantics that follows from my syntactic analysis. Although the semantic analysis seems implausible, I present corroborating evidence in its favour and argue that is more plausible than the alternatives.

Acknowledgements

This document was a long time in the making, and it certainly couldn't have happened without a supervisor with the generosity and intellectual dexterity of Elizabeth Cowper. Elizabeth, who was already retired by the time I asked her to be my supervisor, never seemed to be fazed by supervising or advising on projects that were outside her wheelhouse, even though those projects were decidedly incompatible with each other. But, my history with her goes farther back than the start of my thesis work, to LIN331 where she introduced me to the minimalist program, and later, in that time between undergrad and the MA when she met with me just to talk about syntax. Between the course and those meetings, I'm quite certain I wouldn't be where I am without her.

I am also indebted to Diane Massam and Michela Ippolito, who, with Elizabeth, formed my thesis committee. My meetings with them helped me to ground my research and avoid tunnel-vision, and more generally helped keep me honest. I also want to thank the rest of my defense panel: Susana Béjar, Nick LaCara, and Norbert Hornstein, for making my defense lively and actually fun.

Beyond my thesis committee I am indebted to the entire Linguistics Department at UofT. Thank you to the faculty in general and the syntax and semantics faculty in particular for creating a model of a collegial atmosphere to nurture research and researchers. Thank you to my fellow grad students, who were the best support network a young researcher could ask for. And last but certainly not least, thank you to the administrative staffers who made the department run: Mary Hsu, Jill Given-King, Jennifer McCallum, and Deem Waham.

Research, of course, is only one aspect of graduate school; there's also teaching, TA-ing, and grading to fill our days. There are too many to mention by name here, but I want to thank all the people I TA-ed for, and with, and those people who TA-ed for me. I also must acknowledge my union, CUPE 3902, which tirelessly worked to protect and advocate for me and my fellow workers. I came into grad school skeptical of unions in general but that skepticism evaporated in 2015 when the collective bargaining process and ensuing four week strike revealed the true colours of both the union and the university administration. The union revealed itself to be supportive, open, and protective of all its members, while the administration revealed itself to be almost the opposite. It definitely stung to see the administrators of the university that I had spent the better part of a decade at (undergrad and grad school) outright lie about me and my fellow workers, and try to pit us against each other, undergrads, and faculty, but that sting was lessened by the sense of solidarity I felt on the picket lines with my fellow members of that university. The strike was a difficult time but it did remind me that, as is the case with any organization, the university is defined by its membership, not its administration.

When I was dithering over whether or not to pursue a PhD, wondering if it was worth it, Derek Denis, then just a grad student, pointed out one of the great side-benefits of the

PhD: real lasting friendships. Perhaps it's too early to say, but I think he was right, I've found a good number of people that I feel grateful to call friends. Not least of all, I'm grateful to have had so many wonderful cohort-mates: Becky Tollan, Clarissa Forbes, Tomohiro Yokoyama, Jess Denniss, Emily Blamire, Michelle Yuan, and Jada Fung. I'm grateful to have met them and to have shared the ups and downs of graduate school with them. Thank you also to the many other wonderful people I got to know as fellow grad students from Joanna Chociej, Sarah Clarke, Richard Compton, Ailís Cournane, Radu Craioveanu, Derek Denis, Liisa Duncan, Ross Godfrey, Yu-Leng Lin, Mercedeh Mohaghegh, Alex Motut, Kenji Oda, Will Oxford, Christopher Spahr, Eugenia Suh, and others, who were the towering senior grad students who helped me to figure things out as I got started, to Julien Carrier, Emily Clare, Julie Doner, Erin Hall, Shayna Gardiner, Ruth Maddeaux, Emilia Melara, Patrick Murphy, and Na-Young Ryu, who were junior grad students figuring things out alongside me, to all the soon-to-be senior grad students, some of whom I was lucky enough to befriend, all of whom I'm proud to have shared a department with. I hope that the line of graduate students continues on, being there for each other in rough times, celebrating achievements together, and just blowing off steam at pub night.

Thank you to Dawn Whitwell for teaching me how to write jokes and giving me a creative outlet other than linguistics. Thank you to my friends Shawna Edward, Lisa Feingold, and Dan Reilly (aka "The Bit Acknowledgers"). Thank you to the many people who put out the podcasts that informed me, made me laugh and generally helped me get through the PhD with my sanity: Scott Aukerman, Adam Scott, Paul F. Tompkins, Dave Shumka, Graham Clark, Sean Clements, Hayes Davenport, John Hodgman, Jesse Thorn, Greg Kot, Jim DeRogatis, Howard Kremer, Kulap Vilaysack, Tim Batt, Guy Montgomery, Mike Mitchell, Nick Wiger, Jimmy Pardo, Matt Belknap, Garon Cockrell, Eliot Hochberg, Leo Laporte, Jeff Jarvis, Stacey Higginbotham, Gina Trapani, Julia Prescott, Allie Goertz, Lauren Lapkus, Russ Roberts, Jeremy Scahill, Jessica St Clair, Lennon Parham, and Andy Daly.

I'd like to thank my siblings and their families: my brother Tom, his wife Karen, and their children Declan, Rose, Violet, and Sadie; my brother Mike, his wife Brigid, and their children Imogen and Desmond; my sister Joan and her partner Jared; my brother Peter.

I'd like to thank Zoe, who came into my life only recently, but whose love and support has made this milestone all the sweeter.

Finally, Thank you to my parents Jim and Sheila. I wouldn't be where I am now, I wouldn't have achieved anything worthwhile, I wouldn't be the person I am today were it not for their love and support.

Contents

List of Tables

List of Figures

List of definitions

Chapter 1

Introduction

This thesis asks a seemingly simple question: Why do some but not all languages allow their users to generate adjectival resultatives? I write "*seemingly* simple" for reasons that are likely obvious to anyone reading this, but will clarify the reasons as a way of introducing what follows. There are, as far as I can tell, two complications inherent in questions of the form *Why P?*: one linguistic, one metaphysical. The linguistic complication is that *Why P?* presupposes that *P*. In order to even be justified in asking the question posed above, then, we must first demonstrate that there is a class of expressions that can be called adjectival resultatives, and that they are not found in every language. The metaphysical complication is due to the fact that a given *Why* question has an indefinite number of true responses, yet the appropriate response depends on the level of explanation that is sought. So, in order to answer the question, we must clarify the level of explanation we are seeking.

Beginning with the presupposition: is it the case that some but not all natural language grammars generate adjectival resultatives? The first thing we need to answer that question is a working definition of adjectival resultatives, which I give in definition 1.1 and which, in turn, depends on the definition of *secondary predicate* in definition 1.2.

Definition 1.1 (Adjectival Resultative)**.** An *adjectival resultative* is a secondary predication structure, whose secondary predicate is (*a*) an adjective (phrase) and (*b*) interpreted as describing a state directly caused by the event described by the primary predicate.

Definition 1.2 (Secondary Predication)**.** A *secondary predication* structure is a monoclausal structure containing a constituent consisting of a verb (phrase) (V) an argument (DP) and another element (SP), such that SP is interpreted as a predicate, and DP is an argument of both V and SP.

A canonical example of an adjectival resultative is given in (1.1).

(1.1) Natalie hammered the metal flat.

This is a secondary predication structure in the sense that it contains a constituent *hammer the metal flat* which contains a verbal and an adjectival predicate (*hammer* and *flat*, respectively) and a DP *the metal* which is an argument of both predicates. Furthermore, it is an adjectival resultative because its secondary predicate is the adjective *flat*, which describes a state caused by the hammering event. Resultatives contrast minimally with depictives, like the one in (1.2)—secondary predication structures whose secondary predicate describes a state not caused by the primary predicate.

(1.2) Heather ate the fish raw.

This is an example of secondary predication, with *ate* being the verb, *raw* being the secondary predicate, and the argument *the fish* shared between the two. It is not a resultative because, in the situation it describes, the rawness state is not caused by the eating event. So, part of the presupposition is true: resultatives exist in at least one language. Snyder (1995, 2001) demonstrates that adjectival resultatives exist in a number of other languages, including ASL, Dutch, German, Khmer, Japanese, Korean, Hungarian, Mandarin, and Thai.

(1.3) Examples of adjectival resultatives[1]

a. J-O-H-N PAINT CHAIR RED (ASL, Kentner and Wilbur 2018)
"John painted the chair red."

b. Hans hämmert das Metall flach. (German)
Hans hammered the metal flat
"Hans hammered the metal flat."

c. A munkás lapos-ra kalapácsolta a fémet. (Hungarian)
the worker flat-TRANS hammer-PST the metal
"The worker hammered the metal flat."

d. John-ga teeburu-o kiree-ni hui-ta. (Japanese)
John-NOM table-ACC clean wipe-PST
"John wiped the table clean."

e. Kira wai daik kpaet. (Khmer)
Kira hit metal flat
"Kira beat the metal flat."

f. John-i teibl-ul kekuti tak-at-ta (Korean)
John-NOM table-ACC clean polish-PST-COMP
"John wiped/polished the table clean."

g. Ta ba tie guan da ping. (Mandarin [tones omitted])
(s)he BA iron pipe hit flat
"(S)he beat the iron pipe flat."

h. Ja: t'up lo:haʔ haj bae:n (Thai [tones omitted])
NEG-IMPER hammer metal HAJ (be-)flat
"Don't hammer the metal flat."

[1]Unless otherwise noted, these examples are drawn from Snyder (2001)

Furthermore, Snyder demonstrates that some languages seem to be incapable of generating adjectival resultatives, expressing resultatives periphrastically instead.

(1.4) Examples of periphrastic resultative.[2]

a. Lokoda taroktu el hadide haete? osbohoh mosoto?han. (Arabic, Egyptian)
Lokoda beat the iron until it.became flat
"Lokoda beat the metal *until it became* flat"

b. Gorri-z atz-azal-ak pintatzen ari naiz (Basque)
red-with/in finger.covering-PL painting AUX
"I am painting my finger nails *with/in* red."

c. Jean a martelé le métal jusqu'à ce qu'il était plat (French)
John has hammered the metal until-to that that-it be.PST flat
"John hammered the metal until it was flat."

d. Dani tzavaA ?it ha-bayet bi-?adoom. (Hebrew, Modern)
Dani painted P(ACC) the-house in-red
"Dani painted the house *in* (the colour) red."

e. [Tukang pande-nipun] mande wesi ngantos gepeng. (Javanese)
worker forge-POSS beat iron until flat
"The blacksmith beat the iron *until* (it was) flat."

f. Joe abisi ndako na modobo motani. (Lingala)
Joe he.paint house with paint red
"Joe painted the house *with* red paint."

g. Ivan pokrasil dom v krasnyj tsvet. (Russian)
John paint.PST house in red colour
"John painted the house *in* the colour red"

h. John je ofarbao kucu u crveno. (Serbo-Croatian)
John is painted house in red
"John painted the house *in* (the colour) red."

i. Juan golpeó el hierro hasta que estaba plano. (Spanish)
John beat.PST the iron until that be.PST flat
"John beat the iron UNTIL it was flat."

Our presupposition, then, seems to hold; some, but not all languages exhibit adjectival resultatives.

Our second issue—that of deciding what we mean by *why*—I believe is a far more interesting one, as answering it requires us to be explicit about the broader goals of our inquiry. If our interest is historical linguistics or language variation and change, then we might be interested in migration patterns and language contact situations, and how they do or do not correlate with a language's ability to generate resultatives, or with the social factors linked to resultatives. This thesis, however, is a work of largely theoretical generative syntax, so our *why* question is actually two questions: What essential property or properties do grammars

[2]These examples are drawn from Snyder (2001) unless otherwise noted.

that generate resultatives have, that grammars that do not generate resultatives lack? And how is that property or set of properties acquirable by children from the primary linguistic data? Note that I have framed the acquisition question as dependent on the grammatical question—likely a reflection of my training as a syntactician—but I don't believe that one question is *logically* prior to the other. Snyder (1995, 2001), for instance, takes the grammatical question to be dependent on the acquisition question. I believe the questions are interdependent, meaning that the correct answer to one should at least be consistent with the correct answer to the other. The simplest situation, however, would be that the correct answer to each question entails the correct answer to the other; in other words, that a single statement would provide an answer to both questions.

For reasons that have little to do with the content of linguistic theory or its empirical base and a great deal to do with the social, cultural, and political atmosphere of modern scientific research, the two questions that I have posed above are not commonly treated as interdependent. Syntacticians tend to focus on the grammatical question, and consider the acquisition question to be secondary at best, while acquisitionists tend to consider the reverse to be the case. This leads to syntactic proposals where the acquisition question is ignored or treated as an afterthought, and acquisition studies which do not fully address how their results could be integrated into linguistic theory. With this thesis, I hope to avoid this pitfall. That is, I aim to develop a theoretical explanation of the resultative parameter that takes the acquisition question to be a crucial criterion for the success of my proposal; in other words, I assume that a grammatical theory of resultatives is adequate only if it answers the acquisition question. This statement is likely uncontroversial among generative syntacticians; indeed it is perhaps an unstated criterion of all generative syntax. I make it a stated criterion here as a way of ensuring that readers can hold me to it.

The answer I will argue for here is that a grammar generates resultatives only if it also generates adjectives without φ-features. I argue that this parameter is both acquirable and consistent with Snyder's results. Since φ-features on adjectives manifest themselves as agreement morphology, their presence or absence is directly detectable, and therefore acquirable. The demonstration that Snyder's results can be derived from the presence or absence of φ-features on adjectives, however, is highly technical, based on a non-standard version of so-called minimalist theory. In chapter 2, I argue that the non-standard aspects of the theory I assume, in fact, follow from entirely standard assumptions in contemporary syntactic theory, which I discuss in detail in section 1.1. With these assumptions—standard and non-standard—in place, I assess the current syntactic analyses of adjectival resultatives in chapter 3 before synthesizing these analyses into my own analysis in chapter 4. The analysis that I settle on, shown in figure 1.1, is modified from that of Kratzer (2005). Note three non-standard properties of this analysis: resP is adjoined to VP, the shared argument *the metal* undergoes sideward movement, and that movement is into a θ-position. These

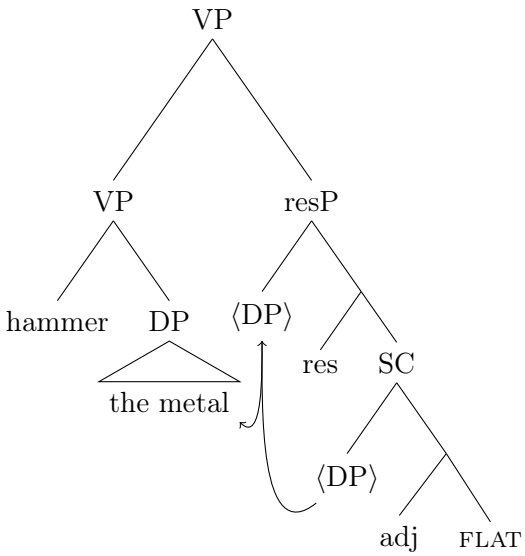


Figure 1.1: The structure of resultatives proposed in chapter 4

properties will be explained in a principled manner throughout the thesis.

Before my ultimate explanation of the resultative parameter in chapter 5, I introduce and discuss label theory (Chomsky 2013, 2015), a new modification to minimalist syntactic theory. Finally, in chapter 6 I show how the assumptions, hypotheses, and analyses that I mention above allow me to derive Snyder's (2016) correlation between resultatives and bare stem compounding.

While this derivation, which ends part I of this thesis, might seem to fully explain the resultative parameter, and therefore fulfill the stated goal of my thesis, I go on in part II to consider the broader implications of the account I give in part I. For instance, my account of the resultative parameter seems to be too strong, eliminating the possibility of expressions derived from small clauses in languages which do not have resultatives. That is, my account would seem to predict that a language like French would fail to generate copular clauses and depictives like the examples in (1.5) and (1.6), respectively.

(1.5) Jeanne est grand -e.
Jeanne is tall -FSg
"Jeanne is tall."

(1.6) Marie mange la viande crue
Marie eats the.FSg meat raw.Fsg
"Marie eats meat raw"

In chapter 7, however, I propose a refinement of the grammatical architecture which places the operation Agree outside the narrow syntax, before Label, and show that this proposal eliminates the faulty prediction made in part I.

Chapter 8 discusses a puzzling aspect of the structure I propose for resultatives (see

figure 1.1)—namely, the fact that sideward movement from [Spec, res] to [Comp, V] seems to be obligatory. I highlight the puzzling nature of this by discussing the analysis of ACC-ing clauses given by Cinque (1996), in which such sideward movement is also obligatory. In order to explain these analyses, I make two modifications to Chomsky's label theory, modifications which I argue are independently motivated. I then show that obligatory sideward movement follows from the resulting version of label theory.

Finally, in chapter 9, I discuss the interpretation of resultatives and constructions with similar structures (depictives and some ACC-ing constructions). I argue that given the right theory of event semantics, these structures with such seemingly different interpretations are, in fact, very similar. They are all species of secondary predication structures in which the primary and secondary predicates take the same eventuality (or situation) argument.

1.1 Theoretical Context

The general theory that I assume here is a variety of what is called *minimalist syntax* after Chomsky's (1995) *The Minimalist Program*. Using the term *minimalism* to refer to a theory of grammar, however, is perhaps incorrect, as minimalism is a metatheoretical position. The contrast between theory and metatheory that I assume here is due to Chametzky (1996).[3]

[3]Chametzky makes a three-way distinction between metatheoretical, theoretical, and analytic work:

> *Metatheoretical* work is theory of theory, and divides into two sorts: general and (domain) specific. *General* metatheoretical work is concerned with developing and investigating adequacy conditions for any theory in any domain. So, for example, it is generally agreed that theories should be (1) consisted and coherent, both internally and with other well-established theories; (2) explicit; and (3) simple. This sort of work is philosophical in nature [. . .]. *Specific* metatheoretical work is concerned with adequacy conditions for theory in a particular domain. So, for example, in linguistics we have Chomsky's (1964; 1965) familiar distinctions among observational, descriptive, and explanatory adequacy. Whether such work is "philosophy" or, in this case, "linguistics" seems to me a pointless question.
>
> *Theoretical* work is concerned with developing and investigating primitives, derived concepts, and architecture within a particular domain of inquiry. This work will also deploy and test concepts developed in metatheoretical work against the results of actual theory construction in a domain, allowing for both evaluation of the domain theory and sharpening of the metatheoretical concepts. Note this well: deployment of metatheoretical concepts is *not* metatheoretical work; it is theoretical work.
>
> *Analytic* work is concerned with investigating the (phenomena of the) domain in question. It deploys and tests concepts and architecture developed in theoretical work, allowing for both understanding of the domain and sharpening of the theoretical concepts. Note this well: deployment of theoretical concepts is *not* theoretical work, it is analytic work. Analytic work is what overwhelmingly most linguists do overwhelmingly most of the time. This is as it should, and indeed must, be: an empirical discipline only exists insofar as there is a community of scientists investigating the domain. For linguistics to be the science of language, this must be

This distinction is evident when one considers the stark contrasts between those theories of grammar that are referred to as *minimalist*. For instance, Borer (2005a,b, 2013), Chomsky (2000a), Epstein and Seely (2006), Frampton and Gutmann (2008) and Hornstein (2009) all develop distinct minimalist theories of syntax. They all, however, share a set of assumptions, likely due to their shared Chomskyan heritage. Since this thesis shares that heritage, it also shares those assumptions which I will list and explain below.

Most fundamentally, minimalist theories of grammar share a model of the language faculty called the *Y Model* or the *T Model*. In this model, a narrowly syntactic "module" operates

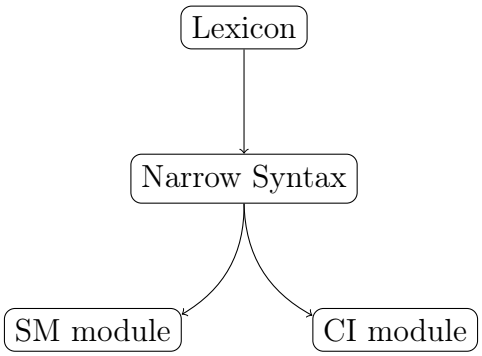


Figure 1.2: The Y Model of the language faculty

on items drawn from a lexicon to generate structures that are evaluated by a pair of modules: the Sensorimotor (SM) module, commonly called the morphophonology, or PF, which is responsible for external expression, and the Conceptual-Intentional (CI) module, commonly called the semantics, or LF, which is responsible for interpreting structures for use in internal systems of thought. Beyond this, there is significant disagreement among minimalist theorists.

Another assumption common to minimalist syntacticians is the syntactic operation Merge. The primary (and, in some theories, only) syntactic operation, Merge combines pairs of syntactic objects (*i.e.*, lexical items or syntactic structures) to form larger syntactic objects. The standard formulation of Merge is given in (1.7)[4].

(1.7) Merge$(\alpha, \beta) = \{\alpha, \beta\}$
iff both α and β are syntactic objects.

Merge is responsible not only for creating new structures but also for syntactic displacement. To my knowledge, this ability was discovered by Chomsky (2004) when he distinguished between the two logically possible cases of Merge: external and internal. An instance of Merge(α, β) is external if neither α nor β contains the other, and internal if α contains β or

where linguists do their work. (Chametzky 1996, xvii*ff*)

[4]Hornstein (2009) departs from the standard formulation, defining Merge as concatenation rather than set formation.

vice versa. Both cases, of course, create a new object, but in the case of internal Merge, the operation has the effect of creating displacement as shown in (1.8).

(1.8) $\text{Merge}(\alpha, \{\ldots\alpha\ldots\}) = \{\alpha, \{\ldots\alpha\ldots\}\}$

Again, beyond the basics discussed above, there is little consensus among minimalist syntacticians about the nature and operation of Merge.

While the Y Model and Merge seem to be the only instances of true consensus among minimalist syntacticians, there is growing agreement about the underlying representation of certain types of words. Specifically, many minimalist syntacticians now assume that a lexical word, like the noun *chair*, minimally consists of an acategorial root and a categorizing head (Borer 2005a; Marantz 1997). This is commonly expressed in the formalism of a vocabulary insertion rule from the theory of Distributed Morphology as in (1.9).

(1.9) $chair \leftrightarrow \left\{n, \sqrt{\text{CHAIR}}\right\}$

There are competing views of the syntactic nature of morphological words, but as of this dissertation's writing this is a standard view, a view to which I subscribe, and therefore, I will not explicitly argue for it.

The theory I assume is not entirely standard, though. There are a number of assumptions that I make, which will certainly raise the eyebrows of many if not most contemporary syntacticians. I discuss these assumptions in chapter 2.

1.1.1 Minimalism and the Strong Minimalist Thesis

The minimalist program can be viewed as an effort to simplify Government and Binding theory without losing its empirical coverage. That is, a minimalist analysis is one that compares two hypotheses that have roughly equivalent empirical power, and chooses the simpler one. However, as Chomsky (1965) discusses, there is no such thing as an absolute measure of simplicity. Consider, for instance, the following equivalent expressions of arithmetic using the more standard infix notation in (1.10) and lambda calculus in (1.11).

(1.10) $3 + 2$

(1.11) $\lambda f \lambda x.((\lambda f \lambda x.f(f(fx)))f(\lambda f \lambda x.(f(fx))(fx)))$

While it may seem obvious that (1.10), which consists of a mere three symbols, is simpler than (1.11) with its 41 characters, it becomes less obvious when we compare, as wholes, the systems that they are drawn from. Performing arithmetic with infix notation requires rote memorization of the results of single-digit addition, multiplication, subtraction, and division, and rather complex algorithms for larger numbers (*e.g.*, long division). The lambda calculus, on

the other hand, uses two very simple operations—α-conversion and β-reduction—requiring no rote memorization for their application. From this standpoint, the lambda calculus is vastly simpler. The point here is that a judgment of simplicity depends on the choice of simplicity metric.

This is not to say that the choice of simplicity metric is arbitrary. On the contrary, since any choice of simplicity metric will be a major factor in deciding between theories, that choice must be justified. The main simplicity metric of the minimalist program is the Strong Minimalist Thesis (SMT) which states that the language faculty is an optimal solution to interface conditions (Chomsky 2001). One of the justifications for SMT, often repeated by Chomsky, comes from evolutionary biology. It begins with two observations, the results of several decades of linguistics research. The first observation is that the human language faculty is unique in the biological world, that "nothing does language like we do," to use Norbert Hornstein's oft-repeated formulation. The second observation is that the language faculty is uniform across our species, that a child born on the island of Inishmore, but raised in Hanoi would acquire Vietnamese with the same ease as a child born and raised in Hanoi. These observations suggest that the language faculty emerged quite suddenly, likely due to a single genetic mutation in a single individual. It follows from this that whatever portion of our cognitive system that is specific to language must be very simple.

The SMT provides the following principles for minimalist syntactic analysis and theorizing:

1. Assume the simplest possible recursive syntax (*i.e.*, one that consists only of simplest Merge).
2. Assume that no other module of the language faculty is capable of recursion.
3. When you encounter a proposed property of the language faculty or principle of linguistic theory either
   (a) show that it can be reduced to Merge,
   (b) show that it can be reduced to interface conditions,
   (c) show that it can be reduced to independent principles, or
   (d) show that it can be reduced to a combination of Merge, interface conditions, and independent principles.
4. It is only when we can show that any such reduction is impossible that we may modify our assumptions.

One might object that this places an unreasonable burden on linguists. The principles, however, follow from the basic principles of science, and they are inapplicable to the study of human language only insofar as human language is immune to scientific inquiry. So, to abandon these principles at the first sight of difficulty is to abandon the scientific approach to understanding the language faculty.

Part I

Explaining Resultatives

Chapter 2

Non-standard Theoretical Assumptions

This dissertation rests on a number of non-standard theoretical assumptions and draws a few non-standard distinctions, which I will defend in this chapter. My defence of the assumptions, however, will not be an argument that they are true, as the truth of any theoretical statement ultimately depends on the empirical facts. Rather, my defence will actually be an offence; I will argue that the standard assumption is, in fact, ill-founded. So, in a sense, I will be rejecting standard assumptions rather than making non-standard ones. So, I will identify problems with the standard assumptions, and present non-standard assumptions which do not face those problems. The distinctions I draw, in contrast, will not be defended, but rather explained and clarified.

2.1 The Θ-Criterion

The θ-criterion as standardly assumed was first formulated by Chomsky in *Lectures on Government and Binding* (LGB) as (2.1).

(2.1) Each argument bears one and only one θ-role, and each θ-role is assigned to one and only one argument. (Chomsky 1981, p. 36)

In a footnote, Chomsky justifies this criterion, saying

> The second clause of [the θ-criterion] is well-motivated. To say that each θ-role must be filled implies, for example, that a pure transitive verb such as *hit* must have an object, that a verb such as *put* or *keep* (with the sense they have in *put it in the corner*, *keep it in the garage*) must have the associated PP slot filled, etc. The additional requirement that each θ-role must be filled by only one argument will, for example, exclude the possibility that a single trace is associated with several argument antecedents, a possibility ruled out in principle under the Move-α theory. (Chomsky 1981, p. 139)

I would agree that the second clause of (2.1), that each θ-role is assigned to a single argument, is well-motivated by the empirical considerations Chomsky cites, and as such I will not reject that portion of the θ-criterion. The first clause, however, is motivated mainly by theoretical concerns of LGB, that is, its empirical motivation is indirect at best.

The nature of the LGB theory is such that its various hypotheses and principles are connected to each other in a web-like network. As a result, the first clause of the θ-criterion depends on various other theoretical statements and various other theoretical statements depend on it. So, rather than attempting an exhaustive enumeration of the links between the θ-criterion and the other theoretical statements of LGB, I will present what I consider to be the best argument in favour of the θ-criterion and argue that its premises have since been rejected within syntactic theory.

The first premise is the now-familiar Y- or T-model of grammar shown in Figure 2.1, which LGB theory continues from earlier theories. According to this model, a syntactic derivation has four levels of representation (D-structure, S-Structure, PF, and LF), and each step in the derivation is performed by the application of a subset of the transformational rules. S-structures are derived by the applying of Move-α to D-structures, LFs are derived by applying QR (and maybe Move-α) to S-structures, and PFs are derived by applying "stylistic rules" to S-structures (Chomsky 1981, p. 18). The exact formal properties of the

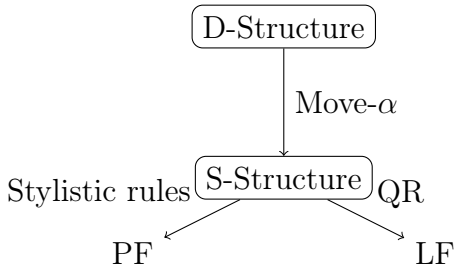


Figure 2.1: The architecture of the grammar in LGB

transformations are not important for this discussion. What is important, is that all syntactic displacement is the result of one of these transformations.

The second premise is the projection principle, which states that lexical properties must be represented at all levels of syntax. Since θ-roles are lexical properties of (at least) verbs, they must be represented at all levels of syntax. Consider the verb *hit*, whose lexical entry specifies that it needs a patient argument. The projection principle requires that at D-Structure and S-Structure *hit* must have assigned a patient θ-role to an argument and therefore, assuming the patient θ-role is assigned to Comp,V, there must be a nominal in the complement position of *hit* at both D-Structure and S-Structure.

With these assumptions, it follows that no single argument can receive more than one θ-role. Suppose there is a derivation in which a single argument X receives two θ-roles Θ1 and Θ2. According to the projection principle, X must be marked with both Θ1 and Θ2

at D-Structure. Since each θ-role is associated with a unique structural position, it follows that X must be in two distinct positions at D-Structure. The only way an argument can be in multiple positions is if it has undergone Move-α. Move-α, however, maps D-Structures to S-Structures. Therefore, an argument cannot be in two positions at D-Structure, and furthermore, cannot be multiply θ-marked at D-Structure. If an argument cannot be multiply θ-marked at D-Structure, then it cannot be multiply θ-marked at all.

Thus we are able to derive the first clause of the θ-criterion from other principles. These principles, however, have either been rejected or problematized since their statement in LGB. Since *The Minimalist Program* (Chomsky 1995), generative theories have largely dispensed with D-Structure and S-Structure. Without these levels of representation, the projection principle (as formulated in LGB) is effectively meaningless, and without the projection principle, there is no more basis for the first clause of the θ-criterion.[1] Therefore, I will not assume the first clause of the θ-criterion.

2.2 Last Resort

The goal of the minimalist program is to clean up syntactic theory by explaining unnecessary principles in terms of necessary principles. In nearly one fell swoop, Chomsky (1995) eliminates a number of the complications built into LGB theory (S-Structure and D-Structure chief among them). In *The Minimalist Program* (MP), however, Chomsky was not able to explain two apparent imperfections: displacement and uninterpretable features.[2] Chomsky formalized displacement as the operation Move, which is more complex than Merge, and proposes that, unlike Merge, which "comes for free," every instance of Move must be triggered by the need to satisfy an uninterpretable feature. The simple operation Merge is preferred for economy reasons, while movement, Chomsky argues, is a "last resort." In work subsequent to MP, however, Chomsky proposes that Move is actually a sub-type of Merge, called Internal Merge. Merge and Move are different names for the same operation, there is no reason to think that Move is more computationally complex, and therefore there is no reason to think that movement is a last resort.

In "Beyond Explanatory Adequacy" (Chomsky 2004) Chomsky makes this rejection of Last Resort explicit:

> [Narrow Syntax] is based on the free operation Merge. [The Strong Minimalist Thesis] entails that Merge of α, β is unconstrained, therefore either external or internal. Under external Merge, α and β are separate objects; under internal

[1] Note that this is far from a knock-down argument against the first clause of the θ-criterion. In fact, the θ-criterion, as it is expressed in LGB may still hold, but in order to show this, we need empirical arguments.

[2] The framework developed in MP also does not explain projection/labelling, but Chomsky does not recognize this as an imperfection until "Problems of projection" (Chomsky 2013). More on this in Chapter 5.

> Merge, one is part of the other, and Merge yields the property of "displacement," which is ubiquitous in language and must be captured in some manner in any theory. It is hard to think of a simpler approach than allowing internal Merge (a grammatical transformation), an operation that is freely available there is. (Chomsky 2004, p. 110)

This line of reasoning bears discussion in light of the fact that Last Resort is still standardly assumed by self-described minimalist syntacticians. In fact, several syntacticians expand Last Resort to External Merge arguing that neither type of Merge "comes free" (Frampton and Gutmann 2008; Pesetsky and Torrego 2006; Wurmbrand 2014; Yokoyama 2015). This proposal is understandable from a historical perspective, but ultimately misguided in my opinion.

Perhaps the most attractive aspect of a constrained Merge syntax is the purported gains in computational efficiency. To illustrate this, Frampton and Gutmann (2008) consider the incomplete product of a doomed derivation in (2.2).

(2.2) it to be believed Max to be happy

In a free Merge syntax, according to Frampton and Gutmann (2008), the derivation must continue for an indefinite time until a phase head is merged. At this point, the derivation will crash due to a Case Filter violation. A constrained Merge syntax, however, could halt as soon as the derivation becomes doomed, say, when *it* is merged. This would save us the indefinite number of steps it takes to merge a phase head and is, therefore, more efficient. This I take to be a species of the argument that free Merge systems are inefficient because, in addition to the infinite array of convergent derivations they must generate, they also generate an infinite array of crashing derivations, whereas constrained Merge systems only generate to "convergent" derivations. However, as I show below, when we investigate the nature of constrained Merge, the purported gains in efficiency in one part of the system come at the expense of another part of that same system. That is, a system with constrained Merge is not inherently more efficient than a system with free Merge, and since there is no efficiency difference between the two systems, the choice between them will have to be based on some other metric.

At a minimum, abstracting away from Agree operations, each version of the syntax will have a Merge operation and a Transfer operation. So, the free Merge syntax will consist of an unconstrained Merge_F and a constrained Transfer_F as defined in (2.3).

(2.3) a. $\text{Merge}_F(\alpha, \beta) = \{\alpha, \beta\}$
b. $\text{Transfer}_F(\gamma) = \langle \textsc{sem}(\gamma), \textsc{phon}(\gamma) \rangle$ iff $\text{Filter}(\gamma) = \text{F}$
(where α, β, and γ are syntactic objects.)

A constrained Merge syntax, then, would consist of a constrained Merge_C and an unconstrained Transfer_C.

(2.4) a. $\text{Merge}_C(\alpha, \beta) = \{\alpha, \beta\}$ iff $\text{Satisfy}(\alpha, \beta) = \text{T}$
b. $\text{Transfer}_C(\gamma) = \langle \textsc{sem}(\gamma), \textsc{phon}(\gamma) \rangle$
(where α, β, and γ are syntactic objects.)

If we assume that both theories can be made descriptively adequate, then the Satisfy predicate will have the same net effect as the Filter predicate required for the free Merge system. So, given any pair of syntactic objects α and β, Satisfy must be able to evaluate whether $\text{Merge}(\alpha, \beta)$ is allowed. And since there are an infinite number of derivable syntactic objects, even in a constrained Merge syntax, Satisfy must be able to evaluate an infinity of possible β's against each possible α. For any given syntactic object, then, there is an indefinite number of objects which will merge with that object and an indefinite number that will not, and the only way to know if Satisfy is true of a pair of objects is to check. So, much like free Merge syntax suffers from an infinity of crashes, constrained Merge syntax suffers from an infinity of failed Merge operations.

A constrained Merge theorist might still object by saying that Satisfy is a local operation, while Filter is a global operation, and local operations are to be preferred if we care about computational complexity. Again, this is an intuitively attractive argument, but not obviously valid. Consider the following thought experiment. Suppose you are a TA charged with grading a quiz by your somewhat maniacal instructor. Part of the instructor's mania is that they require all quizzes to consist of 40 equally weighted questions and be graded out of 10 points. What is the most efficient procedure for assigning a grade to each quiz? Two types of procedure suggest themselves. The first, which I will call the Local Only procedure, is to assign each correct answer a value of 0.25 points and then add up all of the points. The second, which I will call the Local/Global procedure, is to assign each correct answer a value of 1 point, add up all of the points, and divide by 4 (perhaps with a calculator). Since humans are very good at counting by increments of 1, and calculators are very good at dividing by 4, while neither is very good at counting by increments of 0.25, the Local/Global procedure is likely to be more efficient than the Local Only procedure. The moral of this story: One machine's global procedure is another machine's local procedure.

There is also a methodological rationale for preferring a free Merge framework which is slightly counterintuitive so I would like to dwell on it for a moment. My reason for assuming a free Merge framework is that it creates, or rather, lays bare, more problems for us to solve. Why is this preferable? Shouldn't we prefer the theory with fewer problems? Intuitively, we should prefer the less problematic theory, but this depends on how we count a theory's problems. I would like to argue that, while free Merge theories pose a greater number of problems than constrained Merge theories, the sheer weight of the problems posed by each

type of theory is equal to that of the other. Furthermore, the problems of free Merge can be made into empirical questions more readily than those of constrained Merge.

In order to argue in favour of free Merge, I will present one argument against it and show how that argument actually strengthens my claims in the previous paragraph. The argument comes from Frampton and Gutmann (2008) and states that a free Merge theory of grammar must posit "filters" to rule out the non-converging structures that its syntax generates, and the last thing we want is a flourishing of filters. I could not agree more with their assessment, but where they see a bug, I see a feature. So-called "filters" are not attempts at explanation, but descriptions of generalizations in need of explanation.

Take, for instance, the remaining clause of the θ-criterion, given in (2.5).

(2.5) [E]ach θ-role is assigned to one and only one argument. (Chomsky 1981, p. 36)

To propose a θ-filter, then, would be to say that those derivations that violate (2.5) crash at an interface, presumably the CI interface. For a theorist, this filter is actually a question or series of questions: Why is it that only those derived structures that satisfy (2.5) are valid CI objects? That question may not be empirical, but it invites hypotheses which may lead to empirical questions that we don't currently know how to ask. Very likely, due to the interface nature of the questions, their answers will not be narrowly linguistic.

Now, consider the situation constrained Merge puts the theorist in. For Frampton and Gutmann (2008), the θ-criterion is expressed by "selectional features" on heads which must be satisfied immediately. This leads to a number of questions: What is the nature of these selectional features? How are they related to, say, φ-features? Do they exist independently of the narrow syntax? Why do they need to be satisfied first? And so on. I, for one, don't have the slightest clue how to proceed in answering or even sharpening these questions, and there don't seem to be any clues in the offing from constrained Merge theorists.

2.3 Long Distance Agree

Strongly associated with, but distinct from Last Resort is the notion of an Agree operation. Chomsky (2000a, p. 101) introduces Agree as an operation "which establishes a relation (agreement, Case checking) between an LI α and a feature F in some restricted search space (its *domain*)", and virtually every syntactician who identifies as a minimalist assumes that such an operation is part of the syntactic computation. In fact, there is an ongoing debate as to the exact nature of Agree (Bjorkman and Zeijlstra 2014; Preminger 2013; Zeijlstra 2012), but there is nonetheless a broad consensus with respect to the general properties of Agree. All of the versions of Agree have three components which I will call Probe, Match and Satisfy. If we take Agree to operate on a head H and a feature F, then Agree states that

F Satisfies H if and only if H Probes and Matches F.[3] Probe is a predicate that evaluates to true if and only if a structural relation R holds between H and F. In virtually all versions of Agree, R is some combination of c-command and (relativized) minimality. The versions tend to differ on the directionality of R, that is, whether H c-commands F or vice-versa. Match is a predicate that evaluates to true iff H bears some feature of the same type as F. So, a head bearing unsatisfied φ-features will Match φ-features. If Probe and Match succeed, then F can Satisfy H, where satisfaction is generally taken to be feature valuation. So, a copula bears unvalued φ-features which are valued by the inherent φ-features of a DP.

The prototypical evidence for this conception of Agree comes from morphological agreement in existential constructions like (2.6), where the verb agrees with a DP in its c-command domain rather than the expletive *there* in its specifier.

(2.6) There are three plates in the sink.

As I see it, though, the main reason for keeping long-distance Agree in the Narrow syntax is that it seems to save Last Resort theories of movement from a rather vicious look-ahead problem. As I discussed in the previous section, the earliest version of Last Resort held that a movement operation proceeds only if it satisfies some uninterpretable feature. Under this theory, however, feature satisfaction occurs only as a result of movement, which means that in order to know if a movement operation satisfies a feature, we need to perform that movement. This has the effect of saying that Move(α) is defined iff Move(α) is defined. If instead, Move is preconditioned on Agree, then we no longer have an infinite regress.

Within the group of minimalist syntacticians, there is a somewhat disparate minority that rejects the orthodoxy of Long Distance Agree for a number of reasons. In the remainder of this section I will discuss the main theoretical[4] argument against a narrow syntactic notion of Agree, which is that this notion of Agree introduces a redundancy into the language faculty which violates the scientific principle of parsimony.

Hornstein (2009) argues that Agree is redundant with Merge and should, therefore, be rejected as an operation of the Narrow Syntax. Once Move is unified with Merge in the form of Internal Merge, then Merge becomes an operation capable of creating non-local dependencies. Since the creation of non-local dependencies is the sole purpose of Agree, the standard minimalist model of grammar has two operations that create non-local dependencies. This would not be an issue, however, if the types of dependencies created by Merge were distinct from those created by Agree, but they seem to be nearly identical. As I discussed above,

[3]Depending on the theory, there may be additional requirements placed on what sorts of heads and features can undergo Agree. For instance, many theories assume that H or F must be active to undergo Agree, where the precise definition of *active* depends on the theory in which it is situated.

[4]Arregi and Nevins (2013) and Bobaljik (2008) present **empirical** rather than *theoretical* arguments in favour of a post-syntactic conception of Agree, but their data and analyses are too involved to do justice in this section.

each instance of Agree is restricted to the c-command domain of some head, with an added (relativized) minimality restriction. Therefore, the dependencies created by Agree are restricted to those that obey c-command and (relativized) minimality. Dependencies created by Merge have exactly the same restrictions: C-command is definable in terms of Merge, and minimality is traditionally a restriction on movement. The only difference between Agree dependencies and I-Merge dependencies is that Agree dependencies don't appear to show displacement, while I-Merge ones do. Hornstein (2009) argues that this is not enough of a difference to justify an entirely distinct operation. Rather than assuming a novel operation, we can just take Agree to be a variety of movement (and therefore a variety of Merge) in which the lower copy is pronounced. Doing this would eliminate a redundancy, and push our study of agreement to the interfaces, a result I believe to be theoretically attractive.

Of course, theoretical attractiveness is not enough to justify an assertion as true. Only accordance with the facts can justify a theoretical claim. This is difficult for a negative claim, such as the rejection of long-distance Agree, since no amount of observation or exploration can definitively justify it. Rather, the task of the minimalist, given theory T which accounts for the set of facts C, is to show that a simpler theory T′ can also account for C,[5] and the only way to demonstrate this is to assume the simpler theory and show how it is able to explain the facts in question. That will be my strategy in this thesis: I will simply assume that there is no Long Distance Agree in the Narrow Syntax, and go about explaining the facts in question with the simplified theory.

2.4 Terminological notes

The subject matter of this thesis is often called the "syntax-semantics interface," a term which I have discovered is ambiguous, probably due to the fact that it is constructed from three ambiguous terms.

The term *syntax* has (at least) three senses which seem to be used in generative grammar circles. The first sense, which I will call the sociological sense, is that *syntax* is what syntacticians do. For instance, θ-theory belongs to the domain of syntax under this sense, because syntacticians care about it, while semanticists tend not to. However, θ-theory deals at least partially with meaning so it, at least, intersects with semantics. This sense would be useful if this dissertation were an intellectual history of generative syntax, but since this is a work of syntactic theory, I will not use this sense.

The second sense, which I will call the broad sense, is that *syntax* is the study (or description) of the form and arrangement of symbolic representations. Under this sense, the

[5] As Chomsky (1965) notes, however, the notion of simplicity is not quite empirical so any debate over whether T′ is simpler than T will be philosophical rather than scientific. I will take for granted that a theory of syntactic derivations without long-distance Agree is simpler than one with it, all else being equal.

study of syntax would be a part of the study of logic, programming languages, arithmetic, etc. Furthermore, Chomsky (2000b, p. 174), discussing this sense, argues that most of what we call *semantics* and *phonology* would be classified as syntax under this sense.[6] This sense will prove useful in this dissertation so I will retain it.

The third sense, which I will call the narrow sense, is that *syntax* is a mental module characterized by a computational procedure that generates an unbounded array of structured form[7]-meaning pairs. This, I believe, is what generative syntacticians mean when they use the term *syntax*. The "syntax module" is one of the objects of study of this thesis so I will retain this sense.

Since both the broad and narrow senses are useful to me, I will need to make a distinction for the sake of clarity. I will use the term "Narrow Syntax" (or NS) to refer to the narrow sense, that is, the hypothesized mental module, and "syntax" (and derived terms) to indicate the broad sense.

Similar remarks apply to the term *semantics*, which has at least three senses. The first sense, as in the case of *syntax*, is the sociological sense: semantics is what semanticists do. I will not be using this sense for the same reasons as I cited above for the sociological sense of *syntax*.

The broad sense of *semantics* is that of the study of the relation of a symbolic system to some other system. So, to various degrees, we can talk about the semantics of a logical system, a programming language, a natural language, etc. I will use this sense only informally when discussing notions of truth and reference associated with an instance or class of natural language expression.

The narrow sense of *semantics* is that of the mental module (or system of modules) associated with computing the meaning of a linguistic expression. Chomsky often refers to this mental entity as the Conceptual-Intentional (CI) system and stresses that we know very little about it. Insofar as this thesis makes claims or hypotheses about *semantics*, it makes claims or hypotheses about the CI system.

The final ambiguous term I will discuss is *interface*. In recent years it has become common within generative linguistics to write papers, hold workshops, and compile books on *the syntax-semantics interface*, but as I mentioned above, the term is ambiguous. It largely seems to be ambiguous between a *sociological* sense and a *narrow* sense, with the sociological sense dominating discussion.

When used in the sociological sense, the syntax-semantics interface refers to a body of literature that mixes the formalisms and methods used by syntacticians with those used by semanticists. That is, this type of work makes use of tree diagrams and expressions of

[6]Based on my discussion of this sense with phonologists and semanticists, this may be the most controversial claim Chomsky has ever made.

[7]I use the term *form* here to refer to all possible expressive modalities of language

typed lambda calculus. In this sense, the interface is not an object of study per se, but a sub-discipline.

In the narrow sense, the syntax-semantics interface refers to the interface between the Narrow Syntax and the CI system. Assuming this sense results in very different sorts of analyses compared to the standard analyses, analyses that posit computational procedures rather than merely representing expressions in two ways. In many ways, however, the term *interface* in its narrow sense is a misnomer, as there are likely no mental objects that we might call interfaces. As best we can tell, the mind consists of a set of modules and a non-modular central system (Fodor 1983, 2001). An interface, then, emerges wherever two modules interact with each other, or perhaps where a module interacts with the central system. Restricting ourselves to the modules, we can see why positing interfaces as mental objects won't do. Suppose we have two modules, M1 and M2, which seem to interact with each other. Being modules, each will consist in a set of computational operations (P1 and P2) defined over a class of syntactically structured (in the broad sense) objects (L1 and L2). Suppose we posit an interface I1, which consists in an operation P3 that converts objects of L1 into objects of L2. What is I1, then, but a module that has interfaces with M1 and M2? If I1 is a module, are its interfaces with M1 and M2 also modules? If so, then we seem to be stuck with an infinite regress. If not, then interfaces are a special kind of module, but this would raise further questions with respect to their evolutionary origins.

If there are no mental objects that we might call interfaces, then how are we to study them? The answer to this question is that to study an interface, we must study the modules associated with that interface with the added assumption that such an interface exists. So, studying the syntax-semantics interface involves studying the Narrow Syntax and the CI module with the assumption that there is an interface between them. We will get a glimpse of how such a study would work in chapter 5.

2.5 Summary

In this chapter, I have made explicit two of my assumptions which would be considered non-standard among contemporary generative syntacticians. In particular, I do not assume the θ-criterion as it is commonly stated, and I do assume a free Merge syntax. I have also clarified some terminology that many take for granted; specifically, I clarified my use of the term *syntax-semantics interface* and its constituent terms. Now that the reader has a sense of my theoretical assumptions, we can move on to more specific concerns in the following chapters.

Chapter 3

Previous Literature

There's always a siren
Singing you to shipwreck.
Steer away from these rocks.

"There, There"
Radiohead

In this chapter, I will review several previous analyses of adjectival resultatives and the parametric variation associated with them. I will evaluate the analyses against two desiderata. First, I will evaluate whether the variation, as analyzed, is learnable. Second, I will evaluate whether the analysis comports with the theoretical principles of the minimalist program. Before reviewing the analyses, however, I will make these desiderata explicit and justify them.

3.1 Desiderata for an analysis of resultatives

3.1.1 Desideratum 1: Learnability

Most analyses of the structure of adjectival resultatives include an account of the associated parametric variation, and it goes without saying that any aspect of a language that differs from grammar to grammar must be acquirable from the primary linguistic data. While few authors directly address the acquisition of parametric variation, any analysis of variation makes implicit claims about acquisition. Frequently, in discussions of parametric variation, the claims about acquisition are left implicit, as the acquisition task is assumed to be trivial. The nature of adjectival resultatives, however, is such that we must make those acquisition claims explicit. To explain why, I will compare the resultative parameter to the V-to-T parameter.

Analyses of the V-to-T parameter do not need to address learnability because the "parameter setting" is directly learnable from the primary linguistic data. It is directly learnable because constructions like polar questions differ overtly depending on the parameter setting. In a language with V-to-T movement such as German, the language learner will observe that lexical verbs undergo inversion for polar questions as shown in (3.1), while in a language without V-to-T movement such as English, the learner will observe that lexical verbs do not invert for questions.

(3.1) V-to-T movement (German)

a. Trinken sie Kaffee?
Drink.3plPres they Coffee
"Do they drink coffee?"

b. *Tun sie Kaffee trinken?
do they coffee drink

(3.2) *V-to-T Movement (English)

a. *Drink they coffee?

b. Do they drink coffee?

The form of polar questions, then, can be positive evidence for a particular setting of a parameter. Since we can find direct positive evidence for a parameter setting in data like (3.1)–(3.2), the task of the analyst/theoretician, then is merely to formalize the parameter is a way that is consistent with the broader theory. Chomsky (1995), for instance, formalizes the V-to-T parameter in terms of feature strength, while Lasnik (1999c) formalizes it in terms of the presence/absence of inflectional features on lexical verbs. Neither, however, needs to explicitly describe how their parameter is set. Rather, it can be assumed that a certain setting is the default, and the other setting can be deduced from (*e.g.*) the form of polar questions in the language.

Resultatives, on the other hand, are not directly learnable for two broad reasons. First, the task of discriminating between resultatives and depictives is far from trivial. And second, even if a learner is able to distinguish between resultatives and depictives, there is no principled and reliable way for a child to converge on a [-resultative] parameter setting. I discuss each of these arguments in turn below.

Distinguishing resultatives from depictives There are two reasons to say that resultatives are not, in principle, distinguishable from depictives. The first reason is that, on the surface, resultatives, which are parameterized, are indistinguishable from depictives, which appear to be universal. The two construction types are indistinguishable in the sense that both correspond to the string template in (3.3) (setting aside independent crosslinguistic word order variation).

(3.3) Subj V Obj Adj.

This indistinguishability is evident in the fact that one can construct examples that are truly ambiguous between resultative and depictive readings, as in (3.4).

(3.4) a. He fried the fish dry.
  (i) ≈ He fried the fish once it was dry. (**Depictive**)
  (ii) ≈ He fried the fish until it was dry. (**Resultative**)
 b. Grant Wood painted the house white.
  (i) ≈ The house is white in Grant Wood's painting. (**Depictive**)
  (ii) ≈ Grant Wood applied a coat of white paint to the house. (**Resultative**)

Assuming that a child acquiring either French or English encounters sentences with the form of (3.3) in their PLD, there is no obvious way for the child to determine whether a given secondary predicate is to be interpreted depictively or resultatively.

Some might object, arguing that the ambiguous examples above are contrived, and would easily by disambiguated in context. They might insist that the learner would infer a positive setting of the resultative parameter from the use of a secondary predication construction in the presence of a resultative event. So, an English learner, but not a French learner, might be exposed to the context-sentence pairing in (3.5).

(3.5) **Context:** A woman is methodically hammering a lump of metal. A parent draws their child's attention to the hammering event and utters:
A: She's hammering the metal flat.

Even this, however, is not fully unambiguous. Suppose the child has some notion of the link between the word *flat* and the property or state of flatness. In this case (3.5-b) certainly couldn't be interpreted as a depictive, but *flat* could plausibly be interpreted by a child as a manner adverb, modifying *hammering*. The child may have encountered the adverb *flatly* and also encountered adverbs that seem to have an optional *-ly* suffix (*e.g. quick*(*-ly*)), or cannot be suffixed with *-ly* (*e.g. fast*(**-ly*)). Such unavoidable ambiguity would make it difficult to reliably employ any sort of semantic bootstrapping in the acquisition of the resultative parameter.[1]

The second problem comes from the fact that both English-type and French-type languages can express resultative semantics periphrastically as in (3.6).

(3.6) **Periphrastic Resultatives**

 a. Elle a aplati le métal en le martelant
  She has flattened the metal in the hammering

[1]This line of argumentation is inspired by Carey (1985), Gleitman (1990) and Quine (1960)

b. She flattened the metal by hammering (it).

The resultative parameter, then, is a one-or-both parameter, meaning the presence of periphrastic resultatives in the PLD cannot constitute conclusive evidence of either setting of the resultative parameter. The V-to-T parameter, in contrast, is an either/or parameter, meaning, for example, that the presence of *do*-support can be taken as evidence that V does not move.

Setting the parameter There is, in fact, a stronger argument against the direct learnability of the resultative parameter, which I will make in this section. The basic structure of the argument is along the lines of Fodor's (1975) argument against the possibility of concept learning. To begin with, let's assume that a child has all the ingredients for resultatives, whether or not she is acquiring a language with resultatives. This either means the lexical items required for an adjectival resultative are innate or that they are acquired prior to the possible acquisition of adjectival resultatives. Also, let's assume that a child has acquired resultatives if and only if her grammar is able to generate them and further that she can mentally represent a resultative iff her grammar is able to generate them.

Now, consider the classic "switchboard" model of parameter setting where each parameter comes with a default setting which is only altered in the face of PLD which is incompatible with the default. So, if we assume that there is a resultative parameter, we must then determine what its default setting is. Either it is set to generate resultatives by default, or it is set to not generate resultatives by default. I will call these possibilities English-default and French-default, respectively. As we shall see, the English-default option is faulty on its face. The French-default option, on the other hand, seems to fare better, but I will argue that it faces the exact same problems as the English-default option. Since these two options exhaust the logical possibilities of parameter defaults, I will conclude that the acquisition of resultatives in this manner is impossible.

First, consider the English-default option. According to this option, when a child encounters a secondary predication construction, she will entertain two possible parses: a depictive parse and a resultative parse, and choose the one whose interpretation is appropriate for the context. If the child is acquiring English, then the context will sometimes prefer the resultative parse, and other times prefer the depictive parse. If the child is acquiring French, then the context will always prefer the depictive parse. However, in order to change the default setting, the child must make an inductive leap from a series of particular utterances to a generalization across all possible utterances. Specifically, she will have to infer from the absence of data with a preferred resultative parse, that the resultative parse is categorically unavailable. Such an inductive step is problematic, to say the least. In fact, it is generally assumed in generative grammar that language acquisition cannot be an inductive process. (Chomsky

and Fodor 1980; Fodor 1980) Therefore, the English-default option is to be rejected.

Suppose instead that we choose the French-default option. According to this option, when a child encounters a secondary predication construction, she will only entertain a depictive parse by default. If the child is acquiring French, then there will be no issue: Every parse will be appropriate for the context. If the child is acquiring English, then a subset of the parses will be inappropriate for the context. These inappropriate parses, then, would be taken as positive evidence for the alteration of the default setting. This means, however, that the French-acquiring child is able to entertain resultative parses, even though they are dispreferred by default. If the French-acquiring child is able to entertain resultative parses, then she is able to mentally represent resultatives, and if she is able to mentally represent them, then she is able to generate them.

3.1.2 Desideratum 2: Theoretical consistency

The second quality that an analysis of resultatives should have is consistency with the broader theory of grammar. While this is a requirement of all analyses of grammatical phenomena, I will outline the subset of minimalist hypotheses that I will use to evaluate existing analyses of adjectival resultatives. First, I will review what is called the Uniformity of Theta Hypothesis (UTAH) which gives us a baseline for our theory of θ-role assignment. Second, I will look at the Lexical Parameterization Hypothesis, the minimalist hypothesis which will guide our theory of variation. Each of these will be discussed with reference to a particular property of adjectival resultatives.

3.1.2.1 The Uniformity of Theta Assignment Hypothesis

The canonical formulation of UTAH is that of Baker (1988) working in a pre-minimalist Principles and Parameters framework, given below in (3.7).

(3.7) **The Uniformity of Theta Assignment Hypothesis (UTAH)**
Identical thematic relationships between items are represented by identical structural relationships between those items at the level of D-structure. (Baker 1988, p. 46)

While I will be assuming something of this sort in my discussion, the reference to D-structure, which was eliminated in the minimalist program, renders Baker's hypothesis unusable in its original form. I propose the following more precise formulation of UTAH:

(3.8) **Minimal UTAH**
If a thematic relation f holds between items X and Y in distinct expressions S_1 and S_2, then there is a single structural relation g that also holds between X and Y in both S_1 and S_2.

Before considering concrete cases, I would like to discuss a few points about my conception of UTAH. First, I do not make any commitment to the existence of thematic relations, θ-roles, θ-features, or the like as theoretical primitives, rather as descriptions of phenomena. There may be notions of AGENT or THEME in the mind, or in the grammar, or in the world, or maybe not, I believe Minimal UTAH will be useful whatever the case may be. As such, I will not be particularly interested in distinguishing between θ-roles (*e.g.*, whether a given argument is a THEME or a PATIENT) unless such a distinction has consequences for the grammar.[2]

The first point leads to the second point, which is that Minimal UTAH, being an empirical hypothesis, is both a theoretical claim and a methodological heuristic. As a point of comparison, consider a hypothesis at the core of generative syntax: constituency. The constituency hypothesis claims that within a sufficiently complex linguistic expression (longer than 2 words, say) there are sub-parts of that expression which form linguistic expressions, and there are sub-parts that do not. So, for instance, in (3.9), *the roof* is a linguistic expression, but **hit the* is not.

(3.9) hit the roof

This leads to a set of methodological heuristics often called *constituency tests*, which are, no doubt, familiar to anyone who has made it this far into this dissertation.

The third point, which also follows from the earlier points, is that the methodological heuristic, the diagnostic, that follows from Minimal UTAH consists in analyticity, that is, entailments that hold by virtue of linguistic facts rather than facts about the world.[3]

To understand what this means, consider the following sentence pairs:

(3.10) a. Sara broke the bottle.
b. The bottle broke.

(3.11) a. Katie ran.
b. Katie ran a kilometre.

The pair in 3.10 represents a single thematic relation between *broke* and *the bottle*. Minimal UTAH says that there must be a single structural relation that holds between *broke* and *the bottle* in both sentences. Similarly, in 3.11, there is a single thematic relation between *Katie* and *ran*, which means there must be a single corresponding structural relation between the

[2]For instance, it has been shown that AGENT subjects have different grammatical properties than EXPERIENCER subjects do.

[3]These entailments contrast with *synthetic* inferences, which are due to extra-linguistic (and extra-mental) facts. For instance, suppose Hyppolite lived at the mouth of the Credit River in 2017, from this we can infer that Hyppolite lived in Mississauga as the mouth of the Credit River is in Mississauga. The same inference does not hold if Hyppolite lived at the mouth of the Credit River in 1900, as the borders of Mississauga did not include the mouth of the Credit then.

two in both (3.11-a) and (3.11-b). Any structural analysis of these sentences that violates Minimal UTAH, then, can be rejected on theoretical grounds.

Consider, for instance, three possible structural analyses of the pair in 3.10. In the first analysis, represented in Figure 3.1, *the bottle* is base generated in its surface position, which is different in the two sentences. In this analysis, there is no single structural relation be-

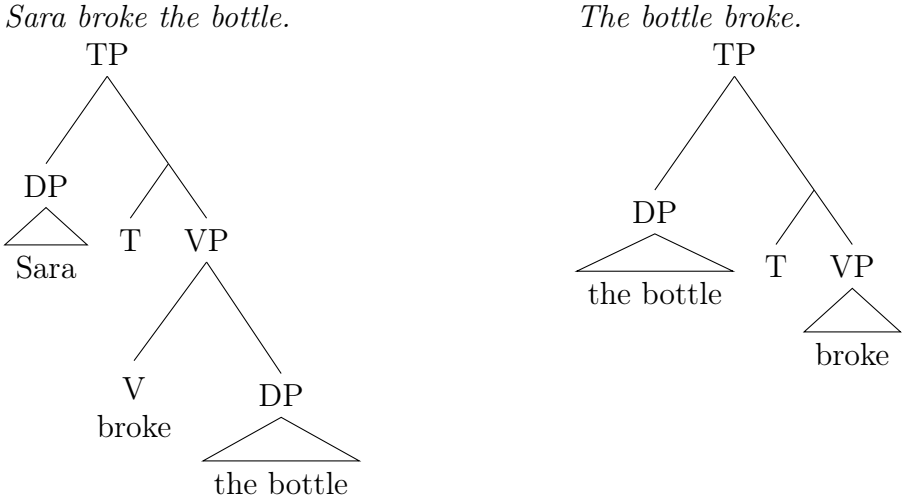


Figure 3.1: A surface analysis of causative-inchoatives

tween *broke* and *the bottle*, corresponding to the single thematic relation. This analysis, then, violates Minimal UTAH and can, therefore, be rejected.

In the second analysis, represented in Figure 3.2, *the bottle* originates in [Comp V] and moves to subject position in the intransitive situation. In this analysis, the single thematic

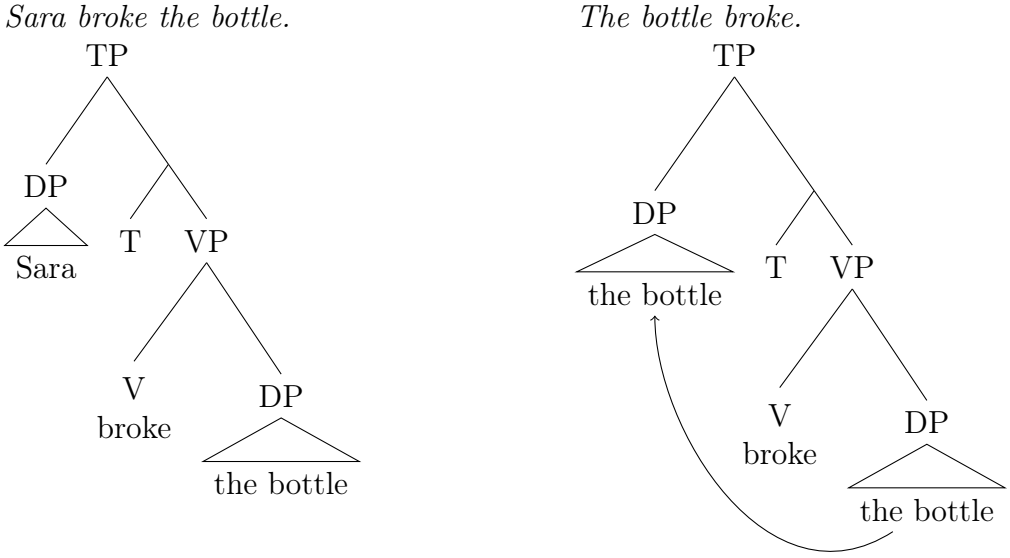


Figure 3.2: A movement analysis of causative-inchoatives

relation between *broke* and *the bottle* corresponds to a single structural relation between the

two. While UTAH considerations clearly lead us to prefer the movement analysis over the surface analysis, such considerations are far from decisive evidence in favour of the movement analysis, as it is not the only possible analysis that complies with UTAH.

Another UTAH-compliant account, which I will call the layering analysis, is represented in Figure 3.3. This type of analysis is proposed by Borer (2005b) and Ramchand (2008).

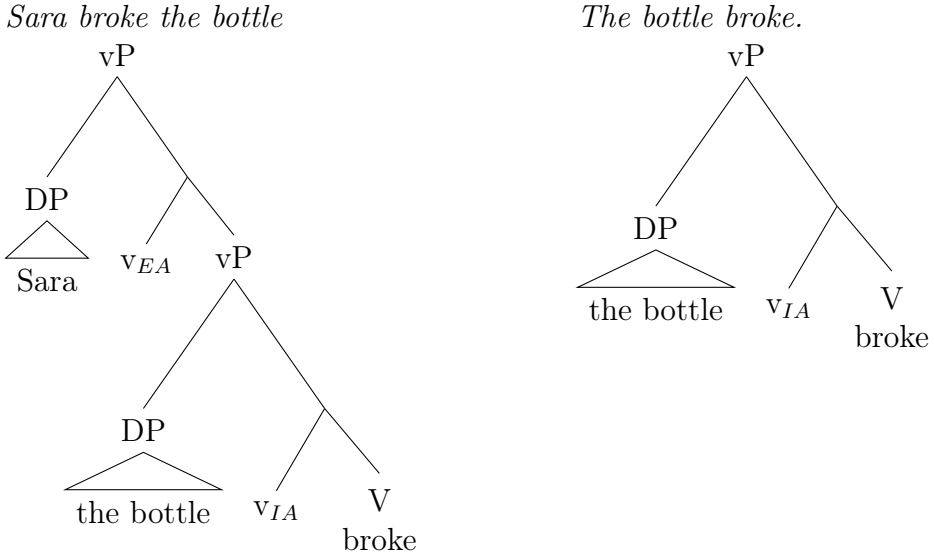


Figure 3.3: A layering analysis of causative-inchoatives

In this analysis, as in the movement analysis, the thematic relation between *broke* and *the bottle* is represented by a single structural relation between the two. This means that this analysis also satisfies Minimal UTAH and, therefore, cannot be rejected immediately.

UTAH is relevant to adjectival resultatives considering the pattern of thematic relations in (3.12).

(3.12) a. Jackie hammered the metal flat.
b. Jackie hammered the metal.
c. Jackie made the metal flat.

There is a single θ-relation that holds between *hammered* and *the metal* in both (3.12-a) and (3.12-b), and there is a single θ-relation that holds between *the metal* and *flat* in both (3.12-a) and (3.12-c). Therefore, there should be a single structural relation corresponding to each of those θ-relations, and any analysis that does not meet this criterion will be rejected.

It could be noted that while a particular entailment pattern clearly and robustly holds of (3.12) ((3.12-a) entails both (3.12-b) and (3.12-c)), the same cannot be said for all versions of adjectival resultatives. Consider, for instance, (3.13) the English version of Kratzer's (2005) central example.

(3.13) a. We drank the teapot dry.
(*cf. Wir haben die Teekanne leer getrunken*)
b.?# We drank the teapot.
c. We made the teapot dry.

While (3.13-a) does clearly and robustly entail (3.13-c), we cannot say the same about (3.13-a) and (3.13-b), as it is generally judged as highly coerced. Presented with this pattern, we might argue that UTAH considerations are only informative regarding the structural relationship between the object and the adjective of resultatives like (3.13-a). The pattern in (3.13) does not, however, negate the pattern in (3.12), which suggested that the resultative object is θ-marked by its verb as well as its adjective. So how are we to deal with this inconsistent data? At this point, there are two lines of reasoning we could pursue: assume that (3.12) and (3.13) represent two distinct resultative structures, or assume that there is only one structure for resultatives and the inconsistency in the data comes from some other difference between the two sentences.

Following the first line of reasoning, there are at least two distinct structures for resultatives: one corresponding to cases like (3.12), and another corresponding to cases like (3.13). To my knowledge, however, there is no corroborating evidence for any structural difference between (3.12-a) and (3.13-a), and, absent any such evidence, we ought to assume they have the same structure.

If we take this second line of reasoning, we must then decide which type of resultative we take as "prototypical" and which one needs extra explanation. If (3.13-a) is our prototype, then the resultative object is not prototypically θ-marked by its verb and we would have to explain why the object of (3.12-a) exceptionally seems to be θ-marked by its verb. There are, in my mind, two possible ways to explain this: we could either retain a single structure for resultatives and argue that *the metal* is really not θ-marked by *hammer* in *hammer the metal flat*, or we could propose multiple structures for resultatives depending on their θ-marking properties.

The first line of argument means that we would need to propose at least one situation in which a thematic relation is established without explicit θ-marking, that is, extragrammatically. This means we would be proposing an exception to UTAH, meaning we would need to explain why resultatives are exceptional. If resultatives are not truly exceptional, then they are the norm and UTAH-compliant cases are the exceptions, meaning we would have to reject UTAH wholesale. Neither of these is a theoretically attractive option, so I will not pursue them any further here.

The second line of argument – that there are multiple structures associated with resultatives – violates the scientific principle of theoretical parsimony. This means that, while we can't reject it outright, we should hold it as a last resort.

So, assuming Kratzer's example to be our prototype leads us to undesirable results. Assuming *hammer the metal flat* to be the prototype, I argue, does not lead us to any undesirable conclusions, though it does present some puzzles. As in the discussion above, assuming one type of resultative to be prototypical means explaining the eccentricities in the non-prototypical version; that is, we must explain the apparent lack of θ-marking in Kratzer's example. While I cannot offer a full explanation, I propose that such an explanation will depend on a fuller understanding of the complex dynamics of coercion. Developing such an understanding is beyond the scope of this thesis, but see Pustejovsky (1998) for some promising proposals in this domain.

3.1.2.2 The Lexical Parameterization Hypothesis

In addition to UTAH considerations, I will evaluate whether the theory of parametric variation attached to a given analysis is formulable in the minimalist model of grammar. This model of the grammar is composed of four parts: a lexicon containing the atoms from which complex expressions are built; Merge, the simplest combinatorial operation for constructing complex expressions; and two interfaces, points at which a constructed expression is transferred to language-external modules for either externalization (the sensorimotor (SM) interface), or interpretation (the conceptual-intentional (CI) interface). In this section I will argue for what is known as the Lexical Parameterization Hypothesis (LPH), which was first formulated by Hagit Borer in the following way:

> In this study, we will propose a model of parameters which restricts the availability of variation to the possibilities which are offered by one single component: the inflectional component. (Borer 1984, p. 3)

Manzini and Wexler (1987) however reformulate this as the hypothesis that "[v]alues of a parameter are associated not with particular grammars but with particular lexical items." (Manzini and Wexler 1987, p. 424). The version of LPH that I assume is given in (3.14).

(3.14) **The Lexical Parameterization Hypothesis**
The lexicon is the only component of grammar that can be the locus of parametric syntactic variation.

In the remainder of the section, I will argue that LPH is a reasonable hypothesis. I will do so first by a positive argument – that the lexicon is a natural place to locate variation – then by a negative argument – that neither Merge, nor the interfaces can be parameterized.

The notion that lexicons vary may seem trivial, but it is worth describing if only to distinguish grammatically relevant variation from grammatically irrelevant variation. This dichotomy roughly seems to follow the distinction between functional elements and more meaningful elements (*i.e.*, roots). Variation between two speakers with respect to vocabulary

(in the non-technical sense) can be fairly profound without necessarily meaning that the speakers are speaking a different language. The variation relevant to this thesis, however, is the kind that leads to systemic differences between grammars. For instance, consider the link between verbal agreement and pro-drop. Synthesizing a broad set of empirical studies, Huang (1984) observes that only those languages with rich agreement morphology (*e.g.*, Italian and Spanish) or no agreement morphology (*e.g.*, Mandarin and Japanese) allow subject pro-drop. If we take the richness of agreement to be a lexical parameter,[4] then we have evidence that lexical variation can lead to systemic variation between grammars, and it is reasonable to consider agreement to be a lexical property.

The English verbal system provides us with good reason to consider richness of agreement to be lexical. In English, main verbs and modals differ with respect to their agreement, with main verbs showing impoverished agreement and modals showing no agreement. Regardless of the details of the analysis of this variation, there must be some lexical property that determines whether or not a verbal element surfaces with agreement morphology. Since agreement morphology can be a variable lexical property within a language, it is reasonable to assume that it can be a variable lexical property across languages. Following this line of reasoning, and assuming that the pro-drop parameter is correlated with agreement morphology, the pro-drop parameter is reducible to a lexical parameter.

Having shown that there can be significant grammatical effects from lexical variation, I will now argue that the remaining components of the grammar (Merge, the two interfaces) are ill-suited to variation.

Merge is the easiest component to eliminate from consideration as the locus of variation due to its simplicity. As defined in (3.15), it is the simplest possible operation for generating hierarchical structures: binary set formation.

(3.15) $\text{Merge}(\alpha, \beta) = \{\alpha, \beta\}$

In order to attribute cross-linguistic variation to Merge, we would need to introduce complexity to it, and it would no longer be maximally simple.

It is worth noting, however, that Merge, as defined in (3.15), is not *the* simplest possible combinatory operation, but the simplest one capable of generating hierarchical structures. Hornstein (2009), for instance, proposes a concatenative version of Merge.

[4]This is by no means the only way to analyze richness of agreement. In fact, a more intuitive analysis would be that English and Italian finite T's are identical with respect to φ-features in the lexicon and narrow syntax, and richness of agreement is a wholly morphological fact. Each of these analyses has its strengths and weaknesses. The morphological analysis provides a straightforward way to represent the differences in agreement paradigms but does not provide a natural explanation for the apparent grammatical effects of the variation in question. The lexical analysis, on the other hand, provides an explanation for the link between agreement and pro-drop (see Chomsky 2015 for such an explanation), but cannot naturally represent the variation in richness of agreement. Since the grammatical effects of variation are of interest to me, I will assume the lexical analysis.

(3.16) $\text{Merge}'(\alpha, \beta) = \alpha^\frown\beta$

Assuming that concatenation is as simple as set-formation, it is still not able to form hierarchical structures without further complication. This is because concatenation, unlike binary set-formation, is associative. That is, it obeys the law in (3.17) as demonstrated in (3.18).

(3.17) **Associative Law**
$(x \centerdot y) \centerdot z = x \centerdot (y \centerdot z)$ for all x, y, z

(3.18) **Concatenative Merge is associative**
$\text{Merge}'(\ \text{Merge}'(\alpha, \beta), \gamma)$
$= (\alpha^\frown\beta)^\frown\gamma$
$= \alpha^\frown\beta^\frown\gamma$
$= \alpha^\frown(\beta^\frown\gamma)$
$=\text{Merge}'(\alpha,\ \text{Merge}'(\beta, \gamma))$

Since concatenative Merge forms essentially flat structures, it is ill-suited to construct linguistic expressions which appear to be fundamentally hierarchical. So, while there are other possible combinatory operations as simple as Merge, there don't seem to be any that can capture the fact that linguistic expressions are hierarchically structured. It, therefore, appears that any variation in Merge can only come from complicating the operation.

Being a hypothesis, the LPH need not be derived deductively. Rather, it can be justified by showing that it is a sufficiently strong, non-falsified hypothesis. To define the strength of a hypothesis, I will adopt Popper's (1959; 2014) notion of *empirical content*. For Popper, the empirical content of a theory, or statement a is the "the class of all basic statements which contradict a" (Popper 2014, p. 315), where a *basic statement* is a statement whose truth-value can be determined by mere observation of the world. A hypothesis H_1 is stronger than another hypothesis H_2 if the empirical content of H_1 is greater than that of H_2. While there is no way to determine the absolute empirical content of a given statement and, therefore, there is no way to directly compare the content of two independent statements, we can compare logically related statements. Popper (2014, p. 295), for instance, compares the content of two logically independent statements, a and b, with that of the conjunction of the two $a \wedge b$, arguing that the empirical content of the conjunct is at least as great as that of each of the independent statements.

(3.19) $Ct(a) \leq Ct(a \wedge b) \geq Ct(b)$

A hypothesis is *sufficiently* strong if there are no hypotheses in the relevant domain that are demonstrably stronger.

So, is LPH a sufficiently strong hypothesis? To answer this we must consider what the

competing hypotheses are. The LPH belongs to a class of single-source hypotheses, those that hypothesize a single source of variation which I give below in (3.20).

(3.20) If x is a variable property of language...

a. then x is a property of the lexicon. (= LPH)
b. then x is a property of the SM interface.
c. then x is a property of the CI interface.
d. then x is a property of the computational system.

We could formulate a class of multiple source hypotheses which would be the class of disjunctions of two or more of the single-source hypotheses as in (3.21).

(3.21) If x is a variable property of language...

a. then x is a property of the lexicon or x is a property of the SM interface.
b. then x is a property of the lexicon or x is a property of the CI interface.
c. then x is a property of the lexicon or x is a property of the computational system.
d. then x is a property of the CI interface or x is a property of the SM interface.
e. *etc.*

Unlike conjunctions, whose empirical content equals or exceeds that of their component statements, the empirical content of disjunctions is less than or equal to that of their component statements.

(3.22) $Ct(a) \geq Ct(a \vee b) \leq Ct(b)$

Since the multiple source hypotheses have less empirical content the single-source hypotheses, they cannot be considered sufficiently strong.

There are hypotheses which might be considered stronger than any of the single-source hypotheses, but these hypotheses face problems. For instance, a stronger hypothesis would be that there are no variable properties of language, but this hypothesis is quite easily falsified. A hypothesis constructed by conjoining two or more of the single-source hypotheses would, on its face, seem like a stronger hypothesis, but in fact, such a hypothesis would be incoherent. A conjoined source hypothesis would be incoherent because the very notion of modularity that our theory assumes requires that modules are independent of each other. If two modules are independent of each other, then they can only share a property under two conditions: either they share the property by accident, or the property is a property of a modular system that contains them both. The former condition is merely a special case of the multiple source hypothesis, while the latter case is irrelevant here, though perhaps relevant to cognitive science more broadly. Therefore, I will set aside the conjoined source

hypothesis.

So, since there are no admissible hypotheses that are stronger than our single-source hypotheses, then we can say that the single-source hypotheses are sufficiently strong. Furthermore, the presence of lexical variation (*e.g.*, variable richness of subject agreement) seems to falsify all of the single-source hypotheses except for LPH, meaning that LPH is the only sufficiently strong non-falsified hypothesis.

Before continuing, I should reiterate that the discussion above is not the deduction of the truth of the LPH but rather of its relative strength. To the contrary, as Popper (2014) notes repeatedly, the strength of a theoretical claim is inversely proportional to the probability of its truth. This means that the LPH is very likely to be false, which means that if it is false, it should be relatively easy to refute it. The strength of a hypothesis, such as the LPH or its competitors, then, is of methodological relevance, while the truth, or verisimilitude, of that claim is ultimately an empirical question.

The two theoretical principles discussed above will allow me to evaluate the previously proposed analyses of resultatives. UTAH considerations will determine if a given structural analysis is tenable in minimalist syntax. The LPH, on the other hand, will be used to evaluate whether a given analysis encodes variation in an acquirable way.

3.2 Previous analyses of adjectival resultatives

In this section, I will address three previous analyses of adjectival resultatives. The first, the complex predicate analysis (Irimia 2012; Snyder 1995), I argue, violates both UTAH and the LPH and is, therefore, not theoretically viable. The second, the cartographic/nanosyntactic analysis (Son and Svenonius 2008), is theoretically sound but arguably unlearnable. The third, that of Kratzer (2005), which I will adapt for my purposes, narrowly violates UTAH.

Complex-predicate analyses of secondary predication structures tend to be defined in opposition to small clause analyses. The two classes of analysis crucially differ with respect to where the shared argument originates. According to a small clause analysis, the shared argument forms a constituent with the secondary predicate at some point in the derivation as in figure 3.4. Kratzer's (2005) analysis of resultatives, discussed below, is a small clause analysis.

According to a complex predicate analysis, however, the shared argument and the secondary predicate never form a constituent that does not include the main verb. Instead, the verb and the secondary predicate (and perhaps additional syntactic objects) form a constituent that does not contain the shared argument. The shared argument (in the case of resultatives, at least) then forms a constituent with this complex predicate. For instance, Irimia (2012) analyzes resultatives as in figure 3.5. This analysis is inconsistent with any strict version of UTAH, as the structural relation between *the metal* and *hammer* in figure 3.5 is

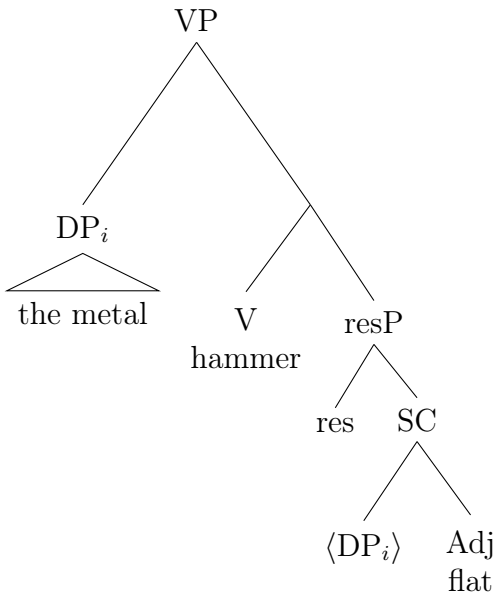


Figure 3.4: A small clause analysis of resultatives

not shared with the simple transitive version as shown in figure 3.6. In the simple transitive version, *the metal* and *hammer* are sisters but not in the resultative version. The complex predicate analysis, then, violates UTAH and should be rejected under our assumptions. Before rejecting them outright, however, I will first consider the strongest argument in favour of complex predicate analyses and an attempt to reconcile complex predicate analyses with UTAH.

Perhaps the strongest argument in favour of the complex predicate analysis comes from the anti-reconstruction effects identified by Williams (1983, pp. 293–296). Reconstruction is the phenomenon of arguments being interpreted lower than their surface position. It is usually diagnosed using scope ambiguities as seen in (3.23).

(3.23) A student seems to be sick.

a. Surface reading: $\exists >$ seem
   $\approx$ There is a particular student who seems to be sick.

b. Reconstructed reading: seem $> \exists$
   $\approx$ It seems to be the case that at least one student is sick.

Generally, reconstruction effects are taken to be evidence for syntactic movement, because movement is a natural explanation for them. As Williams points out, many constructions traditionally analyzed as small clauses do not show reconstruction effects. For instance, (3.24), is unambiguous.

(3.24) A student seems sick.

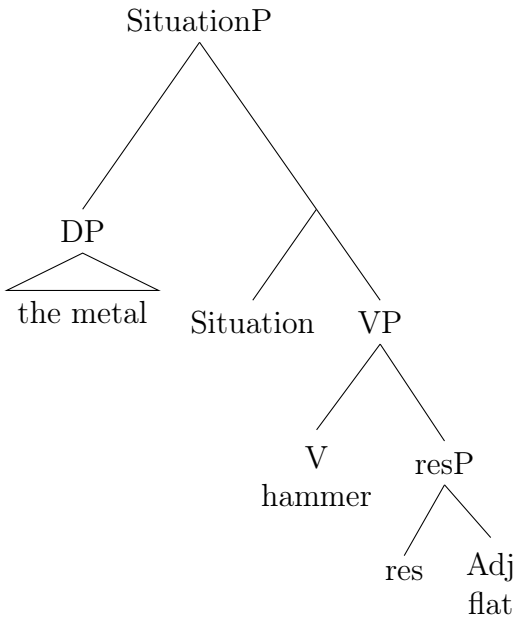


Figure 3.5: Irimia's analysis of resultatives

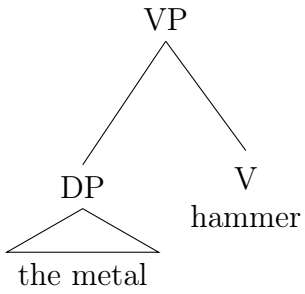


Figure 3.6: A simple transitive VP

a. Surface reading: $\exists >$ seem
   $\approx$ There is a particular student who seems sick.
b. *Reconstructed reading: seem $> \exists$
   $\neq$ It seems to be the case that at least one student is sick.

If there were movement from a small clause in (3.24), the argument goes, we would expect reconstruction effects. Therefore, (3.24) cannot be derived from a small clause structure, and since, according to Irimia (2012), these anti-reconstruction effects are a crucial property of secondary predication structures (including resultatives), we cannot derive these structures from small clauses.

At first blush, this seems to be a solid line of reasoning, and I am not in a position to dispute the evidence adduced in its service, but it hinges on a hidden hypothesis. The basis of this argument is the one-way conditional statement that reconstruction implies movement. The hypothesis that Williams and others put forth is that this conditional is, in fact, a

biconditional. That is, an argument reconstructs if and only if its surface position is derived from movement, and it follows from this that anti-reconstruction effects imply a lack of movement. This, I believe, is too strong, as there are instances of anti-reconstruction effects that are not amenable to analysis as complex predicates.

The strongest evidence against this hypothesis is the fact that quantifier float blocks reconstruction.[5] So (3.25), with an unfloated *all*, shows reconstruction effects while the same sentence with a floated *all* in (3.26), does not show these effects.

(3.25) **All** the students seem to be sick. ($\forall >$ seem, seem $> \forall$)

(3.26) The students **all** seem to be sick. ($\forall >$ seem, *seem $> \forall$)

The hypothesized biconditional would predict that *seem to be sick* is a complex predicate in (3.26) and an ordinary non-finite clause in (3.25). So, in order to maintain the hypothesis, it is incumbent on the complex predicate theorist to explain how and why quantifier float and complex predication interact.

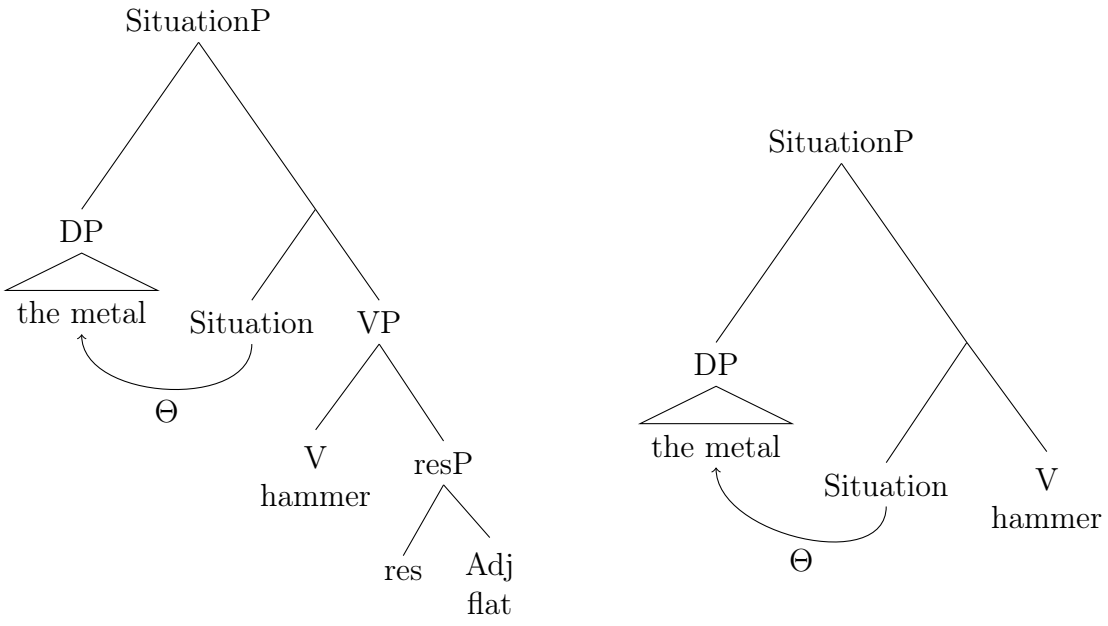
Rather than reject the biconditional outright, however, I will merely weaken it to say that all else being equal, reconstruction effects are a good indication of argument movement. In doing so, I have taken on the burden of explaining why shared arguments of secondary predication structures do not reconstruct into their small clause positions. Such an explanation is beyond the scope of this thesis, though. Therefore, I leave this burden for future research.

As for the approach to θ-roles, Irimia (2012) proposes that the shared argument receives its θ-role from the Situation head which takes the complex predicate as a complement. From a purely formal perspective, this can be made to satisfy Minimal UTAH in the following way. If we assume that resultatives have the structure in figure 3.7a, and simple transitives have the structure in figure 3.7b, then we can say that the complex predicate does not violate Minimal UTAH as the thematic relation that holds between Situation and the DP is mirrored by a grammatical relation between the two in both structures. In both structures, *the metal* receives a θ-role from Situation, and it is the specifier of Situation, so Minimal UTAH holds. This, however, depends on a different conception of θ-relations than the one I assume. Recall that I take thematic relations to be evident in certain types of analytic entailments such as the one expressed in (3.27).

(3.27) Deem drank the cup dry. $\implies$ The cup was dry.

In order to get from the structures in figure 3.7 to the entailment in (3.27), we would need to assume some mechanism by which the θ-role of *dry* is transferred to Situation. Such a mechanism, while perhaps formulable, should be immediately suspect under minimalist

[5] See also Bobaljik and Wurmbrand (2004) for another argument against this hypothesis.



(a) *hammer the metal flat* (b) *hammer the metal*

Figure 3.7: Θ-marking according to Irimia (2012)

assumptions. So, I will set aside the complex predicate analysis of resultatives.

I take Snyder's (1995; 2001; 2012) work on resultatives to be a part of the complex predicate milieu, although a precise structural analysis of resultatives seems to be of secondary interest for him. His primary interest is providing an account of the acquisition and grammatical representation of the parametric variation associated with resultatives. I discuss his account below and argue that, while his analysis of the acquisition of resultatives is convincing, his proposed grammatical basis for the parameter violates the LPH and will, therefore, be set aside.

Snyder (1995, 2001) observes a strong correlation between productive N-N compounding and certain classes of complex predication structures, of which resultatives are one. He demonstrates the correlation using both a cross-linguistic study and an acquisition study. The cross-linguistic study, whose results are summarized in table 3.1, shows that only those languages that allow productive N-N compounding allow adjectival resultatives. The acquisition study bolsters this claim by showing that, for ten English-acquiring children in the CHILDES database, the first recorded utterances of N-N compounds correlates with the first recorded utterances of particle verb constructions.

In later work, Snyder (2016) clarifies the type of compounding relevant to resultatives, arguing that *Bare Stem Compounding* (BSC), rather than N-N compounding, is correlated with resultatives. As the term suggests BSC is the productive process of forming compounds

[6]Snyder 2001 lists Basque as a language that allows N-N compounding disallows resultatives. In later work, however, Snyder (2012) argues that what looks like compounding in Basque actually involves something like a construct state. Therefore, Basque seems to disallow N-N compounding.

	Resultatives	Productive N-N Compounding
American Sign Language	yes	yes
Austroasiatic (Khmer)	yes	yes
Finno-Ugric (Hungarian)	yes	yes
Germanic (English, German)	yes	yes
Japanese-Korean (Japanese, Korean)	yes	yes
Sino-Tibetan (Mandarin)	yes	yes
Tai (Thai)	yes	yes
Basque	no	no[6]
Afroasiatic (Egyptian Arabic, Hebrew)	no	no
Austronesian (Javanese)	no	no
Bantu (Lingala)	no	no
Romance (French, Spanish)	no	no
Slavic (Russian, Serbo-Croatian)	no	no

Table 3.1: Results of Snyder's survey (Snyder 2001, p. 329)

out of bare stems, which Snyder (2016) defines as in (3.28).

(3.28) A bare stem is "any form that (i) could be used as an independent word (or at least, could be so used after the addition of inflectional morphology), and (ii) is the form that inflectional morphology would combine with, but (iii) does not yet bear any inflection." (Snyder 2016, p. 90)

This clarification allows Snyder to more naturally distinguish between the type of compounding we see in English, in which compounds take the form of two juxtaposed bare stems, and the type found in Hebrew, where compounds take the form of one bare stem juxtaposed with a so-called *construct state* stem.

(3.29) a. worm can

b. kufsat tulaAim (Hebrew construct state)
can-of worm
"worm can" (Snyder 2001, p. 338)

If we accept that there is indeed a correlation between the availability of productive BSC and productive resultatives, (and I know of no convincing evidence against it) then it is very reasonable to hypothesize that the experience of one of the correlates in their PLD serves as a triggering experience for a child's acquisition of the other. Since I assume that resultatives are not directly acquirable, the hypothesis must be that compounding triggers resultatives. This is a reasonable hypothesis since there are clear surface signs of compounding. For instance, English nominal compounds are distinguishable from modified NPs by a prosodic pattern

demonstrated in (3.30).

(3.30) a. blue bírd (modified NP)
b. blúe bird (compound N)

The question this hypothesis leads to, and which occupies this thesis, is what it means for a grammar to allow BSC, and how it is that BSC is linked to resultatives. Snyder proposes an answer to this question which I discuss, and ultimately reject, below.

Snyder's proposal is that a language allows resultatives and BSC only if it permits Generalized Modification (GM), a compositional operation defined in (3.31).

(3.31) Generalized Modification (Snyder 2012)
If α and β are syntactic sisters under the node γ, where α is the head of γ, and if α denotes a kind, then interpret γ semantically as a subtype of α's kind that stands in a pragmatically suitable relation to the denotation of β.

Since GM is a compositional principle, that is to say, a function from syntactic objects to meanings, it would be part of the CI interface. Therefore, if this operation were parameterized, the parameter would apply, not to the lexicon but to the CI interface. Such a parameter would violate the LPH and must, therefore, be set aside as a possible account of the resultative parameter given our assumptions.

The fact that his parameter violates the LPH is not lost on Snyder. In fact, he argues that the connection between compounding and resultatives is evidence against the LPH, writing that "[i]t is difficult to imagine any single, independently motivated functional head will find a natural role both in complex predicates and in a morphological compound such as *coffee cup*." (Snyder 1995, p. 62) It is indeed difficult to argue against this quoted assertion as such. What I will do instead is argue against the implicit assertion that under the LPH, a single functional head must necessarily be involved in both resultatives and compound formation. I will do this, by presenting a plausible account of Snyder's correlation that is LPH-compliant.

Another structural and parametric account of resultatives, by Son and Svenonius (2008), follows Ramchand's (2008) decompositional analysis of vP syntax. Under this account, event descriptions are decomposable into several universal features (*e.g.*, Result, Manner, Process), and language variation results from languages lexicalizing these features differently. Son and Svenonius (2008) propose that resultative languages, such as English and German, lexicalize Result in null heads, while non-resultative Romance languages only lexicalize Result in verbs.

(3.32)

Spanish	Proc Res Pred		A
	hacer		*peligroso*
	‘make’		‘dangerous’
English	Proc	Res Pred	A
	pound	∅	*flat*

Note that this account is consistent with the LPH because the variation is, indeed lexical. It is not immediately clear, however, that this lexical parameter could be learned from the primary linguistic data. Since the head in question is phonologically null, there is, by definition, no direct evidence for it, for example, in English. Therefore, the null head would have to be acquired indirectly, but Son and Svenonius (2008) provide no proposals for how that would be done. In the absence of any proposed method for acquiring this parameter, I will set it aside. This means I will also set aside the structural analysis that is associated with the parameter.

The final analysis I will mention is that of Kratzer (2005), which I will adopt with some modifications. In this analysis, the resultative object (*e.g.*, *the metal*) and the adjective (*e.g.*, *flat*) form a small clause, which describes a state. The small clause merges with a *res* head, which encodes a causative relation between events, and the resulting *res*P is merged as the complement of the verb (*e.g.*, *hammer*). The small clause theme is then raised to check accusative Case, and from there the derivation proceeds as normal. The structure this generates is given in figure 3.8. Under the assumption that *the metal* is θ-marked by *hammer*,

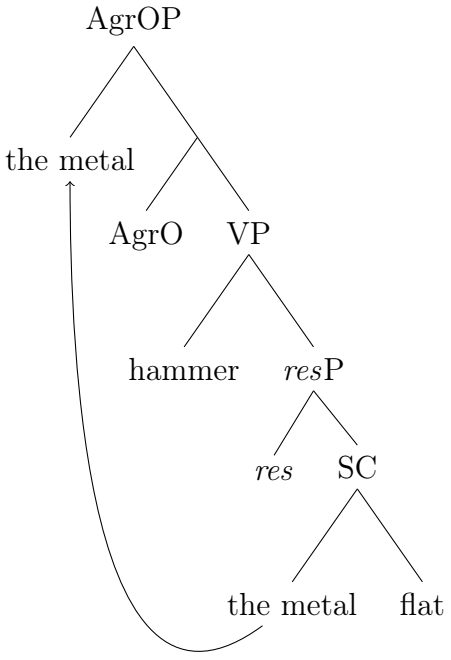


Figure 3.8: Kratzer’s Resultative Structures

this violates UTAH. This violation can be seen when we compare the resultative sentence that is partially represented in figure 3.8 with the sentence in (3.33) and its structure, represented in figure 3.9.

(3.33) The metal was hammered.

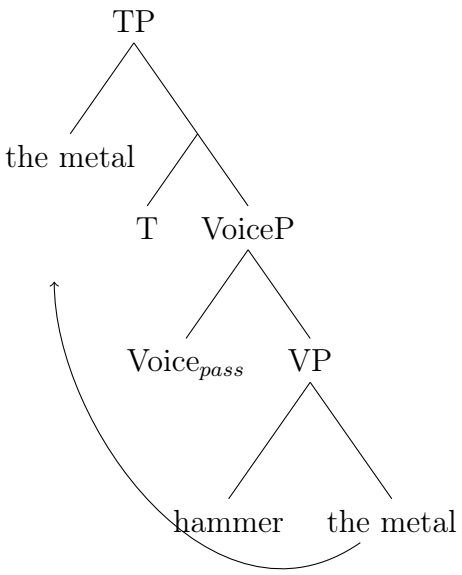


Figure 3.9: The Structure of (3.33)

The sentence in (3.33) shares a thematic relation with its resultative counterpart; in both cases, *the metal* is the theme of *hammer*. According to minimal UTAH, this shared thematic relation should correspond to a shared structural relation, but no such structural relation exists under Kratzer's analysis. This violation of UTAH, however, is not fatal as I will show in the next chapter when I adapt Kratzer's analysis to comply with UTAH.

3.2.1 Summary

In this chapter, I assessed several previous structural and parametric analyses of resultatives according to criteria I developed. The structures proposed for resultatives were assessed in terms of UTAH, while the parameters proposed were assessed based on whether they could be learned from the primary linguistic data, and whether they could be represented as lexical parameters. Several analyses were considered and ultimately rejected for not meeting the criteria I proposed.[7] However, Kratzer's (2005) structural analysis, and Snyder's (1995;

[7]Note that, because the critiques of these analyses were theoretical in nature, they will apply to all applications of the analyses in question.

2012) parametric analysis appear to be promising, for reasons that will be further developed in the next chapter.

While portions of Snyder's and Kratzer's analyses were rejected, they share a certain feature which I will retain in my final analysis: they are acquirable. Snyder's (1995) proposal is that the availability of resultatives in a language is linked to the availability of a productive N-N compounding in that language. Kratzer adopts this link and augments it with a link to predicative adjective agreement. She proposes that resultatives (and N-N compounding) are available only in those languages in which predicative adjectives do not show agreement with the subject.

(3.34) **German**

a. Die Teekanne leer trinken.
the teapot empty drink.
"to drink the teapot dry" (Kratzer 2005)

b. Wurmkanne
"worm+can" (Snyder 2001)

c. Die Teekanne ist leer (*-e).
The.FEM teapot is empty FEM"
"The teapot is empty."

(3.35) **Serbo-Croatian**[8]

a. *Dario je ofarbao kucu crveno
Dario COP painted house red.FEM
"Dario painted the house red."

b. *crve konzerva
"worm+can"

c. Kuća je crven *(-a)
house COP red FEM
"The house is red."

Both compounding, as in (3.34-b), and the lack of predicative adjective agreement, as in (3.34-c), are detectable in the primary linguistic data and, therefore, could reasonably be taken to be the positive data responsible for the setting of the resultative parameter. The account of the resultative parameter which I will develop in the remainder of this thesis will build on the idea that the positive evidence that triggers the parameter setting will be related to compounding and adjectival agreement. Before developing such an account, I will modify Kratzer's analysis to make it theoretically feasible.

[8]These examples are due to Mia Sara Misic (personal communication)

Chapter 4

The Structure of Resultatives

> Die Welt ist eine Glocke, die einen Riß hat: sie klappert, aber klingt nicht.
>
> Johann Wolfgang von Goethe

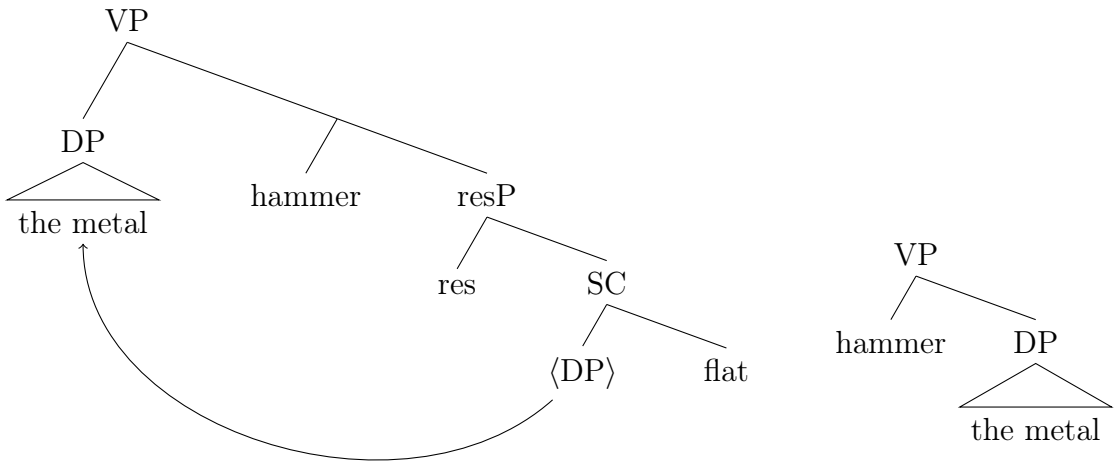
In the previous chapter, I discussed the failings of several previous analyses of adjectival resultatives. In this chapter, I will discuss two of those analyses—one structural and one parametric—and show how they can be modified to address the concerns raised in the previous chapter. The structural analysis, that of Kratzer (2005), was rejected as it stood because it did not comply with UTAH, but it has three features which I will retain: a small clause structure, theme raising, and a result head. The parametric analysis, that of Snyder (1995, 2012), was rejected because it did not comply with the Lexical Parameterization Hypothesis, but it was based on a learnable pattern and so I will be adopting a modified version of it.

4.1 Fixing the UTAH problem

The one issue with Kratzer's analysis is that it seems to violate UTAH. That is, there is a single θ-relation between *hammer* and *the metal* in both sentences in (4.1), that does not correspond to a single structural relation.

(4.1) a. Joe hammered the metal flat.
b. Joe hammered the metal.

According to Kratzer's analysis, *the metal* is the specifier of *hammer* in (4.1-a), but a standard analysis of (4.1-b) will place *the metal* as the complement of *hammer* If we were to modify Kratzer's analysis so that *the metal* is the complement of *hammer*, then we would need to attach the result phrase in a different position. I propose that the result phrase is adjoined to the VP, as shown in figure 4.2, allowing the DP to merge directly with the verb.



(a) *hammer the metal flat* (b) *hammer the metal*

Figure 4.1: Kratzer's (2004) analysis of resultatives

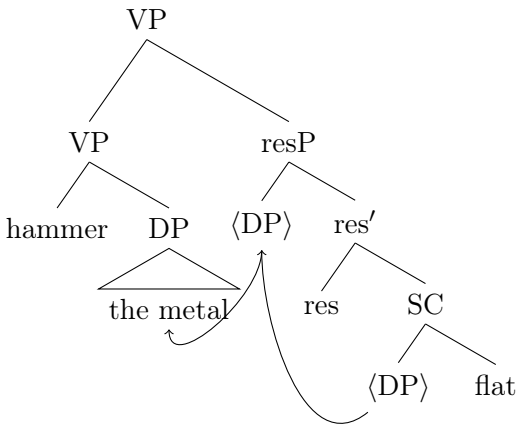


Figure 4.2: The Structure of resultatives

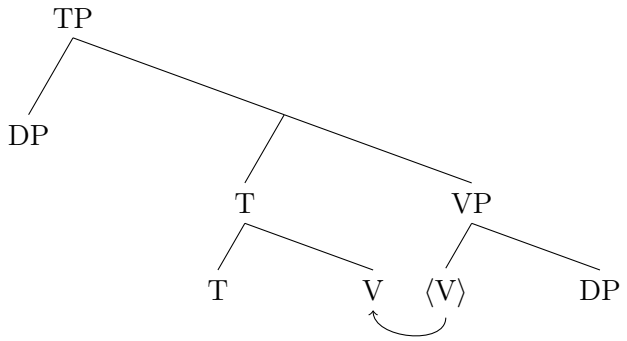
The modified analysis no longer violates UTAH, but it introduces two new issues. First, the movement operation between [Spec, res] and [Comp, V] does not target a c-commanding position. In other words, it is a sideward rather than an upward movement. Second, resP and VP are adjoined, meaning that they compose by conjunction. This is counterintuitive, however, since resultatives are inherently asymmetric, with the verb event causing the adjective state. I will address each of these in turn below.

4.1.1 Sideward movement

In figure 4.2, the object DP moves from [Spec res] to [Comp V]. The movement "chain" this operation forms is problematic because the head of the chain does not c-command the tail. Although this type of so-called sideward movement is generally barred, Nunes (2001)

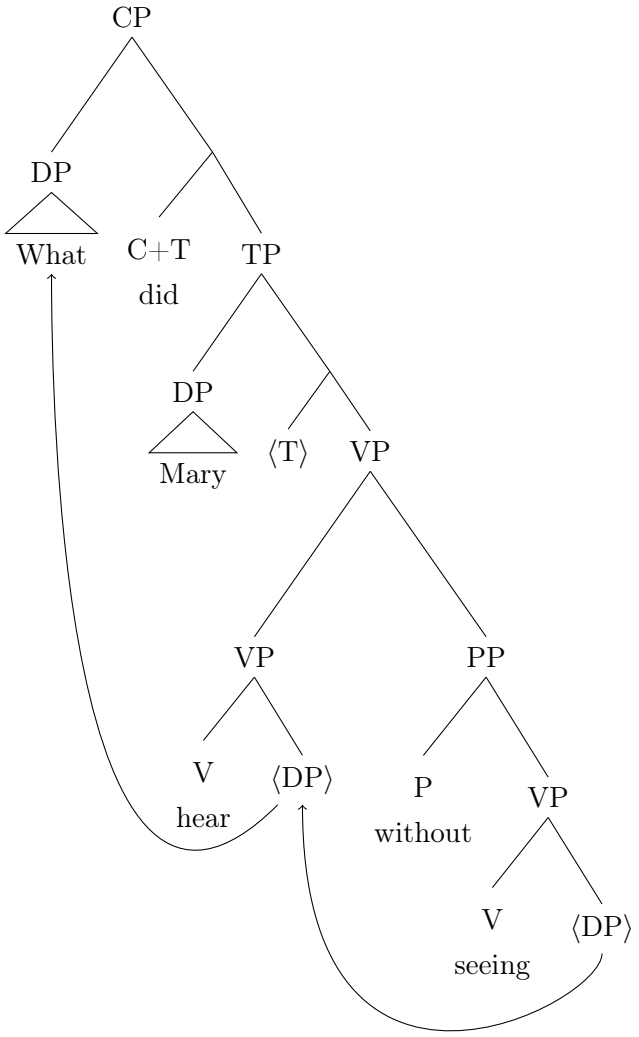
argues for a restricted version of sideward movement. Nunes argues that head movement and parasitic gaps both require a sideward movement operation, as they both create non-c-command dependencies.

(4.2) a. Head Movement



b. Parasitic gaps

What did Mary hear without seeing?



According to the standard definition of Merge, sideward movement should be impossible. The facts of parasitic gaps and head movement, however, suggest that a possibly complex operation with the net effect of sideward movement must be active in the grammar. I adopt the approach developed by Nunes (1995, 2001) as follows.

In order to explain sideward movement, Nunes hypothesizes that a movement operation is composed of a Copy operation followed by Merge. The operation Copy adds an object X to the workspace of a derivation provided that X is contained in an already constructed syntactic object.

(4.3) For a workspace W and a syntactic object X, Copy(W, X) = $W \cup \{X\}$ iff there is a syntactic object Z $\in$ W and Z contains X.

Merge, then, is a simpler operation, which replaces two members of a workspace with the set containing them. To see how a Copy+Merge theory of movement works, consider the derivation of passivization in table 4.1.

Stage	**Workspace**	
1	{[T, [Voice$_{pass}$ [see, [the, boy]]]]}	Copy([the, boy])
2	$\left\{\begin{array}{l}\text{[the, boy]},\\ \text{[T, [Voice}_{pass}\text{ [see, [the, boy]]]]}\end{array}\right\}$	Merge([the, boy], [T . . .])
3	{[[the, boy], [T, [Voice$_{pass}$ [see, [the, boy]]]]]}	

Table 4.1: The derivation of an English Passive

The Copy+Merge theory of movement allows us to derive sideward movement by holding the copied object in the workspace while another tree is built as in the derivation of figure 4.2 in table 4.2.

Stage	**Workspace**	
1	{[[the, metal], [res, [. . .]]]}	Copy([the, metal])
2	$\left\{\begin{array}{l}\text{[the, metal]},\\ \text{[[the, metal], [res, [. . .]]]}\end{array}\right\}$	Select(hammer)
3	$\left\{\begin{array}{l}\text{hammer},\\ \text{[the, metal]},\\ \text{[[the, metal], [res, [. . .]]]}\end{array}\right\}$	Merge(hammer, [the, metal])
4	$\left\{\begin{array}{l}\text{[hammer, [the, metal]]},\\ \text{[[the, metal], [res, [. . .]]]}\end{array}\right\}$	$\text{Merge}\left(\begin{array}{l}\text{[hammer, [the, metal]]},\\ \text{[[the, metal], [res [. . .]]]}\end{array}\right)$
5	{[[hammer, [the, metal]], [[the, metal], [res [. . .]]]]}	

Table 4.2: The derivation of an English resultative VP

Note that at stage 5 of the derivation in (4.3) the syntactic object in the workspace is representable as figure 4.2.

In order to constrain sideward movement, Nunes notes that its immediate, results such as the tree in figure 4.2, are unpronounceable. Assuming that decisions regarding linear order depend on c-command relations, and part of linearization is deciding which copy in a movement chain is to be pronounced, we would be unable to make a definitive linearization statement for the derived structure in (4.3). In order to linearize the movement chain of *the hammer*, there must be a copy which c-commands all other copies, meaning there must be a subsequent move from theme position to grammatical object position, which I represent as [Spec, AgrO][1] in figure 4.3.

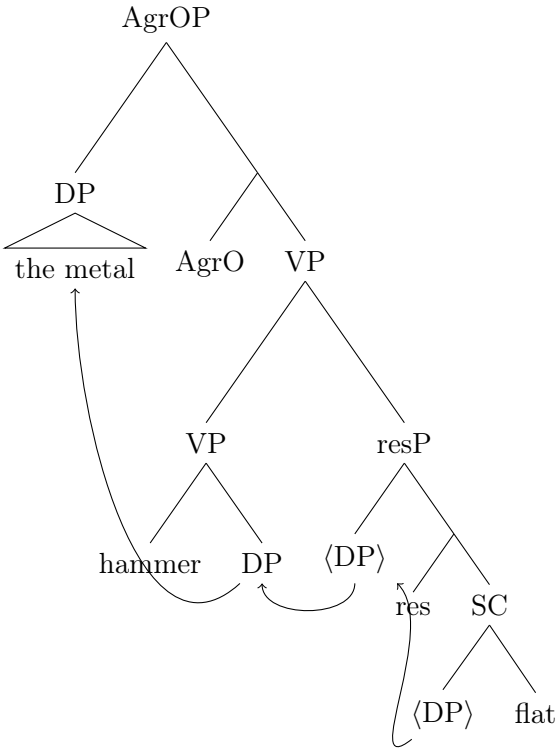


Figure 4.3: An English resultative AgrOP

Since the copy of *the metal* in [Spec, AgrO] c-commands all of the other copies, it will be pronounced and the lower copies will be deleted at the SM interface. So, assuming some mechanism for sideward movement, we are able to modify Kratzer's (2005) analysis of resultatives so that it is compatible with UTAH.

Since we have modified Kratzer's analysis, it is worth asking if our version will still compose semantically to give us the desired interpretation. In the next section, I argue that

[1]The choice to include AgrO in my structures does not indicate a particular commitment on my part to the existence of such a head, but rather to the fact that Object position seems to be distinct from internal argument position, and higher than VP. Further, I assume that either the movement to object position is covert, or there is head raising of the verb, such that it precedes the object.

not only can we retain the proper interpretation, but we can do so while assuming a simpler compositional system.

4.2 Composing resultatives

Kratzer (2005) adopts a neo-Davidsonian semantics for resultatives, meaning that they are analyzed as descriptions of eventualities rather than merely as relations between entities. Her syntactic analysis is repeated in figure 4.4 for reference. According to this analysis, the

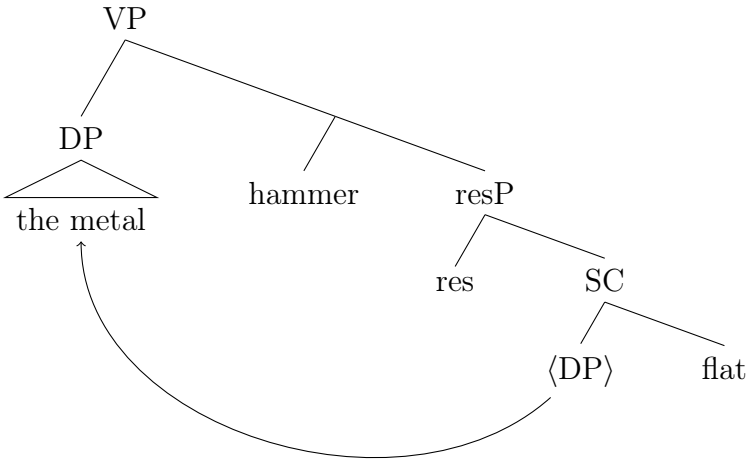


Figure 4.4: Kratzer's (2005) structural analysis of resultatives

small clause *the metal flat* is interpreted as the state description in (4.4), where the domain D_s is the domain of eventualities.

(4.4) $[\![\text{SC}]\!] = \lambda s_s \left[\textsc{state}(s)\&\mathbf{flat}(\mathbf{the\_metal})(s)\right]$

The verb *hammer* is interpreted as a predicate of events.

(4.5) $[\![\textit{hammer}]\!] = \lambda e_s \left[\textsc{event}(e)\&\mathbf{hammer}(e)\right]$

Note that Kratzer analyses resultative verbs as intransitives, meaning they do not take any entity arguments. Finally, she analyses the result head as a higher-order function, which expresses a causal relation between the event expressed by the verb and the state expressed by the small clause.

(4.6) $[\![\textit{res}]\!] = \lambda P_{\langle s,t\rangle} \lambda e_s \exists s_s \left[\textsc{event}(e)\&\textsc{state}(s)\&P(s)\&\textsc{Cause}(s)(e)\right]$

So, for Kratzer, the typed LF of *hammer the metal flat* is as in figure 4.5, and, according to her, *hammer* and the resP compose by an operation she calls Event Identification

(Kratzer 1996) which, in this instance, is equivalent to Predicate Modification generalized to eventualities.

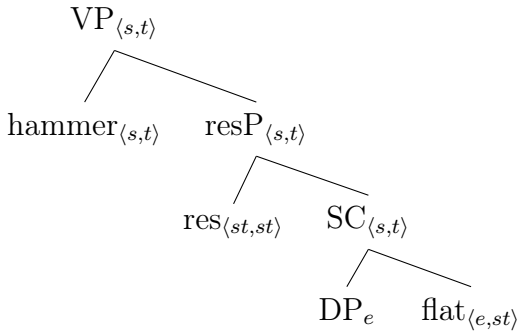


Figure 4.5: The LF of Kratzer's resultatives

(4.7) **Predicate Modification (eventuality version)**
If α is a branching node with daughters β and γ, both of which are of type $\langle s,t\rangle$, the $[\![\alpha]\!] = \lambda e_s[[\![\beta]\!](e)\&[\![\gamma]\!](e)]$

So, the interpretation of the VP in (4.6), can be derived as in (4.8).

(4.8) 1.$[\![\text{VP}]\!]$ = (Predicate Modification)
2.$\lambda e_s[[\![\text{hammer}]\!](e)\&[\![\text{resP}]\!](e)]$ =
3.$\lambda e_s[\textbf{hammer}(e)\&\exists s_s[\textsc{Cause}(s)(e)\&\textbf{flat}(\textbf{the\_metal})(s)]]$

So, the hammering event is identical to the event of causing the flatness state.

The sideward movement structure I propose, represented in figure 4.2, is only slightly different from Kratzer's. The main difference is that, in my analysis, the DP *the metal* is an argument of both the result adjective and the verb. In Kratzer's analysis, The DP is solely an argument of the result adjective. Despite this difference, the VP and resP in my structures are still predicted to be predicates of eventualities. As such, the VP and resP can be combined by predicate modification. As a demonstration of this, see the typed LF for the proposed structure in figure 4.6, and the partial derivation of its denotation in (4.9).

(4.9) 1.$[\![\text{VP}']\!]$ = (Predicate Modification)
2.$\lambda e_s[[\![\text{VP}]\!](e)\&[\![\text{resP}]\!](e)]$
3.$\lambda e_s[\lambda e'_s[\textbf{hammer}(e')\,\&\,\textsc{theme}(\textbf{the\_metal})(e')](e)\,\&\,[\![\text{resP}]\!](e)]$
4.$\lambda e_s[\lambda e'_s[\textbf{hammer}(e')\,\&\,\textsc{theme}(\textbf{the\_metal})(e')](e)\,\&$
$\lambda e''_s[\exists s_s[\textsc{Cause}(s)(e'')\&\textbf{flat}(\textbf{the\_metal})(s)]](e)]$
5.$\lambda e_s[\textbf{hammer}(e)\,\&\,\textsc{theme}(\textbf{the\_metal})(e)\,\&\,\exists s_s[\textsc{Cause}(s)(e)\&\textbf{flat}(\textbf{the\_metal})(s)]]$

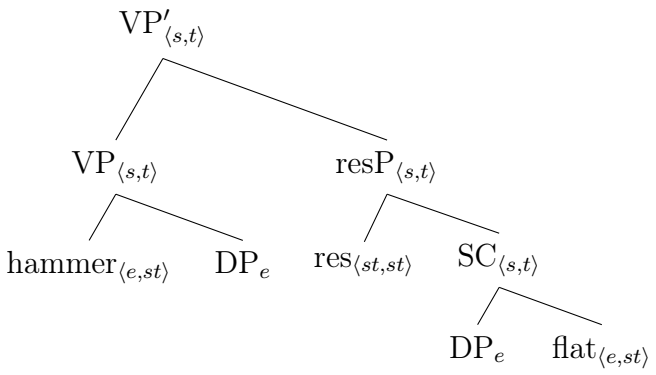


Figure 4.6: The LF of a resultative with sideward movement

Thus, with these adaptations, Kratzer's analysis of resultatives can be made UTAH-compliant. In the remainder of this chapter, I will discuss my parametric analysis of resultatives.

4.3 Where does the resultative parameter come from?

In the previous chapter, I discussed two desiderata for a parametric analysis. First, the parameter must be learnable, meaning there must be some variable in the primary linguistic data which the learner can detect and deduce a particular parameter setting from. Second, the variable must be represented in the lexicon. For clarity, I will refer to the variable detectable in the PLD as the surface variable, and its lexical representation as the lexical variable.

To my knowledge, there is only one proposed candidate for the surface variable in the generative literature, that is, Snyder's (1995; 2012) compounding parameter. According to the latest version of this parameter, a language allows resultatives iff it allows bare stem compounding. As discussed in the previous chapter, Snyder rejects the Lexical Parameterization Hypothesis, meaning that he does not propose a lexical variable. Instead, he situates the parameter in the operations of the CI interface. However, I will propose a lexical variable from which both the (un)availability of bare stem compounding and the (un)availability of adjectival resultatives can be derived.

To make such a proposal, we must make the intermediate hypothesis that a language allows bare stem compounding iff it allows bare stems, meaning there should be no languages that allow for bare stems but cannot compound them together. Now, a bare stem is merely an independent word with no inflectional material. Words are represented in most current theories of syntax as an acategorial root merged with a category-determining functional head (following Marantz 1997, but see also Borer 2005 for a similar proposal). Since roots are, by definition, featureless, any inflectional features on stems must be due to their category-determining heads. It follows from this that the (im)possibility of bare stems derives from the presence or absence of inflectional features on category-determining heads in the lexicon.

So, if we represent inflected category-determining heads as $v_{\varphi}, n_{\varphi}, adj_{\varphi}, etc.$ and their bare counterparts as $v_{\emptyset}, n_{\emptyset}, adj_{\emptyset}, etc.$, then the lexical version of Snyder's compounding parameter can be represented as in (4.10).

(4.10) LEX {includes, does not include} $v_{\emptyset}, n_{\emptyset}, adj_{\emptyset}, etc.$

Note that this is a fairly weak claim. A stronger claim would be that compounding languages have only uninflected category-determining heads. The weak claim, however, is sufficient for present purposes and is therefore adopted.

This version of the compounding parameter is lexical and, therefore, complies with the Lexical Parameterization Hypothesis. Furthermore, it is learnable from the primary linguistic data, since its external manifestation is the presence or absence of inflectional morphology. Since the inflectional morphology is detectable on the surface, its absence must also be detectable or at least deducible. This leaves us with questions regarding the initial state of the lexicon, and which parameter setting is the default, but those questions are beyond the scope of this thesis and will be set aside.

Chapter 5

Label Theory

"A rose by any other name would smell as sweet."
"Not if you called 'em stench blossoms."

"The Principal and the Pauper"
The Simpsons

In this chapter, I discuss a recent development in Chomsky's syntactic theory, which he refers to as *label theory*. Specifically, I discuss the background and content of Chomsky's (2013; 2015) proposal. Label theory will then be used in chapter 6 to explain the resultative parameter. This discussion is separate from the other theoretical background because as I write this thesis, label theory is in its nascent stage. In fact, in section 8.3 I will draw out two questions that Chomsky's label theory leaves unanswered and hypothesize an answer for each, thus modifying the theory.

5.1 Label theory and its motivations

Chomsky begins his proposal of label theory with a discussion of the minimalist program in general. In his estimation, the goal of the minimalist program has been to explain the universal properties of language as simply as possible. The properties he identifies are (i) the structure-dependence of rules, (ii) displacement, (iii) linear order and (iv) projection/labelling. He then argues that if we assume that linear order is a reflex of transfer to the SM interface, properties (i) and (ii) can be explained by assuming that Narrow Syntax is reducible to simplest Merge, as defined in (5.1).

(5.1) Merge(α, β) = $\{\alpha, \beta\}$

Unlike previous versions of Merge, simplest Merge does not include labelling. Chomsky ar-

gues that this is a welcome outcome, because labelling/projection is not as detectible in surface forms as the other properties of language, and has always been a theory internal notion. What's more, Chomsky argues, previous theories that bundle labelling with structure building have always stipulated labelling rather than deriving it. So, for instance, the phrase *see the girl* is stipulated to be a VP rather than a DP.

Chomsky proposes that labels are assigned post-syntactically by a special instance of minimal search called the Labelling Algorithm (LA), which operates as part of the CI system. LA operates iteratively in a top-down manner, searching each syntactic object for a "most prominent" element which can serve as the label. In the simplest case, an atomic element (a head X) merged with a complex object (a phrase YP[1]), the atomic element is found to be the most prominent element and, as such, is the label of the object.

(5.2) $\text{LA}(\{X, YP\}) = X$

The head-phrase case in (5.2) is trivial due to the inherent asymmetry in the structure. Labelling becomes more complicated when symmetric structures are considered, that is, when head-head and phrase-phrase structures are considered. To discover how these structures could be labelled, Chomsky considers examples of head-head and phrase-phrase structures that *are* generated by grammars, and hypothesizes why *they* are generated, while other instances are not. The only head-head structures that surface are those that result from the merger of acategorial roots and category-determining heads. So, structures like (5.3-a) are labellable, but those like (5.3-b) and (5.3-c) are not.

(5.3) a. $\{n, \sqrt{\text{WATER}}\}$
b. *$\{n, v\}$
c. *$\{\sqrt{\text{ICE}}, \sqrt{\text{WATER}}\}$

Chomsky proposes that roots are completely featureless and, therefore, invisible to LA. (5.3-a) thus receives the label n, while neither (5.3-b) nor (5.3-c) can be labelled.

As for phrase-phrase structures, Chomsky identifies two types that can be generated. The first type are what I will call phrase-trace structures. These are phrase-phrase structures in which one of the constituent phrases is a lower copy of a moved element. Following Moro (2000), Chomsky proposes that lower copies are invisible to LA.[2] The labelling of a phrase-

[1]Note that YP is actually a set $\{\alpha, \beta\}$, where α and β are arbitrary syntactic objects. The use of "YP" is not meant to indicate headedness, projection, maximality, or any notions other than the complexity of the object indicated by "YP".

[2]While this assertion might be characterized as a stipulation, I believe it is best understood as a statement of fact in label-theoretic terms. By hypothesis, all convergent SOs are labelled. Since there is a class of $\{XP, \lfloor YP \rfloor\}$ objects that are convergent, those objects must be labelled. The labels of those trace-phrase objects depend only on the unmoved constituent. That is, an object $\{\lfloor DP \rfloor, vP\}$ is labelled v, not D. It follows from this that the labelling algorithm does not take into account the properties of the moved constituent,

trace structure is, therefore, as shown in (5.4).

(5.4) $\text{LA}(\{XP, \lfloor YP \rfloor\}) = \text{LA}(XP)$
(Floor brackets ($\lfloor \cdot \rfloor$) here indicate that YP is a lower copy.)

The second type of phrase-phrase structure that can surface is what I will call agreement structures. These are phrase-phrase structures in which the two constituent phrases agree with one another for some feature. In these cases, the agreeing features serve as the label of the structure as in (5.5).

(5.5) $\text{LA}(\{XP_F, YP_F\}) = \langle F, F \rangle$

If neither of these two situations obtains for a given phrase-phrase structure, it will be unlabellable and result in a crash at the CI interface. To see how this works, consider the raising construction in (5.6) and the ungrammatical version of it in (5.7).

(5.6) a. The dishes seem to be dirty.
b. $\lfloor_\alpha$ The dishes [seem $\lfloor_\beta \lfloor$the dishes$\rfloor$, [to be dirty]]]]

(5.7) a. *It seems the dishes to be dirty
b. $\lfloor_\gamma$ It [seem $\lfloor_\delta$ the dishes, [to be dirty]]]]

The sentence in (5.6) has two relevant phrase-phrase structures which are labellable. The first is a trace-phrase structure, given in (5.8-a), which is labelled by the infinitive *to*, as demonstrated in (5.8-b).

(5.8) a. $\beta = \{\lfloor \text{the dishes} \rfloor, \{\text{to}, \text{be dirty}\}\}$
b. $\text{LA}(\beta) = \text{LA}(\{\text{to}, \text{be dirty}\}) = \text{to}$

The second is the agreement structure, given in (5.9-b), which is labelled by the agreeing φ features as in (5.9-b).

(5.9) a. $\alpha = \{\text{the}_\varphi \text{ dishes } \{T_\varphi, \text{ seem } \{\ldots\}\}\}$
b. $\text{LA}(\alpha) = \langle \varphi, \varphi \rangle$

The derivation of 5.7, however, crashes because the first phrase-phrase structure, shown as δ in (5.10), is unlabellable. The DP *the dishes* has not raised, so it is visible to LA in δ, and there is no φ-agreement between the_φ and $to_\emptyset$.

---

or, in other words, the moved object is invisible to the algorithm.

Note that, this being an empirical statement, it requires explanation. For instance, it remains to be explained how the LA can distinguish upper copies from lower copies. Such an explanation, however, is beyond the scope of this thesis.

(5.10) a. $\delta = \{\text{the}_{\varphi}\ \text{dishes}\,\{\text{to}_{\emptyset}\,\{\ldots\}\}\}$
b. $\text{LA}(\delta) = \text{Undefined}$

Also, Chomsky proposes that heads that bear only a partial set of features (*e.g.* English finite T_{φ}) cannot label unless they agree for those features with some other head. This is in contrast to heads that bear full feature sets (*e.g.* Italian $\text{T}\langle\varphi,\varphi\rangle$) or lack these features altogether (*e.g.* English non-finite $\text{T}_{\emptyset}$) which can label without agreement.[3,4]

At this point, I should note an issue that arises when giving label-based explanations of syntactic derivations. In previous theories, movement operations were described as being "driven" by some need. For example, in Government and Binding theories DPs undergo movement in order to get abstract Case. In minimalist theories, this has been generalized such that all movement is driven by the need to satisfy some feature. This led to debates about the exact mechanism that drives movement. Greed-based accounts, for instance, argue that an object moves if and only if such a move will satisfy one of its own features, while those that assume Enlightened Self Interest take the weaker stance that an object moves if and only if such a move will satisfy some feature on some argument.[5] While these accounts assumed an interface-based theory, they allow syntacticians to explain (un)grammaticality purely in terms of narrow syntax.

Label theory, however, assumes that all operations are free, that is, they do not require a trigger or a driver. This means, however, that an explanation of why an operation occurs or does not occur in a given derivation is slightly more complicated. The well-formedness of a structure is assessed at the interface; this means that entire phases are assessed at once. Consider, for instance, the successive *wh*-movement in (5.11), and how the two types of theories would account for it.

(5.11) Who$_i$ does Mary say t_i that Laura likes t_i?

The accounts of the final movement step ([Spec, C] to [Spec, C]) would be similar, as both theories assume that the highest C needs to agree with a *wh*-word, either for labelling or

[3]In Chomsky's words, English T, with its incomplete agreement feature-set, is "too weak" to label on its own. Contra Gallego (2017), the use of the word *weak* here does not seem to indicate, however, that Chomsky is resurrecting the theory of strong features that he developed in the early stages of the minimalist program (Chomsky 1995, 2000a). According to the theory of strong features, some formal features, by virtue of bearing the diacritic STRONG, needed to be satisfied *overtly*. In the theory presently under discussion, though, the use of the terms *weak* and *strong* are not technical terms but rather descriptive: LIs with incomplete sets of formal features are too weak to label a phrase and must be strengthened by Agree. This theory, of course, is incomplete and certainly has problems, but it does not inherit the problems of the theory of strong features.

[4]The proposal that incomplete feature sets cannot label while complete or null feature sets can label may seem odd at first, but it bears a similiarity to principles from other domains. For instance, in metrical phonology, an odd number of syllables in a word leads to issues in footing. Also, in chemistry, electrons are organized into orbitals, which can each contain up to two electrons and are chemically stable only if they are empty or contain a pair of electrons; unpaired electrons lead to chemical instability.

[5]See Lasnik (1999a) for a discussion of these two types of accounts.

SO	**LA(SO)**	
X	X	(X is not a root, and does not have an incomplete feature set)
$\{X, R\}$	LA(X)	(R is a root, X is not a root)
$\{X, YP\}$	LA(X)	
$\{\langle XP\rangle, YP\}$	LA(YP)	
$\{XP_F, YP_F\}$	$\langle F, F\rangle$	(XP and YP agree for F)
Otherwise	Undefined	

Table 5.1: A summary of the labelling algorithm

feature-satisfaction. The explanations of the first movement ([Comp, V] to [Spec, C]) however, are different. If movement operations were driven by some need for feature satisfaction rather than being free, then we would likely need to posit a feature on the lower C which must be satisfied by a *wh*-word. With free movement, however, the *wh*-word must move to the lower [Spec, C] because, if it doesn't, it cannot move to the higher [Spec, C] without violating Subjacency. Assuming a phase-based theory of subjacency, the first movement operation in (5.11) follows from the proposal that C is a phase head.

To be concrete: suppose that *who* does not move from its base position to its intermediate position in [Spec, *that*]. The phase head *that* would "trigger" the transfer of its complement, which includes every instance of *who*. When C_Q is finally merged, *who* is unavailable, as shown in (5.12).

(5.12) C_Q Mary T say that ~~Laura likes who~~

Since C_Q has only one Q feature, it cannot be the label of (5.12) and the derivation crashes.

To summarize, a syntactic derivation in label theory proceeds as follows. Structures are built by iteratively applying Merge (along with Select and Copy) to syntactic objects. At certain points a portion of a structure (*i.e.*, a phase) is transferred to the interfaces. At the CI interface, the labelling algorithm labels the transferred structure and all of the structures contained within the transferred structure. If the labelling algorithm fails to label any part of the transferred structure, the derivation crashes. A summary of the labelling algorithm is given in table 5.1.

Chomsky (2015) demonstrates that label theory has some empirical advantage over previous theories of syntax, but leaves at least two questions unanswered. The first question is why labels would be required by the CI interface at all, and the second question is how are host-adjunct structures labelled. In Part II I will propose answers to those questions, and consider the ramifications label theory has for the architecture of the grammar.

5.2 Summary

In this chapter, I reviewed Chomsky's (2013; 2015) label theory, according to which labels are assigned algorithmically at the CI interface and are required for proper interpretation at that interface. In the next chapter, I will use label theory to demonstrate that the (un)availability of resultative can be derived from the (un)availability of bare stem compounding.

Chapter 6

Deriving the Resultative Parameter

> All happy families are alike; every unhappy family is unhappy in its own way.
>
> *Anna Karenina*
> Leo Tolstoy

So far, I have clarified and developed my theoretical assumptions and proposed a structural analysis of resultatives, shown in figure 6.1, and an analysis of the resultative parameter, given in (6.1).

(6.1) A language L allows resultatives only if the lexicon of L includes $adj_{\emptyset}$.

Although they share a justification in minimalism, the structural analysis and the parametric analysis are independent proposals, and as such their combination must be justified. That is, I have argued that each component is plausible on its own, but it is altogether possible that they are not consistent with each other, or they are not able to give the correct empirical results. In this chapter, I argue that they are consistent with each other, and that, together, they can provide an explanation of the resultative parameter. I do so by first showing that the structure in Figure 6.1 can be derived in a language with uninflected adjectives ($adj_{\emptyset} \in$ LEX), and then showing how such a derivation fails in a language without uninflected adjectives ($adj_{\emptyset} \notin$ LEX).

In the first section, I will give a derivation of the English resultative VP *hammer the metal flat* and show that it converges at the CI interface (taking SM convergence for granted). In the second section, I will give two possible derivations of the ungrammatical French resultative VP *marteller le métal plat*, and show that deriving that VP leads to a CI crash while avoiding that crash blocks the derivation.

Before describing the derivations, I will reiterate and clarify my assumptions regarding the syntactic derivation. I adopt a slightly simplified version of the formal grammar developed

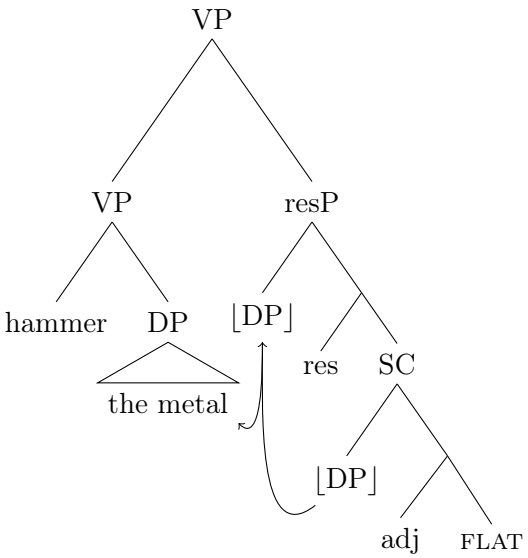


Figure 6.1: The structure of a resultative

by Collins and Stabler (2016) which I will augment slightly based on new assumptions. A *derivation* is defined as a finite sequence of *stages*, $\langle S_1, S_2 \ldots S_n \rangle$. Each stage S_i in a derivation is a pair $\langle LA_i, W_i \rangle$, where LA_i is a set of lexical items called the *lexical array* and W_i is a set of syntactic objects called the *workspace*. The computational operations (Merge, Select, Copy, Transfer) play the role that rules of inference play in deductive systems, that is, they map derivational stages onto subsequent stages. A given stage S_i *derives* a subsequent stage S_{i+1} if and only if some operation, applied to S_i, yields S_{i+1}.

6.1 A successful derivation in English

In many ways, successful derivations, like happy families, are uninteresting, but they still must be demonstrated in order to show where the crashing derivations go wrong.

To begin with, we derive the result phrase. The formal derivation of the resP is given in Table 6.1 and the resulting unlabelled structure is given in Figure 6.2.

Stage	**LA**	**Workspace**	
1	$\left\{\begin{array}{l}\sqrt{\text{FLAT}},\\ adj_{\emptyset},\\ res,\\ \text{DP}\end{array}\right\}$	$\emptyset$	Select($\sqrt{\text{FLAT}}$)
2	$\left\{\begin{array}{l}adj_{\emptyset},\\ res,\\ \text{DP}\end{array}\right\}$	$\left\{\sqrt{\text{FLAT}}\right\}$	Select($adj_{\emptyset}$)

3	$\left\{\begin{matrix} res, \\ \text{DP} \end{matrix}\right\}$	$\left\{\begin{matrix} adj_{\emptyset}, \\ \sqrt{\text{FLAT}} \end{matrix}\right\}$	Merge($adj_{\emptyset}$, $\sqrt{\text{FLAT}}$)
4	$\left\{\begin{matrix} res, \\ \text{DP} \end{matrix}\right\}$	$\left\{\left\{{}_{\alpha} adj_{\emptyset}, \sqrt{\text{FLAT}}\right\}\right\}$	Select(DP)
5	$\{res\}$	$\left\{\begin{matrix} \text{DP}, \\ \left\{{}_{\alpha} adj_{\emptyset}, \sqrt{\text{FLAT}}\right\} \end{matrix}\right\}$	Merge(DP, α)
6	$\{res\}$	$\left\{\left\{{}_{\beta}\text{DP}, \left\{{}_{\alpha} adj_{\emptyset}, \sqrt{\text{FLAT}}\right\}\right\}\right\}$	Select(res)
7	$\emptyset$	$\left\{\begin{matrix} res, \\ \left\{{}_{\beta}\text{DP}, \left\{{}_{\alpha} adj_{\emptyset}, \sqrt{\text{FLAT}}\right\}\right\} \end{matrix}\right\}$	Merge(res, β)
8	$\emptyset$	$\left\{\left\{{}_{\gamma} res, \left\{{}_{\beta}\text{DP}, \left\{{}_{\alpha} adj_{\emptyset}, \sqrt{\text{FLAT}}\right\}\right\}\right\}\right\}$	Copy(DP)
9	$\emptyset$	$\left\{\begin{matrix} \text{DP}, \\ \left\{{}_{\gamma} res, \left\{{}_{\beta}\text{DP}, \left\{{}_{\alpha} adj_{\emptyset}, \sqrt{\text{FLAT}}\right\}\right\}\right\} \end{matrix}\right\}$	Merge(DP, γ)
10	$\emptyset$	$\left\{\left\{{}_{\delta}\text{DP}, \left\{{}_{\gamma} res, \left\{{}_{\beta}\text{DP}, \left\{{}_{\alpha} adj_{\emptyset}, \sqrt{\text{FLAT}}\right\}\right\}\right\}\right\}\right\}$	Transfer(β)

Table 6.1: The derivation of an English resP

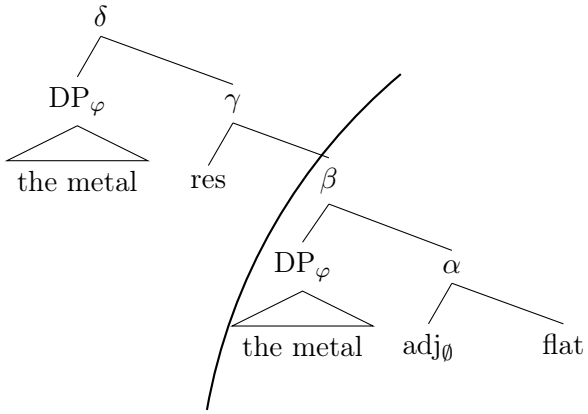


Figure 6.2: An unlabelled resP

Assuming res is a phase head, its complement β is transferred and must be labelled along with the SOs it contains. The small clause β ($\left\{{}_{\beta}\lfloor\text{DP}\rfloor, \left\{{}_{\alpha} adj_{\emptyset}, \sqrt{\text{FLAT}}\right\}\right\}$) is a Phrase-Phrase structure, but since one of its constituent parts, the DP, is a lower copy, that part is invisible to the labelling algorithm. Therefore, only the adjective ($\left\{{}_{\alpha} adj_{\emptyset}, \sqrt{\text{FLAT}}\right\}$) is available to provide a label. Assuming that roots are inert for labelling, and that uninflected categorizing heads can label, $adj_{\emptyset}$ is selected to label β. Since α is a head-root structure, it is labelled by the categorizing head $adj_{\emptyset}$. So, β is successfully labelled and therefore convergent at the CI interface.

(6.2) $\quad \text{LA}(\left\{{}_{\beta}\lfloor\text{DP}\rfloor, \left\{{}_{\alpha} adj_{\emptyset}, \sqrt{\text{FLAT}}\right\}\right\}) = \left[{}_{adj}\ \lfloor\text{DP}\rfloor, \left[{}_{adj}\ adj_{\emptyset}, \sqrt{\text{FLAT}}\right]\right]$

Since γ and δ are not transferred along with β, we do not need to discuss their labels yet.

We then derive the next phase as in Table 6.2. Note that resP and the DP are in the initial lexical array for this derivation. While this stipulation is necessary to derive the next phase, a number of aspects of it are poorly understood. I believe that this lack of understanding is directly related to the nature of the Transfer operation, which seems to be a stand-in for the interfaces. Since a full understanding of the interfaces requires an entire research program, it is decidedly beyond the scope of this thesis, and I will make do with stipulation here.

The unlabelled structure is given in Figure 6.3.

Stage	**LA**	**Workspace**	
1	$\{\text{resP}, \text{DP}, v, \sqrt{\text{HAMMER}}, \text{AgrO}\}$	$\{\}$	Select($\sqrt{\text{HAMMER}}$)
2	$\{\text{resP}, \text{DP}, v, \text{AgrO},\}$	$\{\sqrt{\text{HAMMER}}\}$	Select(v)
3	$\{\text{resP}, \text{DP}, \text{AgrO}\}$	$\{\sqrt{\text{HAMMER}}, v\}$	Merge(v, $\sqrt{\text{HAMMER}}$)
4	$\{\text{resP}, \text{DP}, \text{AgrO}\}$	$\{\{{}_{\alpha}v, \sqrt{\text{HAMMER}}\}\}$	Select(DP)
5	$\{\text{resP}, \text{AgrO}\}$	$\{\text{DP}, \{{}_{\alpha}v, \sqrt{\text{HAMMER}}\}\}$	Merge(DP, α)
6	$\{\text{resP}, \text{AgrO}\}$	$\{\{{}_{\beta}\text{DP}, \{{}_{\alpha}v, \sqrt{\text{HAMMER}}\}\}\}$	Select(resP)
7	{AgrO}	$\{\text{resP}, \{{}_{\beta}\text{DP}, \{{}_{\alpha}v, \sqrt{\text{HAMMER}}\}\}\}$	Merge[1](β, resP)
8	{AgrO}	$\{\{\zeta\ \{{}_{\beta}\text{DP}, \{{}_{\alpha}v, \sqrt{\text{HAMMER}}\}\}, \text{resP}\}\}$	Select(AgrO)
9	$\emptyset$	$\{\text{AgrO}, \{\zeta\ \{{}_{\beta}\text{DP}, \{{}_{\alpha}v, \sqrt{\text{HAMMER}}\}\}, \text{resP}\}\}$	Merge(AgrO, ζ)
10	$\emptyset$	$\{\{{}_{\eta}\text{AgrO}, \{\zeta\ \{{}_{\beta}\text{DP}, \{{}_{\alpha}v, \sqrt{\text{HAMMER}}\}\}, \text{resP}\}\}\}$	Copy(DP)
11	$\emptyset$	$\{\text{DP}, \{{}_{\eta}\text{AgrO}, \{\zeta\ \{{}_{\beta}\text{DP}, \{{}_{\alpha}v, \sqrt{\text{HAMMER}}\}\}, \text{resP}\}\}\}$	Merge(DP, η)
12	$\emptyset$	$\{\{{}_{\kappa}\text{DP}, \{{}_{\eta}\text{AgrO}, \{\zeta\ \{{}_{\beta}\text{DP}, \{{}_{\alpha}v, \sqrt{\text{HAMMER}}\}\}, \text{resP}\}\}\}\}$	...

Table 6.2: The derivation of an English resultative VP

[1]This instance of "Merge" is, in fact, an instance of adjunction. I represent it as Merge in order to maintain

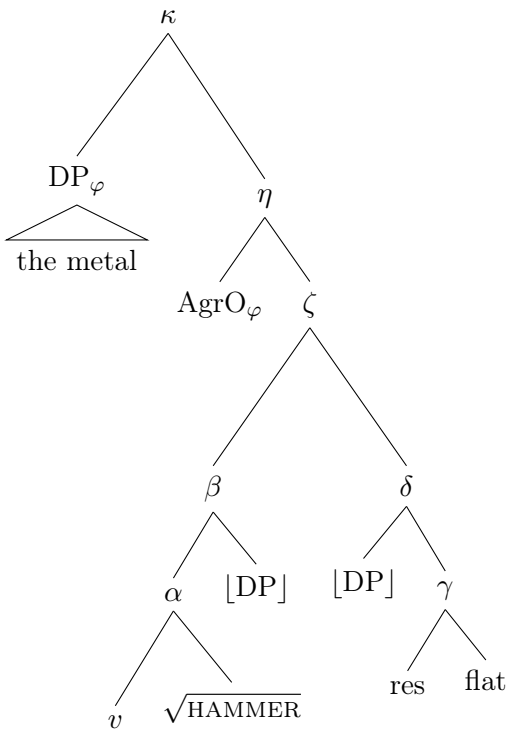


Figure 6.3: An unlabelled English resultative

When this is transferred, triggered, presumably, by the merging of the phase head Voice, it is labelled just as any transitive VP would be. The largest object κ is a phrase-phrase structure with agreeing features, so it will receive a $\langle\varphi,\varphi\rangle$ label. The remaining objects will receive head-labels, with the exception of the host-adjunct structure ζ.

As in previous theories of grammar, host-adjunct structures are problematic in label theory. Structures like ζ are phrase-phrase structures, meaning they can only be labelled if the constituent parts agree for some feature, or one of the constituent parts is somehow invisible. Since, almost by definition, adjuncts are not selected by their hosts[2] it is unlikely that there is agreement between adjuncts and their hosts. If there is no agreement, then the only way for ζ to be labellable is if one of its parts is inert. Since host-adjunct structures, again almost by definition, have the properties of the host and not those of the adjunct, it is reasonable to think that the host is active and the adjunct is inert. I will therefore provisionally assume that the host β provides the label for ζ, and δ is inert. This matter will be addressed in section 8.3.

In this section, we have seen how a convergent resultative is derived in English. The next section, however, is truly where the rubber meets the road. There I show that the same grammar that generates resultatives in English will fail to generate them in French. As we

the simplicity of the formal grammar.

[2]Cartographic approaches to syntax (Cinque and Rizzi 2009, and references therein), however, assume that adjectives and adverbs are selected by functional heads. This assumption does not, to my knowledge, extend to phrase- or clause-sized modifiers though.

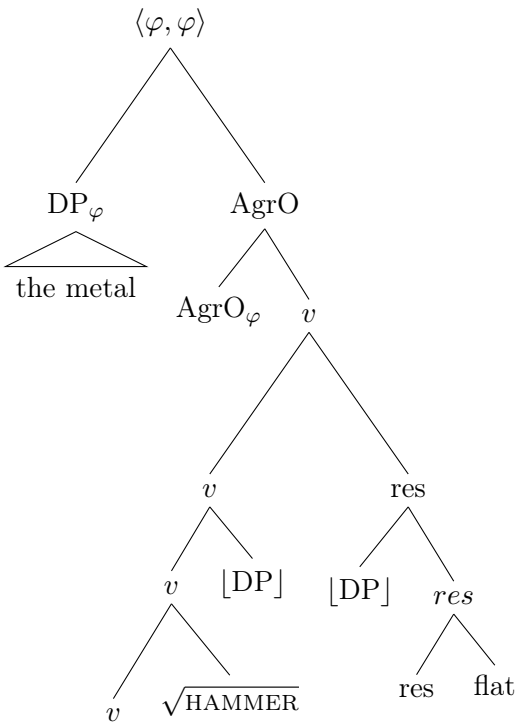


Figure 6.4: A labelled English resultative

will see, the crucial operation, the one which will be blocked in French, is the movement of DP from the adjectival Small Clause.

6.2 Two crashing derivations in French

By hypothesis, the only relevant difference between English and French is that the lexicon of French contains only inflected category heads (specifically $adj_{\emptyset} \notin$ LEX and $adj_{\varphi} \in$ LEX). In this section, I will attempt to derive a resultative with an adj_{φ} and show that such a derivation inevitably either crashes due to failure to label or simply does not derive a resultative. The first attempt will reproduce an English derivation and crash, while the second will avoid that crash but fail to move the object DP into the VP, and thus will be unable to derive the proper structure.

6.2.1 Crashing Derivation

Consider the derivation described in section 6.1 with adj_{φ} replacing $adj_{\emptyset}$. The resP will be derived in the same fashion, as shown in Table 6.3.

Stage	LA	Workspace

1	$\left\{\begin{array}{l}\sqrt{\text{PLAT-}},\\ adj_{\varphi},\\ res,\\ \text{DP}\end{array}\right\}$	$\emptyset$	Select($\sqrt{\text{PLAT-}}$)
2	$\left\{\begin{array}{l}adj_{\varphi},\\ res,\\ \text{DP}\end{array}\right\}$	$\left\{\sqrt{\text{PLAT-}}\right\}$	Select(adj_{φ})
3	$\left\{\begin{array}{l}res,\\ \text{DP}\end{array}\right\}$	$\left\{\begin{array}{l}adj_{\varphi},\\ \sqrt{\text{PLAT-}}\end{array}\right\}$	Merge(adj_{φ}, $\sqrt{\text{PLAT-}}$)
4	$\left\{\begin{array}{l}res,\\ \text{DP}\end{array}\right\}$	$\left\{\left\{{}_{\alpha}adj_{\varphi}, \sqrt{\text{PLAT-}}\right\}\right\}$	Select(DP)
5	$\{res\}$	$\left\{\begin{array}{l}\text{DP},\\ \left\{{}_{\alpha}adj_{\varphi}, \sqrt{\text{PLAT-}}\right\}\end{array}\right\}$	Merge(DP, α)
6	$\{res\}$	$\left\{\left\{{}_{\beta}\text{DP}, \left\{{}_{\alpha}adj_{\varphi}, \sqrt{\text{PLAT-}}\right\}\right\}\right\}$	Select(res)
7	$\emptyset$	$\left\{\begin{array}{l}res,\\ \left\{{}_{\beta}\text{DP}, \left\{{}_{\alpha}adj_{\varphi}, \sqrt{\text{PLAT-}}\right\}\right\}\end{array}\right\}$	Merge(res, β)
8	$\emptyset$	$\left\{\left\{{}_{\gamma}res, \left\{{}_{\beta}\text{DP}, \left\{{}_{\alpha}adj_{\varphi}, \sqrt{\text{PLAT-}}\right\}\right\}\right\}\right\}$	Copy(DP)
9	$\emptyset$	$\left\{\begin{array}{l}\text{DP},\\ \left\{{}_{\gamma}res, \left\{{}_{\beta}\text{DP}, \left\{{}_{\alpha}adj_{\varphi}, \sqrt{\text{PLAT-}}\right\}\right\}\right\}\end{array}\right\}$	Merge(DP, γ)
10	$\emptyset$	$\left\{\left\{{}_{\delta}\text{DP}, \left\{{}_{\gamma}res, \left\{{}_{\beta}\text{DP}, \left\{{}_{\alpha}adj_{\varphi}, \sqrt{\text{PLAT-}}\right\}\right\}\right\}\right\}\right\}$	Transfer(β)

Table 6.3: The derivation of a French resP

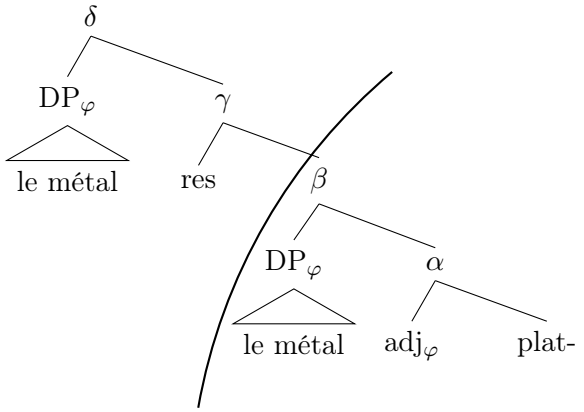


Figure 6.5: An unlabelled resP

Upon Transfer, β must be labelled and since the DP has been moved, it is invisible to the labelling algorithm. The label of β, then will be the label of α ($\left\{adj_{\varphi}, \sqrt{\text{PLAT-}}\right\}$). In the English case, $adj_{\emptyset}$ was able to provide a label, but following Chomsky (2015), the French adj_{φ} is too weak to label without being strengthened by φ-agreement. The DP which is merged with α would agree with adj_{φ}, but since it is a lower copy, it is inert, and therefore cannot

take part in agreement. Since adj_{φ} has not been agreed with, it remains too weak to label α and, by extension, too weak to label β. The derivation, then, crashes due to a failure to label.

So, attempting to derive a resultative in French as we did in English yields a crash at the interfaces. Perhaps, though, there is another way to derive resultatives without causing a crash. In the next section I attempt such an alternative derivation, but ultimately this attempt, while it doesn't crash, will not derive a resultative.

6.2.2 Failed Derivation

The fatal flaw in the previous derivation was moving the DP from the small clause before it could agree with adj_{φ}. Consider the following derivation of Figure 6.6 given in Table 6.4.

Stage	**LA**	**Workspace**	
1	$\left\{\begin{array}{l}\sqrt{\text{PLAT-}},\\ adj_{\varphi},\\ res,\\ \text{DP}\end{array}\right\}$	$\emptyset$	Select($\sqrt{\text{PLAT-}}$)
2	$\left\{\begin{array}{l}adj_{\varphi},\\ res,\\ \text{DP}\end{array}\right\}$	$\left\{\sqrt{\text{PLAT-}}\right\}$	Select(adj_{φ})
3	$\left\{\begin{array}{l}res,\\ \text{DP}\end{array}\right\}$	$\left\{\begin{array}{l}adj_{\varphi},\\ \sqrt{\text{PLAT-}}\end{array}\right\}$	Merge(adj_{φ}, $\sqrt{\text{PLAT-}}$)
4	$\left\{\begin{array}{l}res,\\ \text{DP}\end{array}\right\}$	$\left\{\left\{{}_{\alpha}adj_{\varphi}, \sqrt{\text{PLAT-}}\right\}\right\}$	Select(DP)
5	$\{res\}$	$\left\{\begin{array}{l}\text{DP},\\ \left\{{}_{\alpha}adj_{\varphi}, \sqrt{\text{PLAT-}}\right\}\end{array}\right\}$	Merge(DP, α)
6	$\{res\}$	$\left\{\left\{{}_{\beta}\text{DP}, \left\{{}_{\alpha}adj_{\varphi}, \sqrt{\text{PLAT-}}\right\}\right\}\right\}$	Select(res)
7	$\emptyset$	$\left\{\begin{array}{l}res,\\ \left\{{}_{\beta}\text{DP}, \left\{{}_{\alpha}adj_{\varphi}, \sqrt{\text{PLAT-}}\right\}\right\}\end{array}\right\}$	Merge(res, β)
8	$\emptyset$	$\left\{\left\{{}_{\gamma}res, \left\{{}_{\beta}\text{DP}, \left\{{}_{\alpha}adj_{\varphi}, \sqrt{\text{PLAT-}}\right\}\right\}\right\}\right\}$	Transfer(β)

Table 6.4: The derivation of a French resP with an *in situ* DP

Unlike the case in subsection 6.2.1, the transferred object β will be labellable. The *in situ* DP will φ-agree with adj_{φ}, and β will be labelled with the pair $\langle\varphi,\varphi\rangle$. Furthermore, since adj_{φ} has been strengthened by agreement, it will be able to label α. Thus, the transferred object is labelled as in 6.3.

(6.3) $\quad \text{LA}(\beta) = \left[{}_{\langle\varphi,\varphi\rangle}\text{DP}\left[{}_{adj}adj, \sqrt{\text{PLAT}}\right]\right]$

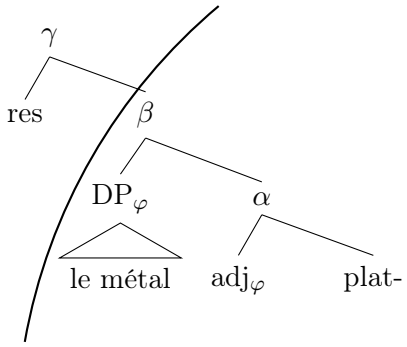


Figure 6.6: An unlabelled French resP with an *in situ* DP

The resP can be derived without a crash, but this will turn out to be something of a Pyrrhic victory. The French small clause is labellable because the DP remains *in situ*, but this same fact means that the DP is now inaccessible to further operations such as Copy and Merge. If we cannot copy and remerge the DP, we will unable to continue the derivation. We can't merge a DP with a verb if that DP is inaccessible to Merge. Thus, our attempt to avert a crash has, in fact, doomed the derivation.

6.3 On Bare Stem Compounding

So far, I have demonstrated that, given my theoretical assumptions, resultatives can be generated only if the result adjective is categorized by a featureless $adj_\emptyset$ head. I also proposed that a child acquires featureless categorizing heads if they encounter productive bare-stem compounding (BSC) in their PLD. In this section, I will propose an analysis of bare-stem compounding that is consistent with these proposals and my theoretical assumptions.

If we restrict ourselves to endocentric bare-stem compounding like *bourbon bar*, for instance, then our analysis must be consistent with the possibility of endocentricity. That is, a proper analysis of *bourbon bar* must naturally explain why it describes a type of bar rather than a type of bourbon. Furthermore, an analysis of BSC must allow for the fact that a language's ability to generate these compounds depends on the properties of the categorizing heads in that language.

I will discuss three possibilities below: one in which a compound is formed by directly merging roots together, a second in which two categorized roots ($\left\{cat, \sqrt{\text{Root}}\right\}$, or *stems*) are directly merged together, and a third in which a stem merges with a root. As we shall see, only the third analysis can account for endocentricity and parametric variation without major stipulation.

6.3.1 Root-root compounding

The first analysis that I will consider is one in which roots merge directly with each other as in figure 6.7. Immediately, we can see that the symmetrical nature of merge (*i.e.* the fact

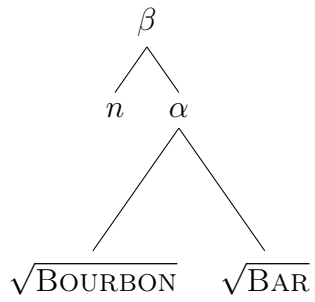


Figure 6.7: A root-root analysis of compounding

that merge creates an unordered set) renders endocentricity impossible; *bourbon bar* would be indistinguishable from *bar bourbon*.

Setting this problem aside for the moment, could this analysis account for the parametric variation? That is, can we show that (6.4) crashes, while (6.5) converges?

(6.4) $*[_{\beta}\, n_F, [_{\alpha}\sqrt{\textsc{Bourbon}}, \sqrt{\textsc{Bar}}]]$

(6.5) $[_{\beta}\, n_{\emptyset}, [_{\alpha}\sqrt{\textsc{Bourbon}}, \sqrt{\textsc{Bar}}]]$

Since n is the least embedded atomic element in both (6.4) and (6.5), it would label both phrases if it were strong enough. The French n_F in (6.4) will, of course, need to be strengthened by Agree, but compare the proposed compound structure with that of a simple noun in (6.6).

(6.6) $[_{\alpha} n_F, \sqrt{\textsc{bar}}]$

Note that in both the simple noun (6.6) and the compound noun (6.4), the categorizing head n_F is an immediate constituent of the phrase in question. Furthermore, in both cases, n_F is the only possible labeller, as all of the other constituents are roots. It follows, then, that β in (6.4) and α in (6.6) are indistinguishable with respect to labelling. Therefore, a root-root analysis of BSC gives us no principled way of ruling out compound nouns in French-like languages, without also ruling out simple nouns.

Since this analysis lacks both necessary properties for compounds, I will set it aside.

6.3.2 Stem-stem compounding

In the second analysis, bare-stem compounds are created by merging two stems (*i.e.*, $\{n, \sqrt{\textsc{root}}\}$). Like the root-root possibility, the stem-stem possibility is symmetrical, as we

can see in figure 6.8 As with the root-root option, the symmetry of this structure precludes

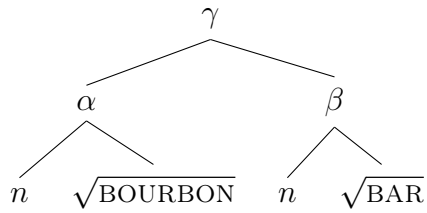


Figure 6.8: A stem-stem analysis of compounding

endocentricity. Again setting this problem aside, let's consider how it fares with respect to parametric variation.

If we consider the compound γ as a whole, we can see that it is a phrase-phrase structure, and therefore labellable in only two situations: either the constituent phrases has moved, or there is agreement between the "heads" of the two constituent phrases (the two n heads in this case). The first is inapplicable since both members of the compound remain *in situ*. As for the second, in the case of an English-type language, there is certainly no agreement between the two n heads, as they have, by hypothesis, no features, which are required for Agree. A stem-stem analysis of compounds, therefore, seems to wrongly predict that English-type grammars do not generate bare-stem compounds; another strike against them. French-like languages, on the other hand, have feature-bearing n heads which are in the ideal structural configuration to agree with each other, but it is unlikely that they would undergo agreement with each other due to the types of features that they have. The standard cases of agreement involve a featural asymmetry between the agreeing heads—an interpretable feature checks an uninterpretable feature; a valued feature values an unvalued feature—but in the structure in figure 6.8, the would-be agreeing heads are of the same type, and therefore have identical featural endowments. Since there is no featural asymmetry, it is likely that there can be no agreement between the two n_F heads in French-type languages. If they cannot agree with each other, then, just as in the case of English-type languages, we would expect the structure in figure 6.8 to be unlabellable. If they can agree with each other, then γ in figure 6.8 should be labelled $\langle F, F \rangle$, and therefore, the structure in figure 6.8 should be a licit compound. So, the stem-stem analysis of compounds may wrongly predict that French-type grammars could generate bare-stem compounds.

Since a stem-stem analysis accounts for neither endocentricity nor the parametric variation with respect to bare-stem compounding, we will set it aside and move on to the third possible analysis.

6.3.3 Root-stem compounding

The final possibility is that compounds are formed by merging a root with a stem, as in figure 6.9. We can immediately see that this is an asymmetric structure, and, as such, one

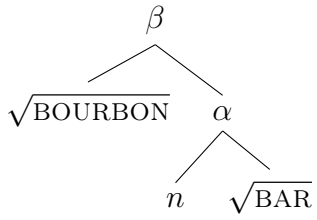


Figure 6.9: A root-stem analysis

that is in principle able to capture endocentricity. Under this view, *bourbon bar* names a type of bar rather than a type of bourbon by virtue of the fact that the *bar* root merges directly with the category-determining head, while *bourbon* does so indirectly. Compare this with the case of a Saxon genitive, such as *the president's men*, which names some men rather than the president. This endocentricity can be captured by the structure in figure 6.10, by virtue of the fact that *men* is the complement, merging directly with the determiner *-'s*, while *the president* is its specifier.

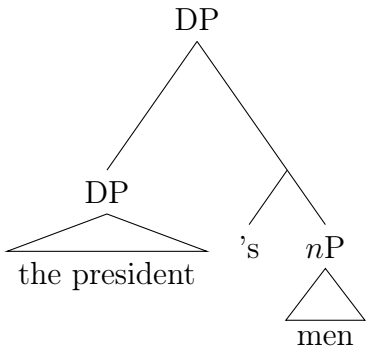


Figure 6.10: The Saxon genitive

Now, what about the parametric variation? Let's consider how the structure in figure 6.9 would be labelled in the case of an English-type language. The least embedded atom in β is the root BOURBON, which is invisible for labelling. Therefore we must look for the next-least-embedded atom, which, assuming the root BAR is invisible, would be $n_\emptyset$. So, the label of β would be $n_\emptyset$, as it would be for α as well.

Now consider the French-type language, where the featureless $n_\emptyset$ is replaced by n_F with an incomplete feature set as in figure 6.11. Recall that n_F cannot label a phrase unless it is strengthened to $n_{\langle F,F\rangle}$ by Agree. So, the structure in figure 6.11, is unlabellable on its own, Furthermore, the n_F head is embedded beneath a root, making it a more remote target for agreement with some head to be merged later. If n_F is too remote to be agreed

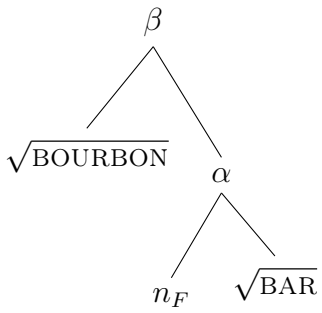


Figure 6.11: A French-like Root-Stem structure

with, then it will be too weak to label α or β, and therefore the structure in figure 6.11 would crash. This difficulty would not, however, emerge in the case of a simple noun (*e.g.*, $\{n_F, \sqrt{\text{BAR}}\}$), because the categorizing head n_F is not embedded, and therefore will be available for agreement with a higher head.

A Root-Stem analysis of BSC, unlike the other two alternatives described above, can capture the fact that languages without featureless $cat_\emptyset$ heads cannot produce bare stem compounds. Since a Root-Stem analysis of BSC seems to be the only one that correctly accounts for both endocentricity and parametric variation, it is likely the correct analysis.[3]

In her analysis of BSC, which she refers to as *primary compounds*, Harley (2009) adopts what is essentially a Root-Stem structure, although her version, demonstrated in figure 6.12 differs from mine in a few interesting ways.

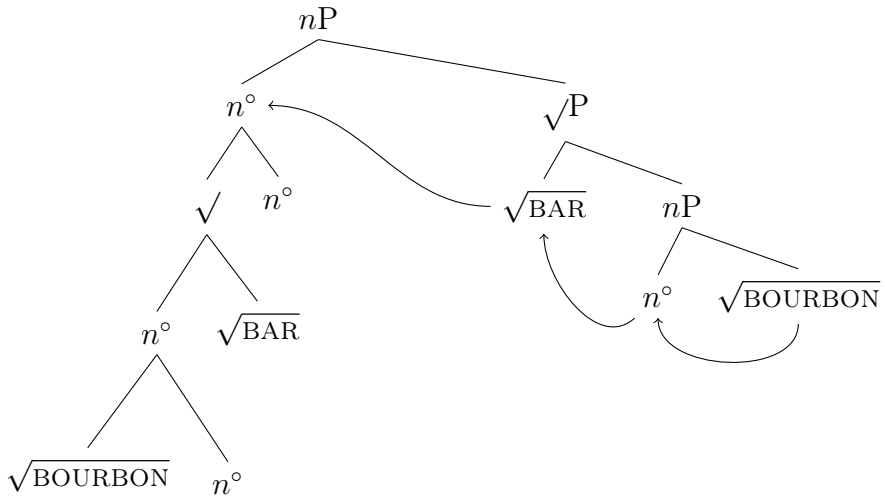


Figure 6.12: Harley's (2009) analysis for *bourbon bar*

Perhaps the biggest difference between Harley's analysis and mine is that the "modifier" (*bourbon*) is merged below the "head" (*bar*). Related to that difference, is that Harley assumes

[3]It is not immediately obvious how this analysis would work for compounds formed by more complex stems such as *attachment disorder*, or *deionized water bottle*. In order to extend my analysis to these cases, though, we would need an explicit theory of morphological derivation. Since such theorizing is outside the scope of this thesis, I will leave it for later research.

that the surface appearance of *bourbon bar* is due to a series of head movement operations, which Harley refers to as *incorporation*. Despite the similarities and differences between this analysis and mine, it would be difficult to compare the respective merits of the two, as they follow from two distinct sets of theoretical assumptions. Harley, for instance, assumes that roots can both project phrases and select complements, assumptions that are explicitly ruled out under Chomsky's (2013) label theory, which this thesis adopts. Comparing the two analyses, then, would require a deep comparison of the assumptions underlying them. Such a comparison is beyond the scope of this thesis so I will set Harley's analysis aside. In this chapter, I have shown that resultatives are, in fact, derivable in a label-theoretic grammar, provided the resultative adjective is categorized by an uninflected head $adj_{\emptyset}$. I then demonstrated that the same derivation runs into difficulties if the resultative adjective is categorized by an inflected head adj_{φ}. I do not claim to have demonstrated that deriving resultatives with adj_{φ} is completely impossible (such a demonstration may be impossible). Rather, I have merely pointed out two ways not to derive resultatives. If we make the hypothesis that the parameter in (6.7) determines whether a language allows resultatives, then the demonstrations in this chapter represent an explanation of the resultative parameter, that is, an answer to the question of how resultatives are acquired, parameterized, and generated.

(6.7) LEX {includes, does not include} $v_{\emptyset}, n_{\emptyset}, adj_{\emptyset}, etc.$

Such an explanation was the stated goal of this thesis, but the hypotheses made in its service open up a number of questions, which I will address in Part II of the thesis.

Part II

Dealing with the consequences

Chapter 7

Small clauses in French-like languages

All the pieces matter

David Simon

In the preceding chapters, I presented an explanation of the absence of adjectival resultatives in French-type languages. Without further comment, however, this explanation seems to wrongly predict that depictives and copular clauses are also ruled out in French. In this chapter, I will refine my proposal so that this prediction is removed. Rather than changing any of the claims and hypotheses I have made thus far, I will refine another part of the grammar, namely, the Agree operation. With this clarification, resultatives in French can be ruled out without also barring depictives and copular clauses. Furthermore, the clarified grammar turns out to provide a straightforward explanation of *wanna*-contraction in English.

7.1 The faulty predictions: *depictives, *copular clauses

In order to account for the lack of resultatives in French-type languages, I argued that DP movement from a small clause is barred in these languages. The fact that French allows copular clauses and depictives, however, means that this restriction, as stated, does not hold across the board. In other words, the grammar must be able to generate (7.1) and (7.2), but not (7.3).

(7.1) Jeanne est grand -e.
Jeanne is tall -FSg
"Jeanne is tall."

(7.2) Marie mange la viande crue
Marie eats the.FSg meat raw.Fsg
"Marie eats meat raw"

(7.3) *Sophie a martellé le métal plat.
Sophie AUX hammered the.MSg metal flat.MSg
"Sophie hammered the metal flat."

Consider the case of the copular clause (7.1), whose (simplified) structure is given in figure 7.1 French *adj* has a single φ-set, meaning it can only label if it is strengthened by agreement.

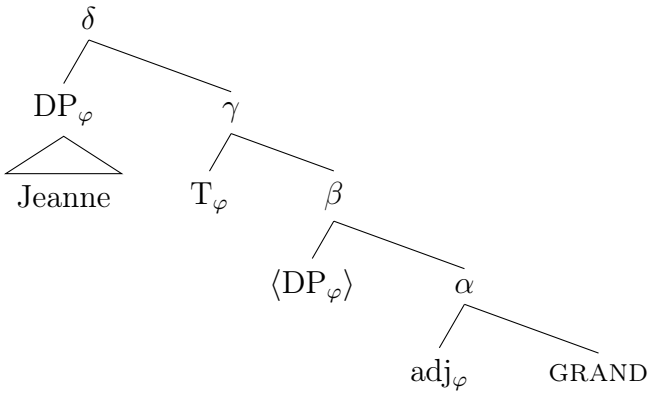


Figure 7.1: Simplified structure of a French copular clause

In the case of resultatives, DP movement out of a small clause bleeds agreement with *adj*. If this were also true of figure 7.1, then α and β would be unlabellable. The same reasoning apply to depictives as well.

This line of thinking is based on the implicit premise that lower copies are invisible to Agree as well as to Label. We could permit (7.1) by hypothesizing that lower copies are visible to Agree but invisible to Label. However, such a move would then predict that French generates resultatives, clearly an unwanted result. I will argue that lower copies are invisible to Agree, but only under certain circumstances. The next section will present a hypothesis regarding the nature of these "certain circumstances" by clarifying the Agree operation.

7.2 The nature of Agree

As stated in chapter 2, I assume that syntactic agreement occurs outside of the Narrow Syntax. Unsurprisingly, this assumption requires additional refinement. If we take Agree to be an operation, we can ask where it fits in the grammatical architecture. By hypothesis, it operates on the output of the Narrow Syntax, and it must operate before labelling. Furthermore, its effects are phonetically overt. These considerations suggest that Agree is part of Transfer, that is, it operates on derived syntactic objects before they are sent to the interfaces. This is represented in figure 7.2.[1]

[1]Despite appearances, this model of the interfaces does not reintroduce grammar-internal levels of representation like S-structure. Agree and Label are computational procedures and therefore each has a specified domain and range, but there are no grammatical constraints or conditions like the EPP or Principle A that hold "at Agree" or "at Label."

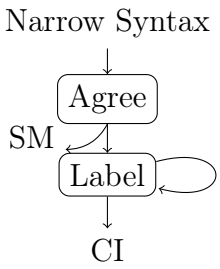


Figure 7.2: The position of Agree in the grammar

This much is an almost unavoidable result of my assumptions and the general observations, but we require further hypotheses to arrive at the predictions we need. The first hypothesis is that Agree feeds some operation that renders lower copies invisible to Label. This could be deletion, impoverishment, or even some kind of cloaking – the details of the operation are not important – but crucially, it is this operation that renders a lower copy invisible to Label. Presumably, the effects of this operation will also be felt at the SM interface. More on that in section 7.4.

If the Agree operation, in a sense, determines whether a given SO is visible to Label, what determines whether an SO is visible to Agree? On the assumption that the input to an Agree cycle is a phase P, I hypothesize that all and only those SO's that are *contained* by P are visible to that cycle. This may seem like a trivial hypothesis, but given the definition of "SO" and "contain" that I will adopt, it makes actual empirical predictions.

My definition will depend on distinguishing a syntactic object and an occurrence of a syntactic object. This type of distinction has been made throughout the development of generative syntax, but perhaps the best-known version is the notion of a chain. In LGB, for instance, each nominal in an S-Structure was associated with a sequence of grammatical functions that represent its derivational history. This sequence was called a function chain. So, in the passive S-structure (7.4), *Jennifer* is associated with the chain (7.5).

(7.4) Jennifer$_i$ was served t_i.

(7.5) $\langle$[NP,S] , [NP,VP]$\rangle$

There is a sense in which the chain was the real grammatical object in LGB and later theories, as filters like the θ-criterion and the Case filter were satisfied by chains rather than by their individual links. So, *Jennifer* in (7.4) has a θ-role because a link in its chain ([NP, VP]) has a θ-role. Following Collins and Stabler (2016), I replace the terms "chain" and "link" with syntactic objects and occurrences defined below.

Definition 7.1 (Syntactic Object)**.** X is a *syntactic object* (SO) iff
X is a lexical item, or
X is a set of syntactic objects. (Modified from Collins and Stabler 2016)

Definition 7.2 (Position)**.** The *position* of $\mathrm{SO_n}$ in $\mathrm{SO_1}$ is a path, a sequence of syntactic objects $\langle \mathrm{SO_1}, \mathrm{SO_2}, \ldots, \mathrm{SO}_n \rangle$ where for all $0 < i < n$, $\mathrm{SO}_{i+1} \in \mathrm{SO}_i$. (Collins and Stabler 2016)

Definition 7.3 (Occurrence)**.** B *occurs* in A at position P iff $\mathrm{P} = \langle \mathrm{A}, \ldots, \mathrm{B} \rangle$. We also say B has an occurrence in A at position P (written $\mathrm{B_{P)}}$.

Consider the abstract syntactic object and its tree representation below in (7.6) and figure 7.3, respectively.

(7.6) $\{\mathrm{X}, \{\mathrm{Y}\,\{\mathrm{X}, \mathrm{Z}\}\}\}$

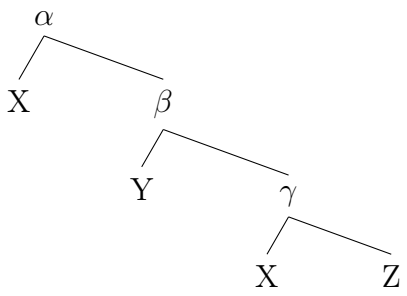


Figure 7.3: A Tree representation of (7.6)

Based on the definitions above, we can say the following things about (7.6): There are six SOs represented in (7.6): three lexical items (X, Y, Z) and three sets (α, β, γ). There is a single SO, X, with two occurrences in (7.6): at $\langle \alpha, \mathrm{X} \rangle$, and at $\langle \alpha, \beta, \gamma, \mathrm{X} \rangle$

With this contrast between SOs and occurrences, we can limit the domain of Agree to complete chains without stipulating the existence of chains. Consider the structure in (7.6), assuming that Y is a phase head, meaning its complement γ has been transferred, rendering it inert. Assume that the computation must track two sets of SOs: the set of SOs in the derivation ($\textsc{Terms}_{\mathrm{SO}}$), which includes the active and inert objects, and the set of active SOs ($\textsc{Active}_{\mathrm{SO}}$), which excludes the inert objects. For (7.6), the two sets are given in (7.7).

(7.7) a. $\textsc{Terms}_{\alpha} = \{\mathrm{X}, \mathrm{Y}, \mathrm{Z}, \alpha, \beta, \gamma\}$
b. $\textsc{Active}_{\alpha} = \{\mathrm{X}, \mathrm{Y}, \alpha, \beta\}$

Since Agree operates on those objects which have been transferred, and, therefore, rendered inert for the purposes of further computation, we can determine the input to Agree, then, by computing the set difference between the two sets in (7.7) as shown in (7.8).

(7.8) $\textsc{Terms}_{\alpha} \setminus \textsc{Active}_{\alpha} = \{\mathrm{Z}, \gamma\}$

This derived set of SOs, I assume, is the input to Agree. Note that X, which has moved to [Spec, Y], is not a member of the input to Agree, despite the fact that there is a member

of the input which has X as a member. As such, X is invisible to Agree and Label. With that understood in the abstract, we can consider the concrete cases of copular clauses and resultatives.

7.3 Correct predictions: French small clauses saved

To begin, let's compare the two relevant structures: the copular clause in figure 7.4, and the resultative adjunct in figure 7.5.

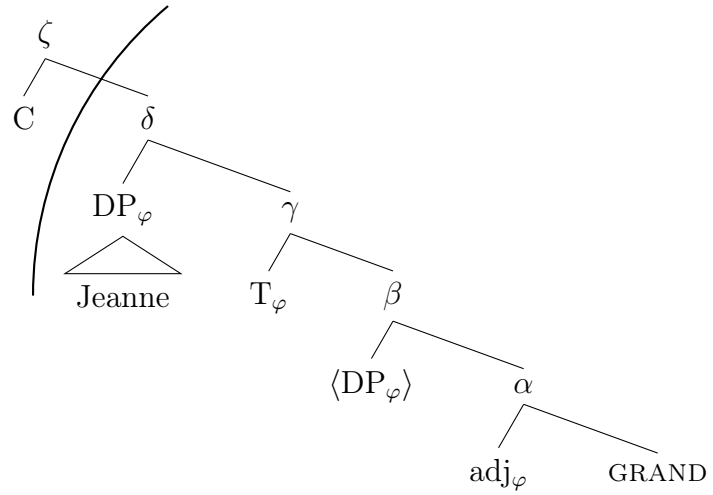


Figure 7.4: An unlabelled French copular clause

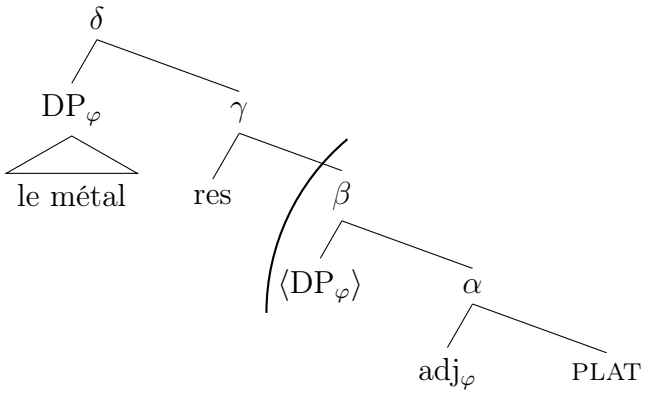


Figure 7.5: An unlabelled French resP (which will crash)

The most salient difference between the DP "chains" in figure 7.4 and figure 7.5 is that the "chain" in figure 7.5 crosses a phase boundary, while the one in figure 7.4 does not. This fact is relevant for Agree's visibility conditions. In figure 7.4, both occurrences of the DP are contained within the phase, meaning the entire syntactic object DP is contained within the phase and therefore is visible to Agree. So, Agree takes δ, which contains a full DP chain, and values φ-features on T and adj with φ-features of DP. This has two relevant effects: first,

it strengthens T and adj so that they can label, and second, it renders the lower copy of DP inactive/invisible for Label. Label then operates on the output of Agree and successfully sends a labelled phrase marker to CI. Thus, the derivation of a copular clause converges in French.

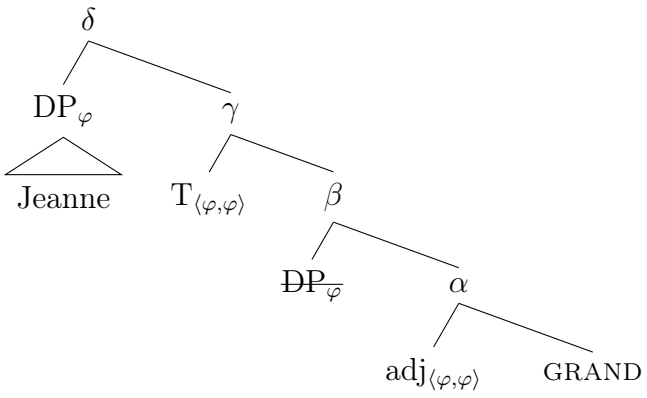


Figure 7.6: The output of Agree for a copular clause

(7.9) The result of Label for figure 7.6

a. Label(δ) = $\langle\varphi,\varphi\rangle$

b. Label(γ) = T

c. Label(β) = Label(α) = adj

Next, consider the resultative adjunct in figure 7.5, which does not converge in French. We start with a small clause which we merge with res, a phase head, forming γ. We then merge the DP with γ, and commence our phase operations on β. The phase complement, β, unlike that of the copular clause in figure 7.4, contains only one occurrence of the DP. Since Agree operates only on SOs, the DP, *le métal*, is invisible to it. There is, therefore, no feature transfer between D and adj, and the DP, therefore cannot be "deleted" yet.[2] The output of Agree, then, is passed to Label, which fails to produce a labelled structure for CI. Specifically, the adjective α cannot be labelled because its would-be labeller adj_φ

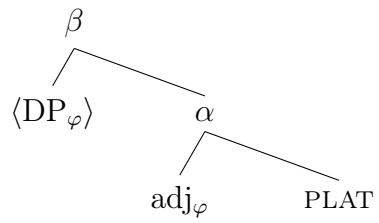


Figure 7.7: The output of Agree for a resP adjunct

has not been strengthened to provide a label. Furthermore, the small clause β cannot be

[2]Recall from section 7.2 that, along with feature transfer, one of the things that Agree does is render lower copies invisible to Label. It follows from this that only those objects which are visible to Agree can be rendered invisible to Label.

labelled, as it is a phrase-phrase structure which cannot be labelled under either of its available labelling strategies. Since DP and α don't agree, β cannot receive a feature-pair label. Even if the DP were somehow rendered invisible by Agree, then the label would be the most prominent element in α, but, as I just mentioned, that would be the too-weak adj_{φ}. Therefore, the derivation will crash. To sum up, if we separate Agree from Label, we can fix the apparent under-generation, provided that Agree operates on full syntactic objects, rather than occurrences.

Turning to depictives, we can see how they would be allowed in French, given our assumptions. Taking resultatives and depictives to be minimally different, we can begin to investigate the source of the contrast in grammaticality between the two. Both are secondary predication constructions, consisting of an eventive VP and a stative small clause. The difference between the two is the semantic relation between the event and the state. Roughly speaking, resultatives describe an event causing a state, while depictives describe an event coinciding with a state. I have chosen, following Kratzer (2005) and Pietroski (2005), to assume a *res* head that encodes causation, but I see no compelling reason to assume a *dep* head to encode coincidence, though most syntacticians assume a *dep* head following Pylkkänen (2008). Pylkkänen, however, does not argue for the presence of a *dep* head but posits the head based on the assumption that it is necessary to encode depictive semantics. I, on the other hand, will assume that no such *dep* head is required.[3] A depictive VP thus has the structure represented in figure 7.8.

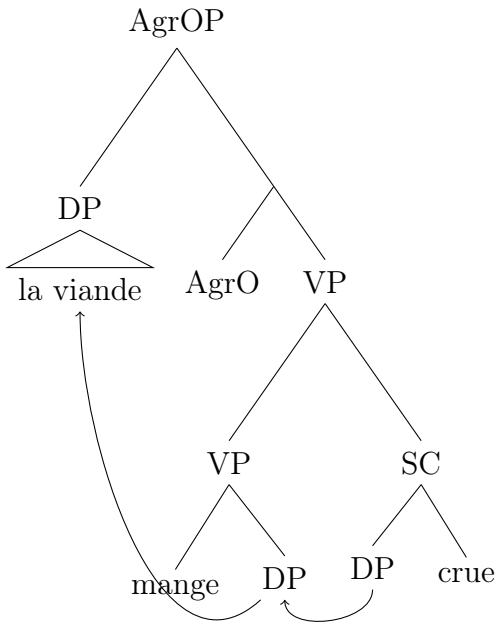


Figure 7.8: A depictive VP

[3]I will postpone the discussion of the coincidence interpretation until chapter 9.

Note that, as in copular clauses, and unlike resultatives, the DP movement "chain" does not cross a phase boundary. Therefore, the same reasoning that I applied to the copular clause, can be applied to the depictive case. In other words, depictives are allowed in French for the same reason that copular clauses are.

7.4 On *wanna*-contraction

The proposal that Agree operates on syntactic objects, as opposed to occurrences, gains support when we consider a fact about A-bar traces. As has been noted by several authors (Hornstein 1999; Jaeggli 1980; Lightfoot 1976), A-bar traces block *wanna*-contraction.

(7.10) a. Who$_i$ do you want to visit t_i? → Who do you wanna visit?
b. Who$_i$ do you want t_i to visit Emma? → *Who do you wanna visit Emma?

The derivation of (7.10-b) involves movement of *who* across a phase boundary, creating a chain that is invisible to Agree. Consider the structure of (7.10-b) in (7.11).

(7.11) $[_\gamma$ Who$_i$ $[_\beta$ do$_C$ $[_\alpha$ you want $\lfloor$who$_i\rfloor$ to visit Emma]]]?

Upon γ being formed, phase operations are performed on α. When Agree operates on α, only the tail of the A-bar chain $\langle$Who$_i$, $\lfloor$who$_i\rfloor\rangle$ is available, meaning it is invisible to Agree.[4] Since Agree, in addition to valuing features, also deletes copies, $\lfloor$who$_i\rfloor$ will remain in α when it is spelled out, until the rest of γ is spelled out. Assuming morphophonological processes operate on the output of Agree, the input to the contraction process will be the string/structure in (7.12).

(7.12) you want who to visit Emma.

And assuming adjacency is a precondition for contraction, we wouldn't expect contraction to occur in (7.12). When the next phase, which includes C$_Q$ and the higher copy of *who*, undergoes Agree, then the *who* chain can be reduced—its lower copy is deleted. Since the

[4]The addition of A-bar chains to my analysis of Agree may conflict with the presence of French *wh*-questions in copular clauses and depictives as in (i), where *wh*-movement does not bleed predicative agreement.

(i) a. Quelle femme est grand -e?
Which.FSg woman is tall -FSg
"Which woman is tall?"
b. Quel genre de viande est-ce que Jeanne a mangé frit -e?
Which.MSg type of meat(Fem) is it that Jeanne has eaten fried -FSg
"Which type of meat did Jeanne eat fried?"

I set this possible contradiction aside for further development of the theory of Agree to resolve.

lower phase has already undergone linearization, etc., *want to* cannot be contracted, even though the intervening material is gone. Thus, we get the string in (7.10-b).

In the underlying structure of (7.10-a), given in (7.13), a trace of *you* intervenes between *want* and *to*.[5]

(7.13) $[_{\gamma}$ Who$_i$ $[_{\beta}$ do$_C$ $[_{\alpha}$ you$_j$ want $\lfloor$you$_j\rfloor$ to visit $\lfloor$who$_i\rfloor]]]$

Upon γ being formed, phase operations are performed on α. When Agree operates on α, only the tail of the A-bar chain $\langle$Who$_i$, $\lfloor$who$_i\rfloor\rangle$ is available, meaning it is invisible to Agree. However, the entire A chain$\langle$you$_j$, $\lfloor$you$_j\rfloor\rangle$ is available, and Agree deletes the lower copy $\lfloor$you$_j\rfloor$, which intervenes between *want* and *to*. Thus, the input to the contraction is the string/structure in (7.14).

(7.14) you want to visit who

In this case, *want* and *to* are adjacent (or at least, no phonologically overt material intervenes between them), meaning contraction can occur.

7.5 Summary

The analysis of the resultative parameter developed in part I seems to under-generate for languages that lack resultatives. Specifically, it seems to predict that languages like French should have no copular clauses and no depictives. In this section, however, I have shown that the apparent under-generation was due to the fact that the grammatical architecture was not explicitly described. In particular, once the nature of the Agree operation and its position in the language faculty was made explicit, it became clear that a grammar that rules out resultatives need not rule out copular clauses and depictives.

In this architecture, Agree is taken to be part of Transfer, operating after the Narrow Syntax and before Label. The proposition that Agree is post-syntactic is assumed, but the hypothesis that it is part of Transfer follows from the fact that the effects of Agree are seen at both interfaces. I further hypothesized that Agree operates only on complete syntactic objects, as opposed to occurrences of syntactic objects. This means that if an object has moved across a phase boundary (as is the case for resultatives) then it will be invisible to Agree, and if agreement is required for labelling, then such a movement operation will bleed labelling. Since the movements required for copular clauses and depictives do not cross phase boundaries, they do not bleed labelling, and therefore do not crash the derivations. To further justify my hypotheses, I showed that this conception of Agree can be used to give a straightforward account of *wanna*-contraction.

[5]Control is movement. See section 2.1. See also Hornstein (1999)

Chapter 8

Movement from Specifier of resP

Another issue with my account has to do with the sideward movement operation. Recall that in order to derive an adjectival resultative, a DP must move sideward from the resP adjunct to the VP as in figure 6.1. Note also that that sideward movement operation seems to be

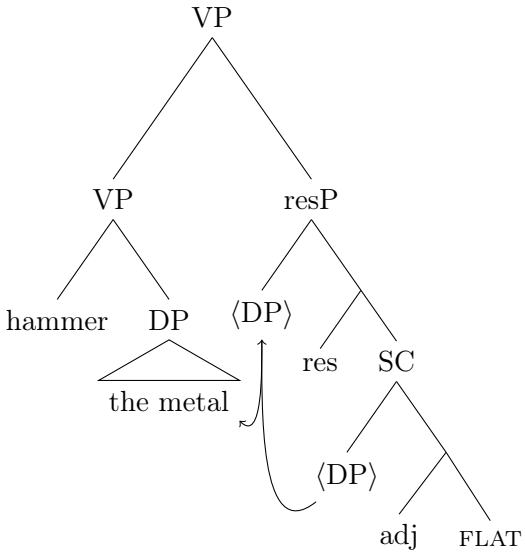


Figure 6.1: The structure of a resultative (repeated from page 60)

obligatory; that is, the DP that originates in the resP must also appear as the theme of the VP. This obligatoriness can be seen in the fact that (8.1) is ungrammatical.

(8.1) *Sam [$_{VP}$ hammered the nail] [$_{resP}$ the planks together].
(≈ Sam hammered the nail and, as a result, the planks were fastened together)

An easy way of accounting for this would be to hypothesize that it is due to some property of the res head, and if this obligatory sideward movement were particular to resultatives, then, indeed, this would likely be the best way to proceed. However, sideward movement seems to be obligatory in other cases. First, there is the case of depictives, which differ

from resultatives only in the fact that they lack a res head. Also, as I will argue in this chapter, there is obligatory sideward movement in certain so-called ACC-ing clauses in direct perception reports such as the embedded clause in (8.2).

(8.2) We heard [them shouting at the top of their lungs].

Furthermore, I will argue that this obligatory sideward movement is, in fact, a property of adjoined phrases. Specifically, the generalization in (8.3) seems to hold.

(8.3) Internally merged specifiers of adjoined phrases must move to the host phrase.

Such a generalization, I argue in section 8.2, cannot be accounted for in a theory of grammar based on feature satisfaction. Label theory, however, is able in principle to derive this generalization, but in order to do so, it must be modified and extended. I will perform such an extension and show that the resulting modified label theory can derive (8.3).

8.1 On ACC-ing clauses

Cinque (1996) discusses ACC-ing clauses (ACs) under direct perception verbs as in (8.4) and argues that they are ambiguous, having the two structures in (8.5).[1]

(8.4) I saw Mario running at full speed.

(8.5) a. I [saw [$_{\text{ProgP}}$ Mario [$_{\text{Prog}'}$ -ing [$_{\text{VP}}$ run ...]]]].
b. I [[saw Mario] [$_{\text{ProgP}}$ *ec* running ...]].

I argue in section 8.2, the fact that a single grammar can generate both of these structures presents a serious problem for standard theories of grammar. I will, therefore, discuss them

[1]The main object of Cinque's study, in fact, is pseudo-relatives such as those in (i), which he argues are ambiguous between the three structures in (ii).

(i) a. Ho visto Mario che correva a tutta velocità. (Italian)
b. J'ai vu Mario qui courrait á tout vitesse. (French)

(ii) a. Ho [visto [$_{\text{NP}}$ Mario [$_{\text{CP}}$ che correva ...]]]
b. Ho [visto [$_{\text{CP}}$ Mario [$_{\text{C}'}$ che [$_{\text{IP}}$ correva ...]]]]
c. Ho [[visto Mario] [$_{\text{CP}}$ *ec* che correva ...]]

He mentions ACs briefly in order to point out that his remarks and claims about pseudo-relatives largely apply to ACs. The main difference between the two constructions is that ACs are not analyzable as nominals, but a related form with nominal morphology serves this function.

(iii) [Their singing of the national anthem] caused an international incident.

in greater detail in the remainder of this section.

In one structure, represented in figure 8.1, the AC is merged as the complement of the perception verb. The interpretation of this structure is one in which the running event was seen and by virtue of the meaning of *run*, seeing a running event generally entails seeing the agent of that event. In the second structure, represented in figure 8.2, the ACC-ing subject is

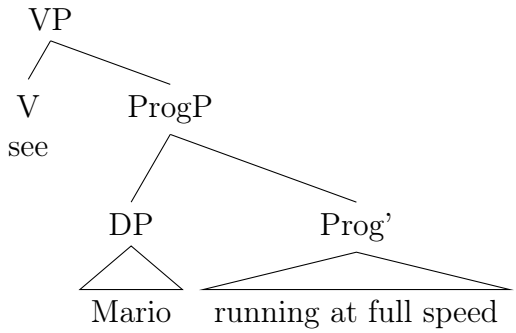


Figure 8.1: Complement ACC-ing structure

merged as the complement of the perception verb, while the AC (with a controlled subject) is adjoined to the VP. The interpretation of this structure is one in which *Mario* is seen and the event of *Mario* being seen coincides with an event of *Mario* running. Again in this interpretation, due to the meaning of *run*, seeing the agent of a running event generally entails seeing the event itself. By the assumptions made here, the argument *Mario* can only

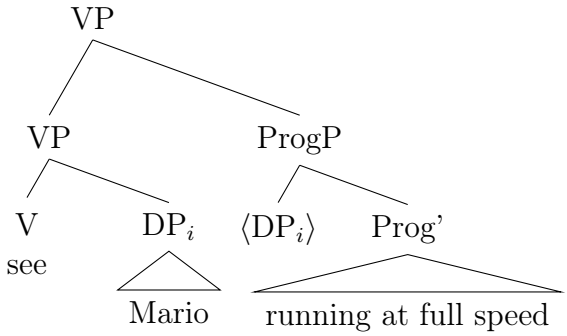


Figure 8.2: Adjunct ACC-ing structure

be shared by the verb *see* and the verb *run* if it is merged with both, meaning it must move from [Spec, Prog] to [Comp, V].[2] In the case of the complement AC in figure 8.1, however, the argument *Mario* seems to stay in situ in [Spec, Prog], suggesting that the movement operation represented in figure 8.2 is, in fact, optional. If the movement operation is optional, however, we would expect two additional structures for (8.4): one, represented in figure 8.3, in which the ProgP is the complement of *saw* and *Mario* has moved from [Spec, Prog], and another, represented in figure 8.4, in which the ProgP is an adjunct, but *Mario* does not move from [Spec, Prog]. If we consider the consequences of this proposed

[2]Cinque assumes that [Spec, Prog] is occupied by a controlled PRO.

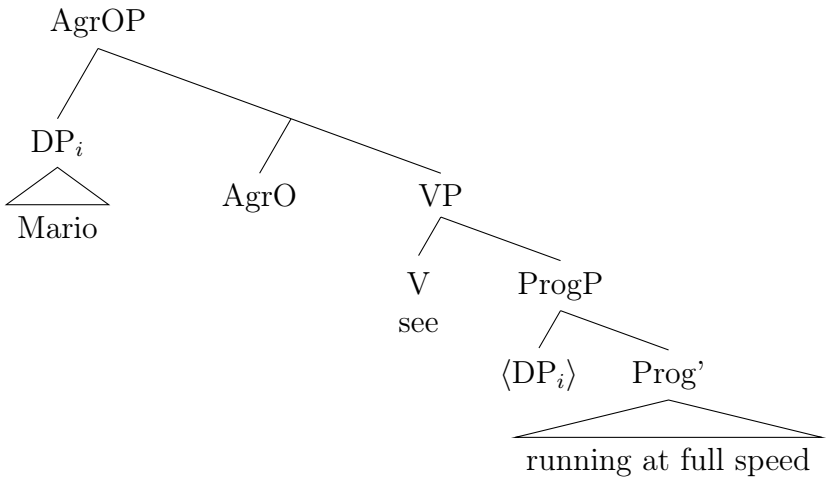


Figure 8.3: Complement ACC-ing structure with object raising

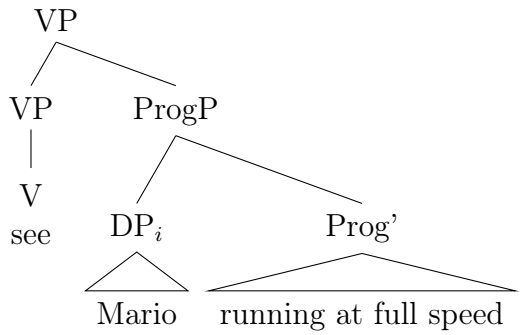


Figure 8.4: Adjunct ACC-ing structure without DP movement

optionality hypothesis, we can see that it cannot be true.

First, consider the structure in figure 8.4, in particular, the fact that *see* does not have an internal argument. This is not *per se* problematic, as verbs may be optionally transitive, but we would expect that *see* in this structure could have an internal argument other than *Mario*. That is, if figure 8.4 is a possible structure for (8.4), then we would expect (8.6) to also be a licit sentence.

(8.6) *I [$_{\text{VP}}$ [$_{\text{VP}}$ saw Sue] [$_{\text{ProgP}}$ Mario running at full speed]].

If *Mario* can remain in [Spec, Prog], then we have no way to rule out *Sue* merging with *see* and behaving as a direct object. Of course, (8.6) is ungrammatical, suggesting that *Mario* cannot remain in [Spec, Prog] if ProgP is adjoined to VP. Movement from [Spec, Prog], then, cannot be optional, strictly speaking. If it is not optional, perhaps it is obligatory.

If movement from [Spec, Prog] is obligatory, then we must revise Cinque's analysis of complement ACs. Suppose, then, that *Mario*, in the complement ACC-ing analysis of (8.4), must raise to object. In other words, suppose figure 8.3 is a possible structure of (8.4) and figure 8.1 is not. Note that in figure 8.3, *Mario* is the grammatical object but not the theme

of *see*. If this is the case then we expect that *Mario* can become the subject of a passive derived from figure 8.3.

Indeed, subjects of ACs can become passive subjects as in (8.7), but it is not immediately obvious whether (8.7) is derived from an Adjunct AC structure or a Complement AC structure.

(8.7) Mario was seen running at full speed.

If (8.7) had been derived from a Complement ACC-ing structure, however, then *Mario* would not have been θ-marked by *see*. So, in order to test whether passives like (8.7) can be generated in which the subject is not interpreted as the theme of the verb, that is, we need a clause of the form in (8.8) where the event of V-ing was perceived without the individual X being perceived.

(8.8) X was {seen/heard/felt} V-ing ...

There are certain classes of predicates which we can use as diagnostics due to a non-canonical event/argument structure. Consider the ACC-ing versions of weather reports and clausal idioms, for instance, as given in (8.9) and (8.10)

(8.9) Bill saw it snowing. $\not\Rightarrow$ Bill saw it.

(8.10) Bill heard all hell breaking loose. $\not\Rightarrow$ Bill heard all hell.

Consider, also, the predicates *be slandered* and *be parodied*. Events of parodying or slandering an individual x generally do not include x as a participant the way, for instance, events of hitting x or speaking to x do. This is demonstrated in (8.11) and (8.12)

(8.11) We heard the writer being slandered. $\not\Rightarrow$ We heard the writer.

(8.12) They saw the singer being parodied. $\not\Rightarrow$ They saw the singer.

Unlike most direct perception reports with ACs, then, (8.9) to (8.12) are not ambiguous. Rather, they have only the complement AC structures. Therefore, if the perception reports in (8.9) to (8.12) can be passivized, this will be evidence for the proposal that DPs are able to move out of complement ACs. In fact, it seems that they cannot be passivized.

(8.13) *It was seen raining.

(8.14) *All hell was heard breaking loose.

(8.15) *The writer was heard being slandered.

(8.16) *The singer was seen being parodied.

I can think of no principled explanation of these facts except to propose that DP movement out of a complement AC is barred. Absent any evidence or argument to the contrary, then, I will assume that Cinque's initial analysis of complement ACs was correct.

Thus we are led to the following generalization with respect to ACs: If an AC is adjoined to a VP, then its subject must move, but if an AC is merged as the complement of a verb, then its subject cannot move. This is an unexpected, perhaps unprecedented syntactic generalization. In fact, I will argue in the following section that such a pattern is predicted to be impossible by any theory that defines grammaticality solely in terms of feature satisfaction, as standard minimalist theories do.

8.2 Feature satisfaction cannot account for ACC-ing clauses

As I discussed in chapter 2, standard minimalist theories tend to assume that a derived syntactic object converges at the interfaces iff it contains no unsatisfied features. There is, of course, debate as to what it means for a feature to be unsatisfied, and what sort of operations are able to satisfy these features. Therefore, I will attempt to abstract away from the details of particular theories and discuss what I take to be their shared assumptions.

The first common assumption (or set of assumptions) is about features. Lexical items bear or consist of features, each of which is inherently either satisfied or unsatisfied,[3] and each of which is also specified for the information it encodes (person, gender, tense, etc.). For instance, a determiner may bear a satisfied definiteness feature and an unsatisfied Case feature.

The second common assumption is that some computational operation converts a token of some unsatisfied feature into a token of the corresponding satisfied feature under the influence of some other feature token. In order for feature token F on lexical item token X to satisfy feature token G on lexical item token Y, X and Y must stand in some structural relation to each other, and F and G must be of the same type. Furthermore, feature satisfaction is automatic; if the conditions are met for F to satisfy G, then F satisfies G.

The third common assumption is that there is no operation that undoes feature satisfaction. This is never made explicit, but it is nonetheless assumed to be true. Such an assumption, in effect, ensures a certain monotonicity in syntactic derivations, in that if at some derivational stage S_n feature token F is satisfied, then there is no later stage $S_{n+i} (i > 0)$ at which F is unsatisfied.

The fourth common assumption, which I have already alluded to, is that a syntactic object is well-formed iff it contains no lexical items with any unsatisfied features. This will be used as a diagnostic for satisfied features. If a sentence, phrase, or word is well-formed,

[3]The two most common versions of unsatisfied features and satisfaction operations are *uninterpretable* features which must be checked, and *unvalued* features which must be valued.

then it must contain no unsatisfied features, and if an expression is ill-formed, then it must contain unsatisfied features.

With these assumptions in place, consider the complement AC case in (8.17), which is well-formed; this means that all of its sub-parts are well-formed.

(8.17) They [$_{\text{VP}}$ saw [$_{\text{ProgP}}$ the raccoon [walking across the street]]].

Since the ProgP is well-formed, it must contain no unsatisfied features, and, therefore, the DP *the raccoon* must contain no unsatisfied features. Assuming that *the raccoon* is in [Spec, Prog], rather than its base position in [Spec, v/Voice], then at least one of its unsatisfied features can only be satisfied in [Spec, Prog].

(8.18) a. *They [$_{\text{VP}}$ saw [$_{\text{ProgP}}$ walking [$_{\text{DP}}$ the raccoon] across the street]].
b. *They [$_{\text{VP}}$ saw [$_{\text{ProgP}}$ [-ing [[$_{\text{DP}}$ the raccoon] walk across the street]]]].

In other words, *the raccoon* is licensed in [Spec, Prog].

However, when we consider the adjunct AC in (8.19), we come very quickly to a contradiction.

(8.19) They [$_{\text{VP}}$ [$_{\text{VP}}$ saw the raccoon][$_{\text{ProgP}}$ t walking across the street]].

The fact that a lower copy/trace occupies [Spec, Prog] indicates that *the raccoon* is not licensed there. In fact, the data adduced above in (8.6) and the discussion thereof indicates this fact even more forcefully. If *the raccoon* is not licensed in [Spec, Prog], this means it bears some feature which cannot be satisfied there. This is a direct contradiction of the conclusion I came to above, and yet this contradiction arises from an analysis of facts based on an axiom set. The conclusion I draw from this contradiction is that the axiom set (a.k.a. the feature-satisfaction theory of syntax) is fundamentally flawed.

In order to pinpoint this flaw, we should consider how the facts could be generalized. I believe a proper expression of the generalization is given in (8.20) and (8.21).

(8.20) A syntactic object X (= $\{t, \{\text{Prog, YP}\}\}$) is well-formed only if X is an adjunct.

(8.21) A syntactic object X (= $\{\text{DP}, \{\text{Prog, YP}\}\}$) is well-formed only if X is an argument.

This is a problem for the feature-satisfaction theory because implicit in the theory is the claim that the well-formedness of an object depends solely on its internal structure. The generalizations here, however, make reference, not just to the internal structure of an object, but also to the larger structure that the object is a part of. This focus on internal structure is to be expected if the grammaticality of an expression is solely determined by whether the narrow syntax operates properly. That is, if a given operation in the NS must be justified

locally, it follows that it cannot be justified on the basis of an operation which has yet to occur.

However, A theory which bases grammaticality, at least partially, on interface conditions, as label theory does, will be able to account for the generalizations in (8.20) and (8.21). Specifically, label theory must treat adjuncts and arguments differently and therefore, provides a good candidate for an explanatory theory of the generalization in question. Since host-adjunct structures are always a species of phrase-phrase structures,[4] we would expect a labelling paradox as with other phrase-phrase structures. Unlike other phrase-phrase structures, however, there does not seem to be an agreement-based "repair strategy" for host-adjunct structures. Almost by definition, adjuncts neither move nor agree. Adjuncts, it seems, simply do not enter into the labelling calculation. However, label theory as formulated in Chomsky 2013 and Chomsky 2015 does not address host-adjunct structures at all and, therefore, is not quite suitable for my purposes. Thus I will modify it in the following section.

8.3 Modifications to label theory

In this section, I will address two questions that Chomsky (2013, 2015) largely leaves open. First, there is the question of how to label Host-Adjunct structures, which I address in section 8.3.1. Second, there is the question of why labels are required at all. I address this question in section 8.3.2.

8.3.1 Labelling Host-Adjunct Structures

To understand how adjuncts behave with respect to labelling, let's consider their basic properties: optionality, iterativity, and freedom of order. These can be demonstrated in the series of sentences in (8.22).

(8.22) a. The protester was brought to the police station.
b. The protester was brought to the police station, against her will.
c. The protester was brought to the police station, against her will, after the demonstration.
d. The protester was brought to the police station after the demonstration, against her will.

If we assume that the adjuncts in (8.22) are adjoined to TP, then the TPs in each of the sentences in (8.22), are, in some sense, grammatically indistinguishable. If we take this much

[4]The phrasal nature of clausal and PP adjuncts, for instance, is uncontroversial, but adjectives and adverbs are also phrasal given the theory of categories assumed here. An adverb, for instance, consists minimally of a root and an *adv* head.

for granted, then we can view the task of developing a theory of adjuncts to be the task of making explicit the sense in which the sentences in (8.22) are indistinguishable. Assuming label theory, we can make a fairly trivial explication of the indistinguishability of these sentences: the TPs in (8.22) are indistinguishable in the sense that they are labelled identically. If the sentences in (8.22) are constructed purely by Merge (and Select, and Copy), then they have the structures in (8.23).

(8.23) a. $[_{\alpha}$ The protester was brought to the police station].
b. $[_{\beta}[_{\alpha}$The protester was brought to the police station], [against her will]].
c. $[_{\gamma}[_{\beta}[_{\alpha}$The protester was brought to the police station], [against her will]], [following the demonstration]].
d. $[_{\eta}[_{\delta}[_{\alpha}$The protester was brought to the police station] [following the demonstration]], [against her will]].

If we take α to be the TP without adjuncts, then its label will be the basis for the label of the modified TPs β, γ, δ and η. Since α is a finite TP with a subject, its label will be $\langle\varphi, \varphi\rangle$, and by assumption, the label of each of the modified TPs will be $\langle\varphi, \varphi\rangle$. If this is the case, then the adjoined phrases contribute nothing to the labelling algorithm; in other words, they are invisible to LA. The invisibility of adjuncts cannot, however, be the same phenomenon as the invisibility of lower copies, as the latter arises from a movement operation, and there is no reason to think that adjunct phrases as a class undergo movement. Furthermore, even if adjuncts did move, this would only explain why lower copies are invisible; we would still need to explain why higher copies are invisible.

If the invisibility of adjuncts cannot be derived syntactically, perhaps it is inherent. That is, perhaps the set of adjuncts is a natural class of objects which are invisible to LA. This suggestion, however, runs into problems almost immediately due to the fact that there are phrases which can be arguments, predicates, or adjuncts as in (8.24) to (8.25).

(8.24) a. The green room. (*green* as an adjunct)
b. The room is green. (*green* as a predicate)

(8.25) a. Meryl swam [in the pool]. (PP adjunct)
b. Cameron fell [in the pool]. (PP complement)

It seems that, without a significant amount of stipulation, this is not a promising approach so I will not pursue it further.

Adjunction, then, cannot be reduced to simplest Merge. This leaves two broad options for assimilating it into our theory. The first is to propose a new operation in Narrow Syntax (NS) that generates adjunction structures. Chomsky (2004) proposes an operation of pair-Merge, which, given a host object β and an adjunct α, creates the object $\langle\alpha, \beta\rangle$ (α adjoined to β).

The new object $\langle \alpha, \beta \rangle$, however, has all of the syntactic properties (c-command relations, θ-roles, selectional properties, etc) of the previously generated object β. So, as far as NS is concerned, $\langle \alpha, \beta \rangle$ is equivalent to β. For preciseness, I will use the term *σ-equivalent*: an object $\langle \alpha, \beta \rangle$ (created by pair-Merge) is σ-equivalent to β. There are two problems with this proposal, which I address in turn in the following paragraphs.

The first problem with adding an operation of pair-Merge to the system arises from the question of whether that change violates SMT. Recall that SMT states that the language faculty is an optimal solution to interface problems, meaning that a minimalist theory of grammar should only admit complications if they are required due to interface conditions. So, is pair-Merge required by one of the interfaces? The fact that the information expressed by pair-Merge can be expressed periphrastically, as shown in (8.26) suggests that pair-Merge is extraneous.

(8.26) a. I'd like a large burger with ketchup.
b. I'd like a burger. I'd like it to be large. I'd like it to have ketchup.

The series of sentences in (8.26-b) express the same proposition as the single sentence in (8.26-a) and they would do so without any instances of pair-Merge. The same cannot be said about structures formed by set-Merge; an expression constructed by set-Merge cannot be paraphrased without set-Merge. Since it is not required by the interfaces, the addition of pair-Merge to NS would constitute a violation of SMT.

The second problem with pair-Merge arises from concerns about economy of derivation. There are two facts about pair-Merge that are relevant to this issue. First, pair-Merge is a more complex operation than set-Merge, as the former induces order, while the latter does not. And second, when we adjoin an object to a host by pair-Merge, the resulting object is σ-equivalent to the host without the adjoined object, in the sense that a noun phrase with an adjective adjoined to it has all of the same syntactic properties as that same noun phrase without any adjunct. Consider the two sub-derivations in (8.27) and (8.28). The results of the two derivations are σ-equivalent to each other, but the first derivation is more complex than the second one.

(8.27) $\text{pMerge}(\text{X}, \text{Y}) = \langle \text{X}, \text{Y} \rangle$
$\text{Merge}(\text{Z}, \langle \text{X}, \text{Y} \rangle) = \{\text{Z}, \langle \text{X}, \text{Y} \rangle\}$

(8.28) $\text{Merge}(\text{Z}, \text{Y}) = \{\text{Z}, \text{Y}\}$

From the view of NS, then, pair-Merge does a lot of work to no effect; the object derived in (8.27) with an adjunct is syntactically indistinguishable from the object derived in (8.28) without an adjunct. Therefore, for every derivation D that uses pair-Merge, there is a simpler derivation D′ such that the result of D is σ-equivalent to that of D′. This is exactly the type

of situation that derivational economy rules out.

So, if, as I argue above, adjunction does not occur in NS, then it must occur after NS—the second broad option for incorporating adjunction into our theory. However, if we make the standard minimalist assumption that NS is the only module capable of recursively combining expressions to form larger expressions, then there can be no recursive combinatory operation outside of NS. Therefore, adjunction – being outside of NS – cannot be a recursive combinatory operation. That is, adjunction does not create new syntactic objects. This means that our way of representing adjunction in tree structures is misleading.

Consider, for instance, the modified VoiceP in (8.29) as represented in figure 8.5.

(8.29) Mary sang the song with gusto.

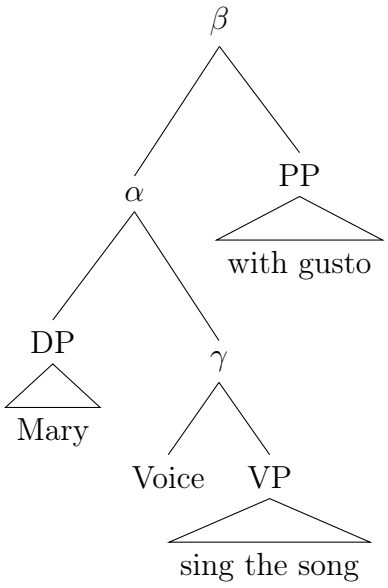


Figure 8.5: A standard representation of a modified VoiceP

The object β is usually taken to be created by adjoining the PP *with gusto* to α, but, as I argued above, adjunction cannot create new objects. Therefore, there is no object β. This is, no doubt, a surprising conclusion, yet it follows from the basic facts of adjunction and SMT, so it behooves us to entertain it as a possibility.

This conclusion, in fact, does resolve the immediate question of how Host-Adjunct structures are labelled, not by answering it but by dissolving it. If LA is a function from unlabelled SOs to labelled SOs and Host-Adjunct structures are not SOs, then they are outside the domain of LA. However, there are a few caveats that bear mentioning. The first is that, while Host-Adjunct structures are not properly SOs, the same cannot be said for the adjuncts *per se*. So, to consider a concrete case, β in figure 8.5 is not an SO, but the PP *with gusto* is an SO, meaning it was derived in NS and labelled by LA. The second caveat is that not everything that looks like a Host-Adjunct structure is one. For instance, the topicalized PP

in (8.30) is likely an argument of, say, a functional projection Topic.

(8.30) With sorrow in her heart, Mary sang the song.

The third caveat, which perhaps is more of a promissory note, is that asserting that Host-Adjunct structures are not SOs leaves us with the question of what they are. This is far from an easy question to answer, and I will not attempt a complete answer here. Instead, I will stipulate that, at the CI interface, a host-adjunct structure is a complex object which is asymmetric and unlabelled. It is asymmetric in the sense that the host is more prominent than its adjuncts. A proper theory of adjunction, if one exists, will derive this asymmetry from intrinsic properties of the host and adjuncts, but I will stipulate it here. It is unlabelled because only SOs are labelled, and host-adjunct structures are not SOs.

The notion that there can be complex linguistic objects which are not generated by Merge may seem to contradict the evolutionary version of SMT, which states that the evolution of the language faculty consists in the sudden appearance of Merge. If adjunction is a non-Merge method for constructing complex linguistic expressions, then we would expect there to be a language faculty even without Merge. While this expectation is not, strictly speaking, borne out, there does seem to be an extra-linguistic cognitive system that makes use of complex language-like representations. Consider the system of propositional attitudes that Fodor (1975) discusses, and the structures employed in the study of discourse pragmatics.

For Fodor, all cognition involves several sets of propositions that a given organism has certain attitudes toward. For instance, every animal has a set of beliefs and a set of desires, which are populated by propositions. The fact that humans have language means that these propositions can be of arbitrary complexity, but they are still beliefs and desires. Also, perhaps the most central notion of discourse pragmatics is the common ground, which is a set of propositions believed to be shared between discourse participants. While the common ground interacts with linguistic expressions, it does not seem to be one itself. These examples are complex cognitive objects that are non-linguistic, and, like host-adjunct structures, they "compose" by conjunction.

While these proposition sets are not usually considered to be compositional, it seems rather obvious that holding an attitude towards a set of propositions is logically equivalent to holding that same attitude towards the conjunction of the proposition in the set.

(8.31) $\textsc{bel}(p) \;\&\; \textsc{bel}(q) \leftrightarrow \textsc{bel}(p \;\&\; q)$

Note, of course, that this is a logical equivalence but not a representational equivalence,[5] and since, according to the computational theory of mind, representations matter, I would not like to claim that holding an attitude towards a set of propositions P necessarily requires

[5]Fodor's (2010, pp. 50–100) discussion of referential opacity, provides, I believe, an excellent argument for distinguishing logical equivalence from representational equivalence.

holding that attitude towards the conjunction of every subset of P. Rather, I claim that if two propositions, p and q, are members of the same set in the mind, then the operation of adding the conjunction of those propositions, p&q is available, but an operation of adding other possible compositions of p and q (p∨q, p→q, etc.) is not available. So, perhaps the process of interpreting a host-adjunct structure involves adding the conjunction it expresses to some proposition set.

To summarize, the above discussion was a long-winded way of hypothesizing that host-adjunct structures not only are not labelled by LA but are not processed by LA.[6]

8.3.2 Why are labels needed at all?

The second question is why labels should be required by the CI interface at all. My proposed answer is that the label of a complex object determines how that object composes semantically. While this may seem *ad hoc*, it is actually a fairly reasonable hypothesis. Consider Chomsky's labelling hypothesis as phrased in (8.32), and the more standard theory of the CI interface in (8.33).

(8.32) A syntactic object is a valid CI object iff it is labellable.

(8.33) A syntactic object is a valid CI object iff it composes semantically.

At first glance, these hypotheses are incompatible, giving us three options for resolving the conflict. The first option would be to reject one of the conflicting hypotheses. There is no strong evidence, however, for rejecting either (8.32) or (8.33) so I will not choose this option. The second option is to conjoin the iff clauses as in (8.34).

(8.34) A syntactic object is a valid CI object iff it is labellable and it composes semantically.

This option is unattractive for reasons of theoretical parsimony so I will not choose it. The third option is to hypothesize that labelling and composition are two sides of the same coin, and, therefore, the conflicting hypotheses are equivalent. We can, then, replace our two conflicting statements with the two compatible statements in (8.35) and (8.36) below.

(8.35) A syntactic object composes iff it is labellable.

(8.36) A syntactic object is a valid CI object iff it composes.

This move is theoretically attractive partially due to the fact that it mirrors the logic of antisymmetry on the SM interface (Kayne 1994). In the case of antisymmetry, Kayne identifies

[6] See also Citko (2005) for a similar analysis.

asymmetric c-command with linear order, and there is no compelling reason to think that the CI interface should be more complex than the SM interface.

So, what would it mean for labelling and composition to be two sides of the same coin? Again, it is helpful to consider the SM interface, where asymmetric c-command and linear order are associated because they are isomorphic. We should expect a similar isomorphism to hold between composition and labels, and, in fact, there seems to be good reason to think that there is such an isomorphism. Consider the main modes of composition generally assumed by semanticists (*e.g.*, by Heim and Kratzer 1998), given schematically in (8.37).

(8.37) a. **Lexical insertion**
$\textsc{sem}(\alpha) = \alpha'$

b. **Function application**
$\textsc{sem}([\alpha, \beta]) = \textsc{sem}(\alpha)(\textsc{sem}(\beta))$

c. **Predicate modification**
$\textsc{sem}([\alpha, \beta]) = \textsc{sem}(\alpha)(x)\& \ \textsc{sem}(\beta)(x)$

d. **Predicate abstraction**
$\textsc{sem}([\alpha, \beta]) = (\mathrm{Op}x)(\textsc{sem}(\beta)(x))$

Each of these modes of composition has a corresponding structure type as identified by the version of label theory developed here. Abstracting away from phrasal idioms, lexical insertion operates on a single syntactic atom, *i.e.*, a head, which label theory necessarily distinguishes from other syntactic objects. Predicate modification is the next most complex operation: it conjoins two (possibly complex) objects without requiring or inducing any ordering of the two, exactly isomorphic with the output of merge: unlabelled and unordered syntactic objects. Function application, likewise, requires two objects, but these objects are ordered. Unlike conjunction structures created by predicate modification, which are commutative ($X\&Y = Y\&X$), the function-argument structures created by function application are inherently asymmetric ($X(Y) \neq Y(X)$). This matches with head-labelled structures, which encode a pair of objects (the contents of the structure) and an ordering statement (the label). Finally, predicate abstraction, which creates structures similar to quantifier structures, requires the content of the two expressions, an ordering between the two, and a variable. Pair-labelled structures provide this information.

First, consider those structures labelled by heads. The classes of structures that get head labels are given in (8.38).

(8.38) **Head-labelled structures**

a. $\{\mathrm{X},\ \textsc{root}\} \xrightarrow{\textit{Label}} [_{\mathrm{X}}\mathrm{X},\ \textsc{root}]$

b. $\{\mathrm{X},\ \mathrm{YP}\} \xrightarrow{\textit{Label}} [_{\mathrm{X}}\mathrm{X},\ \mathrm{YP}]$

c. $\{t_{\mathrm{ZP}}, \{\mathrm{X},\ \mathrm{YP}\}\} \xrightarrow{\textit{Label}} [_{\mathrm{X}}t_{\mathrm{ZP}}, \mathrm{XP}]$

I propose that in these cases, the objects compose by function application, with the label being the function and the non-labelling constituent being the argument. So, for instance, a DP is interpreted as the function D, with NP as an argument.

(8.39) $\textsc{sem}([_{the}\textit{the, ball}]) = \textsc{sem}(\textit{the})(\textsc{sem}(\textit{ball}))$

Next, consider the structures labelled by feature-pairs. These structures tend to be the result of internal Merge, which is generally associated with operator-variable structures.[7] I hypothesize, then, that feature-pair labels signal that a complex object is to be interpreted as an operator-variable structure. For instance, the Wh-question structure in (8.40-a) is interpreted as in (8.40-b).

(8.40) $\textsc{sem}(\left[_{\langle Q,Q\rangle}\, \textit{Who}_Q, [\mathrm{C}_Q + \textit{did}, [\textit{Mary see } t_{\textit{Who}}]]\right]) = (\mathrm{Wh}x)(\textsc{sem}(\textit{Mary saw } x))$

Finally, we come to the case of unlabelled structures, which is identical to the case of Host-Adjunct structures. As I discussed in section 8.3.1, however, Host-Adjunct structures are unlabelled because they are not SOs and the domain of the Labelling Algorithm is restricted to SOs. These structures are given the default interpretation of conjunction, as discussed in section 8.3.1.

I have identified Kayne's (1994) theory of the SM interface as an inspiration for my proposal, and I would like to say a little more about the similarities between his and my hypotheses. Rather than positing an active process, be it simple or intricate, for linearizing a hierarchical structure, Kayne suggests that linear order is the product of a passive isomorphism. Since asymmetric c-command *is* a linear, or total, order, a hierarchical structure can be mapped to a linear string based purely on properties of that structure. The idea that an interface between mental modules should be passive is in keeping with the very idea of modularity. If the SM module and the Narrow Syntax module are truly independent, then we would not expect there to be any specialization of one in order to interact with the other. My proposal for the CI interface is, I believe, a step towards a passive interface. Although evidence for the nature of the CI module is not as readily available as evidence for the nature of the SM module, the working, albeit tacit, assumption seems to be that the CI module deals in representations that are formally very similar to formulas of predicate logics. If we assume a predicate logic with operators[8] ($\forall$, $\exists$, Wh, M,...), functions/predicates ($P, Q, f, g,...$), and variables ($x, y, z,...$), then we can see how there could be an isomorphism between labelled syntactic objects and formulas of this logic.

[7] I use the hedges *tend* and *generally* here to indicate that I was not able to perform an exhaustive enumeration of all pair-labelled structures. I perhaps could have formulated this generalization without hedges, in which case this footnote would be explaining that perhaps I should have hedged the generalization slightly.

[8] Since the term *operator* already has a particular meaning in generative syntax, I will refer to the operators of predicate logic as *l-operators*.

While it seems that SM objects are linear structures (strings), CI objects seem to be more complex, including notions of scope and variable binding. I, therefore, cannot able to give a simple order-theoretic explanation of the CI interface as Kayne (1994) gives for the SM interface. What I can do, however, is give a coarse-grained mapping between labelled SOs and expressions in predicate logic. The class of labelled SOs and that of complex expressions of predicate logic can each be divided into two sub-classes. Labelled SOs can be head-labelled or pair-labelled, while expressions of predicate logic can be function-argument expressions or l-operator expressions. Furthermore, these sub-classes seem to map to each other as shown in (8.41).

(8.41)

Labelled SO	Predicate logic
Head labelled	Function-argument
{in, {in, the snow}}	**in**(**the_snow**)
Pair-labelled	L-operator expressions
$\{\langle Q, Q\rangle, \{\text{Who}_Q, \{\text{C}_Q, \text{fell}\}\}\}$	Whx(**fell**(x))

This mapping is, of course, a first approximation of a theory of the CI interface. Being a first approximation, it will face empirical and theoretical challenges. For instance, there are likely to be cases where an expression's label does not seem to map to its interpretation. These cases, however, might be cases in which our syntactic or semantic analysis is incorrect.

Furthermore, even if the hypothesized mapping in (8.41) were shown to be empirically adequate, it would still require theoretical explanation. That is, we would need to explain why that particular mapping holds. This would require us to show that there is a mathematically sound isomorphism between head-labelled SOs and function-argument expressions, and between pair-labelled SOs and l-operator expressions. Such a demonstration, however, is beyond the scope of this thesis.

8.4 Explaining ACC-ing clause subjects

With these modifications of label theory in place, we can consider how to account for the distribution of ACC-ing clause (AC) subjects. First, I will explain the fact that, in the case of complement ACs, the subject DP cannot move. This explanation will be essentially the same as the explanation of ECP effects given in Chomsky 2015. I will then explain the adjunct AC case, in which the subject DP must move. This explanation will require the modified version of label theory and will support the general restriction in (8.3).

As in Chomsky's (2015) analysis of *that*-trace effects and my explanation of the non-generation of resultatives in French-like languages in section 6.2, a DP can be blocked from moving out of a phrase {DP, XP} if the following conditions hold. First, the head X of XP bears features which must be valued by a DP in order to label XP. That is, X must

bear an incomplete set of, say, φ-features. Second, if the DP moved, then that movement would occur before DP and XP agree for the features in question. In other words, if a movement operation bleeds agreement and labelling requires that agreement relation, then that movement operation will be ruled out. Applying this logic to the complement AC, I hypothesize that the English Prog head has an incomplete φ-set which must be strengthened by Agree in order to provide a label. If this is the case, then we can explain the impossibility of movement from complement ACs if such a movement would bleed agreement.

The case of adjunct ACs, however, requires a more complicated explanation. The explanation will be split into two subsections: Section 8.4.1 will show that a DP can move out of an adjunct AC without causing a problem, and section 8.4.2 will argue that an in-situ DP in an adjunct AC runs into problems.

8.4.1 DPs can move from adjunct ACC-ing clauses

In order to show that movement from an adjunct AC is permitted, I will demonstrate, in some detail, how an adjunct AC structure is derived and interpreted. First, let's consider how the AC in (8.42) is derived.

(8.42) Joanna teaching her students.

Since the "contentful" portion of the AC—the portion that encodes lexical information and thematic structure—is almost entirely independent of the larger AC structure, I will assume its derivation is uncomplicated. That is, I will take for granted that if the complement of Prog is a VoiceP, and that that VoiceP is derived and labelled as it would be in a grammatical finite clause. Further support for this assumption comes from Harwood (2015) who argues that Prog is a phase head and, therefore, its introduction triggers the transfer of its complement, although not before the ACC-ing subject raises and merges with ProgP. So, the AC which is to be adjoined to the host VP has the structure in figure 8.6 The next stages of the derivation requires the DP *Joanna* to move out of the AC, leaving us with an unlabellable structure as per the discussion immediately preceding this subsection.[9] However, a derivation doesn't necessarily crash if it's derived structure is unlabellable. Rather it crashes when LA tries to label an unlabellable structure. Since the AC in figure 8.6 is adjoined to VP in this case, and, by hypothesis, the labelling algorithm operates only on the host of a host-adjunct structure, the AC is not processed by the labelling algorithm. If the AC is not processed by the labelling algorithm, then it doesn't matter whether the AC is labellable or unlabellable. The content of an adjunct does not matter in determining the labelling of the host.

[9]Prog bears a single φ-feature set, and cannot label unless that feature set is enriched by agreeing with a DP in its specifier. Since the DP moves before agreement can happen, Prog's feature set is not strengthened and cannot label.

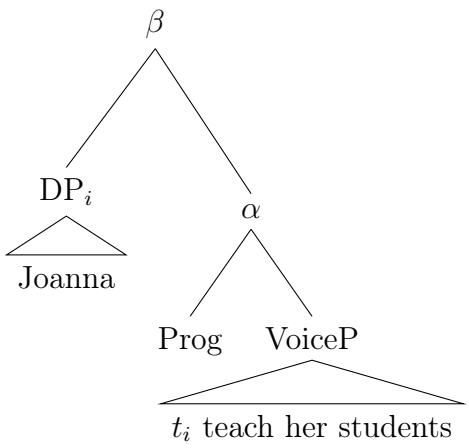


Figure 8.6: The structure of a ProgP

8.4.2 DPs must move from adjunct ACC-ing clauses

Thus far, I have argued that the labellability of an adjunct is immaterial, which suggests that DP movement from an adjunct AC would be optional. The facts of ACs, however, suggest that DPs must move from adjunct ACs. In order to explain why the movement operation in question is obligatory, we must consider how adjunct ACs are interpreted. Consider the illicit adjunct AC structure in figure 8.7 corresponding to the ungrammatical string in (8.43).

(8.43) *We can Mario the woman teaching her students.

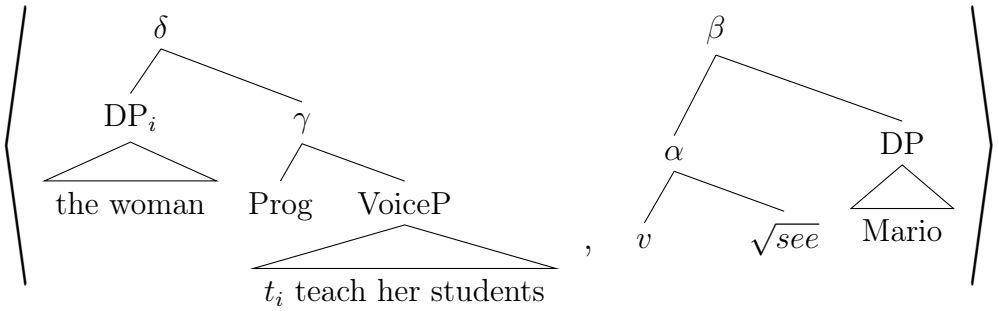


Figure 8.7: An illicit adjunct AC structure (δ adjoined to β)

When the host-adjunct structure is finally transferred as part of a larger phase, only the host β will be labelled. Neither the pair $\langle \delta, \beta \rangle$, representing δ adjoined to β, nor the adjunct δ itself will undergo labelling at this stage. As such, both of these objects will be null-labelled at CI, and therefore, interpreted conjunctively. This is expected for host-adjunct structures, but when we consider the interpretation of δ, we can see an issue with the structure in figure 8.7.

The adjunct δ will be null-labelled, as represented in (8.44) and, therefore, interpreted

conjunctively as shown in (8.45).

(8.44) $[_{\emptyset}$ [the woman]$_i$, $[_{\emptyset}$ Prog, $[_{\text{VoiceP}}$ the woman$_i$ teach her students]]]

(8.45) $\textsc{sem}(\delta) = \textsc{sem}(\text{the woman})(e)\&\textsc{sem}(\text{Prog})(e)\&\textsc{sem}(\text{VoiceP})(e)$

So, DP *the woman* and the VoiceP *the woman teach her students* are predicated of the same extra-mental entity e. In other words, there is some entity e which is both *the woman* and an event of the woman teaching her students. Notice, however, that the two instances of *the woman* in δ are not distinct SOs but occurrences of a single SO. Despite the fact that these two expressions are identical, they are interpreted as being predicated of two distinct entities. The upper copy is predicated of some event e, while the lower copy is predicated of some entity x which participates in the event of teaching students. If these copies are supposed to be identical, it seems like a contradiction to say that they are predicated of distinct entities. This contradiction, I propose, is the reason that DPs must move from specifiers of adjunct ACs.

If, on the other hand, the DP in [Spec, Prog] is a lower copy, then it can either be ignored by the CI system or treated as a variable. In either case, it will not be treated as a predicate, and, therefore, cannot be predicated of two distinct sorts of entities.

8.4.3 Generalizing the ACC-ing results

In the previous section, I offered an explanation for the fact that if an AC is adjoined to a VP, then its subject must move out of the AC. No part of the explanation, however, depended on any inherent property of the AC, but rather on the fact that the AC is adjoined, and the fact that the subject DP was internally merged in subject position. So, this leads us to the generalization in (8.3), which I restate schematically in (8.46).

(8.46) $*\langle\{\text{DP}_i\,\{\ldots t_i \ldots\}\}, \text{XP}\rangle$
$(^{\text{OK}}\langle\{t_i\,\{\ldots t_i \ldots\}\}, \text{XP}\rangle)$

Note that this seems to be violated by sentences like (8.47), where the modifier AC has its subject *in situ*.

(8.47) Her order having arrived late, Kinza was in a sour mood.

The AC in this sentence, however, is a sort of topic, and, by hypothesis, topical expressions are merged in a Topic projection rather than adjoined. So, the structure of (8.47) is given in figure 8.8.

The DP *her order* and Prog will Agree, meaning the AC γ will be labelled $\langle\varphi,\varphi\rangle$, and its constituent β will be labelled Prog. The process of labelling δ is a slightly more complicated

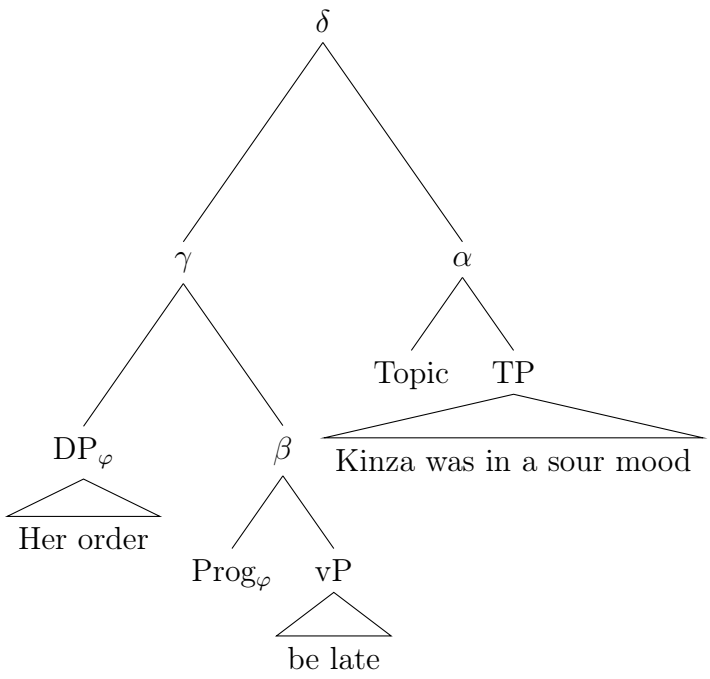


Figure 8.8: The unlabelled structure of (8.47)

case, though. Recall that the labelling algorithm is the process of finding the most prominent element in a syntactic object. Setting aside the cases that result in a feature-pair label, the most prominent element is the least embedded atomic object. So, what is the least embedded atomic object in δ? The likely candidates are Topic, Prog, or D (*her*), of which Topic is the least embedded. Prog is dominated by 3 nodes (β, γ, δ), as is D (DP,γ, δ). Topic, on the other hand, is dominated by 2 nodes (α, δ). Therefore, the label of δ is Topic, and, since (8.47) is grammatical, we can safely assume that Topic is strong enough to be a label. It follows, then, that the label of β (a head-phrase structure) will also be Topic. The end result of the labelling process, then, is represented in figure 8.9.

Turning back to resultatives, we can see that the proposed structure included the adjunction of a resP to a VP. Therefore, the restriction in (8.46) would apply to resultatives, and the DP in [Spec, res] would be required to move. Thus we have an explanation for the ungrammaticality of (8.1).

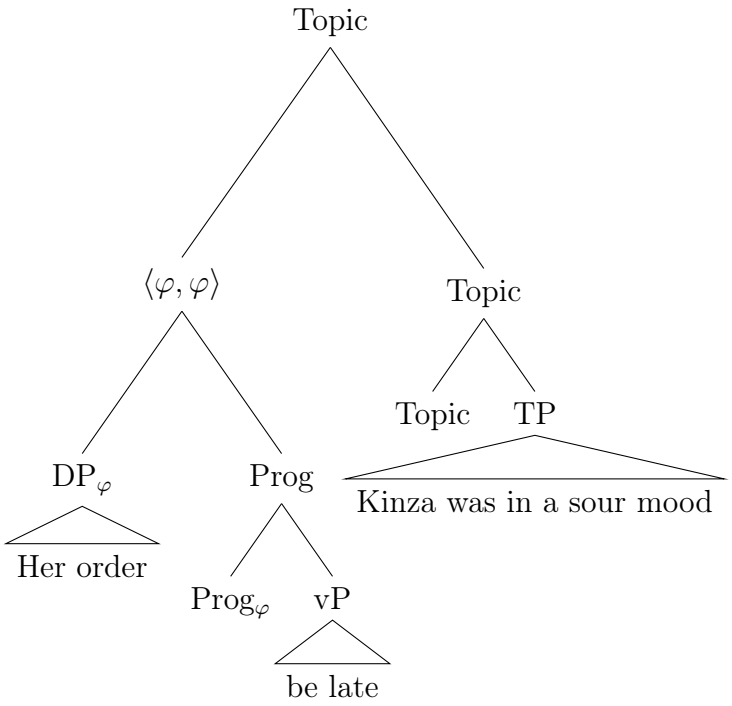


Figure 8.9: The labelled structure of (8.47)

Chapter 9

Coincidence

In this thesis, I have made liberal use of a novel class of sideward movement structures which I schematize below in figure 9.1. This structure was used in the analysis of adjectival

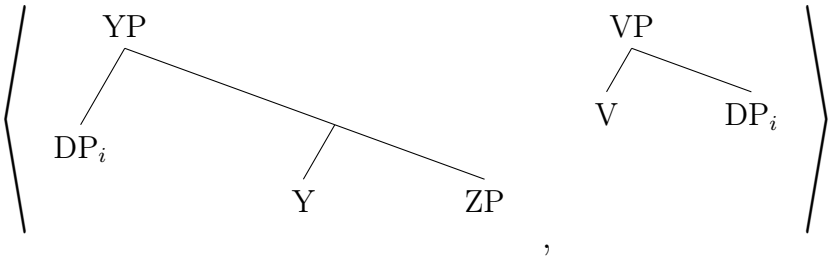


Figure 9.1: A schema of the sideward movement structure

resultatives and depictives, and for one available structure for direct perception reports with ACC-ing clauses. The distinction between these constructions is due to the choice of head Y. For resultatives, Y is instantiated by res, while for perception reports, it is instantiated by Prog. In depictives, the adjoined phrase YP is a small clause, so Y is either absent or instantiated by a Pred head. Note that, while the complement of Y, ZP, also varies from construction to construction, this variation can be derived from the selectional requirements of Y.

If we take VPs to be event descriptions, and we assume that host-adjunct structures compose by predicate conjunction, then a VP adjunct (SC for depictives, resP for resultatives, and ProgP for DPRs) must also be event description. That is, in figure 9.1 the VP and YP are both interpreted as predicates of events, and, because they combine by predicate conjunction, they both describe the same event. In this chapter, I discuss the interpretation of depictives and adjunct ACC-ing clauses. As for the interpretation of resultatives, nothing needs to be added to the discussion in section 4.2.

In the case of depictives, the interpretation is mostly straightforward, though not entirely without complications. Consider (9.1), which describes an eventuality in which Natasha eats

the fish while that fish is raw.

(9.1) Natasha ate the fish raw.

According to our analysis, this interpretation would be derived from the fact that the eating event and the rawness state are taken to be identical. That is the VP in (9.1) has a logical form as derived in (9.2)

(9.2) $\textsc{sem}(\textit{eat the fish raw}) =$
$\lambda e[\textsc{sem}(\textit{eat the fish})(e) \,\&\, \textsc{sem}(\textit{the fish raw})(e)] =$
$\lambda e[\mathbf{eating}(e) \,\&\, \mathbf{raw}(e) \,\&\, \textsc{Theme}(\mathbf{the\_fish})(e)]$

One might object that this LF is incoherent, as eating is an event, while rawness is a state, and an eventuality cannot be both an event and a state. This objection, however, does not hold up under scrutiny. Suppose we take the externalist perspective, according to which the entities that natural language expressions are predicated of are external to and independent of the mind, and the predicates and concepts of natural language correspond to natural kinds. From this perspective, eventualities are regions of space-time, some of which are events, while others are states. So, for instance, to utter (9.3) truthfully is to refer to a particular region of space-time.

(9.3) The officer ticketed the car.

Now, according to the objection at hand, the region of space-time referred to by (9.3) is an event, and, therefore, not a state. However, it is entirely reasonable to assume we could truthfully utter (9.4), a state description, referring to the same space-time region.

(9.4) The car was parked illegally.

It seems, then, that, if there is an event/state contrast, it does not originate in the extra-mental world, or else (9.3) and (9.4) could not possibly refer to the same region of space-time.

Furthermore, many adverbs describe states, yet may modify event descriptions. If adverbs are adjuncts, then they are interpreted as conjoined with their host, meaning that they will provide a partial description of an event, rather than a state. Therefore, there doesn't seem to be any contradiction in my analysis of depictives.

The case of adjunct ACC-ing clauses is slightly more challenging, due to a subtlety in their meanings which I will discuss below. Ultimately, however, their meaning can be explained from their structure, given in figure 9.2, and an assumption about the nature of eventualities. As a first pass, we can say that the interpretation of the structure in figure 9.2 is a description of an event of the dog being seen and an event of the dog running. This alone is not sufficient, as an English speaker's intuition regarding the (9.5) is that the *we* referent saw both the dog

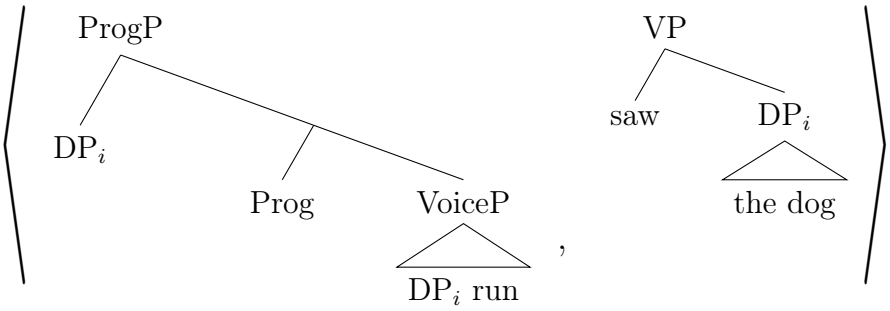


Figure 9.2: An adjunct ACC-ing structure

and the event of the dog running.

(9.5) We saw the dog running.

This intuition seems to be inescapable; English speakers cannot seem to entertain an interpretation of (9.5) in which *we* saw the dog but not the running event, or the event but not the dog. This suggests that both complement ACC-ing and adjunct ACC-ing versions of (9.5) are interpreted as both the individual and the event being seen. The strong version of UTAH that I assume in this thesis, however, predicts that the interpretation *x was seen* can only be encoded if the expression denoting x is merged as the complement of the verb *see*. In adjunct ACC-ing analysis of (9.5), as shown in figure 9.2, the event denoting expression, ProgP, is adjoined to VP, yet we interpret it as meaning that the event was seen. Since this interpretation is not directly encoded, we must infer it from what is directly encoded.

To see how we would infer the perception of the event, consider what is directly encoded. First, the VP is interpreted as a description of a seeing event which the dog is the theme of.

(9.6) $\textsc{sem}(\text{VP}) = \lambda e[\mathbf{see}(e) \,\&\, \textsc{theme}(\mathbf{the\_dog})(e)]$

The interpretation of the ProgP, represented in (9.7), however, is more complicated.

(9.7) [⟨the dog⟩, [Prog [$_{\text{VoiceP}}$⟨the dog⟩ run]]]

I will make the simplifying assumption that the copy of *the dog* in [Spec, Prog] is semantically vacuous[1] and discuss the Prog-VoiceP structure. The VoiceP is unremarkable, so I assume its meaning is the complete but tenseless event description in (9.8).

(9.8) $\textsc{sem}(\text{VoiceP}) = \lambda e[\mathbf{run}(e) \,\&\, \textsc{doer}(e)(\mathbf{the\_dog})]$

Prog, then, takes this description as an argument and ascribes progressive aspect to it. The

[1]At this stage, this is purely stipulative. I suspect some version of this assumption is true, but a full investigation and justification of it is beyond the scope of this thesis.

standard, though perhaps naïve analysis of progressive aspect (as proposed in Klein 1994) is that Prog takes a description of event e as an argument, introduces a topic time t, and asserts that t is included in the run-time of e. This predicts that ProgP encodes the predicate of times in (9.9)

(9.9) **First approximation:**
$\text{SEM}(\text{ProgP}) = \lambda t \exists e[t \subseteq \text{TIME}(e) \,\&\, \mathbf{run}(e) \,\&\, \text{DOER}(e)(\mathbf{the\_dog})]$

However, since ProgP adjoins to VP and the resulting structure is interpreted as a conjunction, ProgP must be interpreted as a predicate of events. Therefore, I will modify the semantic analysis of Prog such that it introduces an event e' and asserts that e' is included in e. The final interpretation of the ProgP, is given in (9.10) (cf. Bjorkman, Cowper and Siddiqi 2018).

(9.10) **Second approximation:**
$\text{SEM}(\text{ProgP}) = \lambda e' \exists e[e' \subseteq e \,\&\, \mathbf{run}(e) \,\&\, \text{DOER}(e)(\mathbf{the\_dog})]$

This interpretation will properly compose with (9.6) to yield the interpretation of the host-adjunct structure in (9.11).

(9.11) $\text{SEM}(\langle \text{ProgP, VP} \rangle) = \lambda e' \exists e[\mathbf{see}(e') \,\&\, \text{THEME}(\mathbf{the\_dog})(e')$
$\&\, e' \subseteq e \,\&\, \mathbf{run}(e) \,\&\, \text{DOER}(e)(\mathbf{the\_dog})]$

So, the event of seeing the dog is included in the event of the dog running, meaning that the seeing occurred in the same space-time region as the running and therefore we can infer that the running event was seen. This denotation along with the very nature of seeing and running allows us to infer from (9.11) that we saw the running event.

One could argue that this analysis is implausible as it requires that the seeing event is a part of the seemingly independent running event. On its face, this seems to imply an interdependency between the two events, and, while it seems reasonable to say that the perception event depends on the perceived event, it is far from obvious that the perceived event depends on the perception event. This line of argumentation, I believe, is not enough to rule out the logical form in (9.11).

Saying that the perception of an event is a sub-part of that event, does imply that the event is dependent on it being perceived, but it does so in a very weak way. It is perhaps a truism of set theory and mereology to say that two complex objects are identical only if they consist of the same parts. So, if x is a part of e but not a part of e', the $e \neq e'$. Similarly, a particular running event which is seen by some individual x, cannot be identical to a running event which is not seen by x. Note that this does not mean that the unseen running event is not a running event, only that it is a not a seen running event.

Furthermore, there is linguistic evidence that an event of x being seen can be construed as a part of an event of x running. Consider, for instance, sentences of the form *In the course of XP, S*, such as (9.12) and (9.13).

(9.12) In the course of her morning run today, Sadie saw three new coffee shops.

(9.13) In the course of her investigation, Kima interviewed two hotel receptionists.

Sentences of this form seem to entail that the event described by S is a proper subpart of the event described by XP. So seeing the new coffee shops was a part of the running event, as interviewing receptionists was part of the investigating event. Compare these to the infelicitous uses of this construction in (9.14) and (9.15).

(9.14) #?In the course of her morning run today, Sadie met her writing partner for coffee.

(9.15) #In the course of her investigation, Kima got married.

In both cases, the would-be subevent is judged to be either incompatible with the would-be superevent, or not a natural subevent of the would-be superevent. In the case of (9.14), a morning run is a more-or-less uninterrupted event, and meeting someone for coffee would constitute an interruption. In the case of (9.15), even if the getting married event occurred while the investigation were ongoing, the sentence as given means that the marriage is somehow part of the investigation. For instance (9.15) would be true if Kima got married as part of an undercover operation. Now consider (9.16), which entails that an event of someone seeing Sadie is included in an event of Sadie running.

(9.16) In the course of her morning run today, Sadie was seen by Declan.

Since this sentence is not judged to be odd, we can infer that there is no semantic or conceptual reason to rule out the interpretation in (9.11).

That being said, I will now entertain two alternative analyses and discuss their flaws.

Suppose, for instance, that the semantics of Prog is about time rather than eventualities, as proposed by Klein (1994). This can be attained without the compositionality issues discussed above if we hypothesize the denotation in (9.17).

(9.17) $\text{SEM}(\text{Prog}) = \lambda P_{\langle s,t\rangle}\lambda e'\exists e[\tau(e') \subseteq \tau(e)\&P(e')]$

This denotation would predict that the sentence in (9.5), under the adjunct ACC-ing interpretation, would mean that there was an event $e1$ of us seeing the dog, and an event $e2$ of the dog running, and that the time of $e1$ is included in the time of $e2$. And while one could certainly infer that if one sees a dog at the same time as the dog is running, then one sees the running, such an inference does not hold in other cases. Consider the proposed adjunct

ACC-ing interpretation of (9.18), given in (9.19).

(9.18) I [[$_{\text{VP}}$ saw Ronaldo$_i$][$_{\text{ProgP}}$ t_i playing soccer on TV]].

$$\begin{aligned}\textsc{sem}((9.18)) \;=\; & \exists e, e'[\textsc{Exper}(e)(\mathbf{speaker}) \,\&\, \mathbf{see}(e) \,\&\, \textsc{theme}(e)(\mathbf{R}) \\ & \,\&\, \textsc{Agent}(e')(\mathbf{R}) \,\&\, \mathbf{playing\_soccer\_on\_TV}(e') \\ & \,\&\, \tau(e) \subseteq \tau(e')]\end{aligned} \tag{9.19}$$

While this hypothesized interpretation is consistent with the actual interpretation of (9.18), it is also consistent with some non-existent interpretations. For instance, (9.19) is consistent with a situation in which the speaker sees Ronaldo, while a replay of one of his matches plays on the TV in the other room.

To be more forceful, this hypothesized interpretation of Prog predicts that the illicit sentences in (8.15) and (8.16), reproduced below, should be licit.

(9.20) *The writer was heard being slandered.

(9.21) *The singer was seen being parodied.

Since the coincidence between the perception event and the slandering/parodying event is strictly temporal, there is no reason to require that the two events occur in the same room, or even in the same hemisphere. So, suppose I were sitting at a café with some singer, while at the same time, "Weird Al" Yankovic was recording a parody of that singer in a studio across town. If Prog merely encodes temporal coincidence, then (9.21) should be licit and true of this situation. These examples, however, are illicit, and therefore the temporal coincidence hypothesis does not fare as well as my hypothesis.

We could, of course, seek to stipulate that the ProgP event (*e.g.*, the event of the dog running) is perceived, but I argue that such a move is untenable. Suppose that the logical form in (9.22) is the final logical form of the VP in (9.5).

$$\begin{aligned}\textsc{sem}((9.5)) \;=\; & \lambda e \exists e'[\mathbf{see}(e) \,\&\, \textsc{theme1}(e)(\mathbf{the\_dog}) \,\&\, \textsc{theme2}(e)(e') \\ & \,\&\, \exists e''[e' \subseteq e'' \,\&\, \mathbf{run}(e'') \,\&\, \textsc{doer}(e'')(\mathbf{the\_dog})]]\end{aligned} \tag{9.22}$$

There are, as far as I can tell, three possibilities for where the novel THEME2 predicate is encoded. It is either (*i*) encoded in Prog, (*ii*) encoded in the perception verb, or (*iii*) encoded on an independent functional head, which I will call Θ. I will discuss each of these below in turn and show that they all fail empirically.

First, consider the option of encoding a theme predicate in the Prog head. This obviously wouldn't do for the case in which the present participle is the main verb of the clause (*e.g.*, *The dog is running.*), so we need to hypothesize a new head which I will call Prog$_\Theta$, with the denotation in (9.23).

(9.23) $\text{SEM}(\text{Prog}_\Theta) = \lambda P \lambda e \exists e', e''[\text{THEME}(e)(e') \,\&\, e' \subseteq e'' \,\&\, P(e'')]$

Now consider the adjunct ACC-ing interpretation of (9.5) under this hypothesis. The Prog_ΘP interpretation will be (9.24), and the host VP interpretation will be (9.25).

(9.24) $$\begin{aligned}\text{SEM}(\text{Prog}_\Theta\text{P}) \;=\; & \lambda e \exists e', e''[\text{THEME}(e)(e') \,\&\, e' \subseteq e'' \\ & \,\&\, \text{DOER}(e'')(\mathbf{the\_dog}) \& \mathbf{run}(e'')]\end{aligned}$$

(9.25) $\text{SEM}(\text{VP}) = \lambda e[\text{THEME}(e)(\mathbf{the\_dog}) \,\&\, \mathbf{see}(e)]$

If these expressions were conjoined by adjunction, then we would get the expression in (9.26), where the seeing event has two distinct themes.

(9.26) $$\begin{aligned}\text{SEM}(\langle \text{Prog}_\Theta\text{P}, \text{VP}\rangle) \;=\; & \lambda e \exists e', e''[\mathbf{see}(e) \\ & \,\&\, \text{THEME}(e)(\mathbf{the\_dog}) \,\&\, \text{THEME}(e)(e') \\ & \,\&\, e' \subseteq e'' \,\&\, \text{DOER}(e'')(\mathbf{the\_dog}) \& \mathbf{run}(e'')]\end{aligned}$$

This is problematic since there seems to be a uniqueness restriction on thematic roles: in a given event description, each thematic role predicate must be true of at most one (possibly plural) individual. We can see this in the case of the agent role in (9.27) and the beneficiary or recipient role in (9.28).

(9.27) * Angela duped Phyllis by Pam.

(9.28) ? Toby baked Kevin a cake for Oscar.[2]

We could stipulate that Prog_Θ encodes a secondary theme predicate THEME2, but this would be entirely ad hoc, as this role would only be used for these direct perception reports. Note, of course, that this line of argumentation would apply to the other two possibilities investigated below. I could, therefore, cut the discussion short and declare the analysis in (9.22) irreparably flawed. There are other arguments against the analysis though so I will present those below.

Consider the option of encoding the secondary theme in the perception verb. Our new version of *see*, then, is given in (9.29).

(9.29) $$\begin{aligned}\text{SEM}(\text{see}_\Theta) \;=\; & \lambda x_e \lambda P_{\langle s,t\rangle} \lambda e_s \exists e'[\mathbf{see}(e) \,\&\, \text{THEME1}(e)(x) \\ & \,\&\, \text{THEME2}(e)(e') \,\&\, P(e')]\end{aligned}$$

The issue with this hypothesis is that it requires the ACC-ing clause to be not an adjunct, but an argument. Note that the second argument of the denotation of see_Θ is a predicate of events, which is satisfied by the ACC-ing clause. However, by hypothesis (see section 8.3), an

[2]This sentence is acceptable if Kevin is the recipient and Oscar is the beneficiary or vice-versa, but not if Kevin and Oscar have the same role.

expression can only be an argument of a verb if it is merged with, rather than adjoined to the VP. Therefore, this proposal requires the ACC-ing clause to be merged with the VP, rather than adjoined to it, as in figure 9.3. This structure, however, is untenable without significant

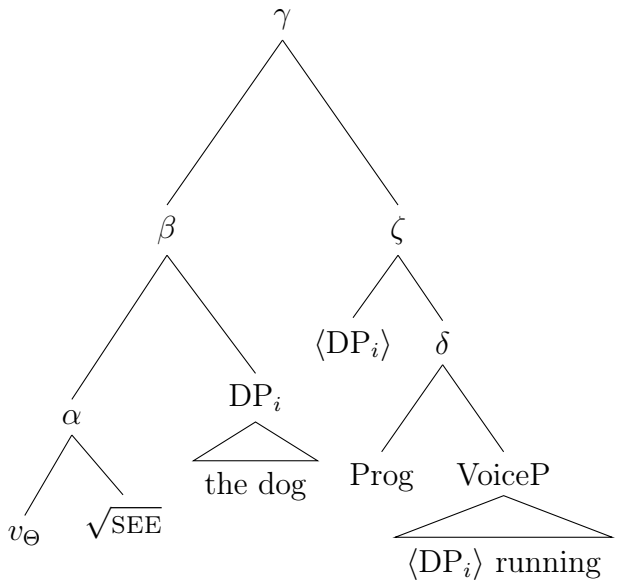


Figure 9.3: The unlabelled structure of (9.5) required by the *see*$_\Theta$ proposal

additional stipulation for two reasons. First, it is unlabellable. Consider how we would label γ, which is a phrase-phrase structure. The labelling algorithm would search γ and find the least-embedded, still-visible, functional LIs. In this case, v_Θ and Prog are equally embedded and thus are returned by LA. This can result in an unambiguous label only if we make one of two additional stipulations. Either (*i*) v_Θ and Prog agree for some feature, or (*ii*) the ProgP ζ undergoes movement. Stipulation (*i*) requires at least two additional stipulations: a stipulation as to the identity of the agreeing features, and a stipulation that agreement can occur at such a structural distance. I strongly suspect that these sub-stipulations will require stipulations of their own, and so on, but I stop here. Stipulation (*ii*), requires us to stipulate a landing site for the movement operation; one that ensures the proper word order. This new landing site implies a new functional head, whose identity will also likely be stipulated. As with stipulation (*i*), this stipulation likely begets more stipulations.[3] The level of stipulation required in order to label the structure in figure 9.3, I believe, renders it untenable.

Even if we were to assume that the structure in figure 9.3, or a structure derived from it, is labellable, there is another reason that this analysis cannot be maintained. Recall that in

[3]A skeptical reader might point out that this is the very nature of scientific inquiry: Explaining one phenomenon invariably brings to light several phenomena in need of explanation. This view might suggest that one person's stipulation is another's hypothesis. I would argue, however, that these hypotheses do not raise new areas of inquiry, but rather allow us to avoid raising new areas of inquiry. Recall that the impetus for these stipulations was to avoid a vague unease with the semantic analysis in (9.11) that followed naturally from the syntactic analysis assumed in this chapter. Such an unease, I believe, does not justify this level of stipulation.

chapter 8 I showed that, when an ACC-ing clause is merged as an argument, its subject must remain in situ. This predicts, not only that ungrammatical sentences with two DP objects like (9.30) should be grammatical, but also that (9.5) should be ungrammatical.

(9.30) *We saw the dog the cat running.

In view of the problems with this proposed analysis, I will set it aside.

This brings us to our final possibility: encoding the THEME2 on a specialized functional head Θ. Under this analysis, (9.5) would have the structure in figure 9.4.

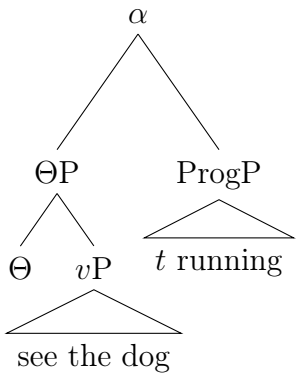


Figure 9.4: The structure of (9.5) with an independent Θ head

Assuming Θ takes vP as an argument, and ProgP is adjoined to the resulting ΘP, lets consider a possible meaning for Θ in (9.31)

(9.31) **First approximation:**
$\text{SEM}(\Theta) = \lambda P_{\langle s,t\rangle}\lambda e\exists e'\,[P(e)\,\&\,\text{THEME2}(e)(e')]$

This SEM says that there is some event which is the secondary theme of the event described by the vP. Now, when this combines with the vP, it yields the predicate of events in (9.32).

(9.32) $\text{SEM}(\Theta\text{P}) = \lambda e\exists e'\left[\textbf{see}(e)\,\&\,\text{THEME}(e)(\textbf{the\_dog})\,\&\,\text{THEME2}(e)(e')\right]$

If this is then conjoined with the ProgP, we get the predicate of events in (9.33).

(9.33) $$\begin{aligned}\text{SEM}(\alpha) \;=\; & \lambda e\exists e'[\textbf{see}(e)\,\&\,\text{THEME}(e)(\textbf{the\_dog})\,\&\,\text{THEME2}(e)(e')\\ & \&\, e\subseteq e'\,\&\,\text{DOER}(e')(\textbf{the\_dog})\&\textbf{run}(e')]\end{aligned}$$

This interpretation has two flaws. First, it asserts that the entire running event e' is the secondary theme of seeing. This is far too strong of an assertion; (9.5) is compatible with only seeing a sliver of the running event. Second, it asserts that the seeing event e is a subpart of the running event e', which is precisely the assertion we are trying to avoid. These issues can be avoided by swapping the variables e and e' in our denotation of Θ as in (9.34).

(9.34) **Second approximation:**
$\text{SEM}(\Theta) = \lambda P_{\langle s,t\rangle} \lambda e' \exists e\, [P(e)\,\&\,\text{THEME2}(e)(e')]$

The resulting denotation of the structure in figure 9.4 is given in (9.35)

(9.35) $$\begin{aligned}\text{SEM}(\alpha) \;=\; & \lambda e' \exists e[\mathbf{see}(e)\,\&\,\text{THEME}(e)(\mathbf{the\_dog})\,\&\,\text{THEME2}(e)(e') \\ & \&\, e \subseteq e'\,\&\,\text{DOER}(e')(\mathbf{the\_dog})\&\mathbf{run}(e')]\end{aligned}$$

This no longer has the flaws of the denotation in (9.33), but it has introduced a new one. The denotation in (9.35) is a predicate of events, but it no longer describes a seeing event, but a running event. So, if we continued the derivation of (9.5), and introduce the external argument *we*, then that argument will be construed as the experiencer/agent of the running event, not the seeing event.

The independent Θ head analysis, it seems, cannot save us, and therefore, the secondary theme analysis cannot save us. Since our attempts to avoid the analysis in (9.11) have only made things worse, I see no alternative to adopting our first analysis, despite the sense of unease it may evoke. Note that we are tied to the analysis in (9.11) only insofar as we adopt the admittedly simplistic semantic analysis of progressive aspect used here. I leave the task of exploring the consequences of more precise semantic analyses of progressive aspect to others.

All of this is just to say that (9.11) is a *plausible* semantic analysis for adjunct ACC-ing VPs. Whether or not it is the correct analysis is, of course, an empirical question, one which may require quantitative experimental methods to answer. Although such an investigation is beyond the scope of this thesis, I can make the question more precise. The question comes down to a contrast between how argument ACC-ing structures and adjunct ACC-ing structures are interpreted, respectively, and therefore asking the question will require us to disambiguate the two structures. We've seen that, while active perception clauses, such as (9.5) (repeated below), are ambiguous between the two readings, passive perception clauses, such as (9.36) seem to only admit the adjunct ACC-ing reading.

(9.5) We saw the dog running.

(9.36) The dog was seen running.

By hypothesis, then, passives like (9.36) should only admit readings under which the perceived event "contains" the perception event, While actives like (9.5) should also admit readings under which the perceived event does not contain the perception event. The best way to do this is to construct active-passive pairs like (9.5) and (9.36) in which the passive, but not the active is judged to be contradictory. More precisely, we need to construct these active-passive pairs such that the source of the would-be contradiction is the hypothesized containment relation.

To this end, I propose the use of direct perception reports in which the perceived event

is potentially imperceptible. For instance, mental activities are potentially imperceptible so there is no contradiction in (9.37).

(9.37) Koorosh$_i$ saw Sahar thinking but he$_i$ didn't realize it.

Similarly, there are a host of physical events that are imperceptible to the naked eye—events such as paint drying, water heating, or skyscrapers swaying. For this reason, neither the sentences in (9.38) are contradictory, nor are the situations they describe impossible.

(9.38) a. We saw the paint drying imperceptibly.
b. She saw the water heating up imperceptibly.
c. You saw the skyscraper swaying imperceptibly.

If my hypothesis is correct, though, the passive versions of (9.37) and (9.38) should entail that an imperceptible event contains its own perception—a clear contradiction. Therefore, the sentences in (9.39) should be contradictory.

(9.39) a. Sahar was seen thinking by Koorosh, but he didn't realize it.
b. The paint was seen drying imperceptibly.
c. The water was seen heating up imperceptibly.
d. The skyscraper was seen swaying imperceptibly.

My own judgments about these sentences are unclear and, of course, likely tainted by self-interest, which is why I believe they should be tested experimentally. As of the writing of this thesis, however, I lack the time, resources, and know-how to perform such an experiment. I, therefore, leave it as a task for future research.

Chapter 10

Conclusion

The parametric variation of resultative, what I call the resultative parameter, presents a puzzle for linguistic theory: Children acquire their language's parameter-setting despite a seeming lack of direct evidence in the primary linguistic data. Since there is no direct evidence, the parameter setting must follow from indirect evidence. This line of reasoning leads to a two-part research question: What aspect of a child's PLD provides indirect evidence for the setting of the resultative parameter, and how is the parameter setting deduced from that aspect? Each part of that question, it turns out, calls for a dissertation-length answer. The first part is largely answered by William Snyder's dissertation (Snyder 1995) and refined in his later work (Snyder 2001, 2012, 2016). Snyder's answer is that children use the availability of bare stem compounding in their PLD as indirect evidence for the availability of resultatives in their target grammar. My dissertation begins with this result, and aims to provide an answer to the second part of the question: A language may generate both bare stem compounds and adjectival resultatives only if its lexicon has categorizing heads without φ-features.[1]

Answering the question "How are resultatives linked to bare stem compounding?" is a theoretical task, and, as with any theoretical task, it begins with an explicit litany of theoretical assumptions. I make many of the assumptions standardly made in early 21st century generative syntax (Merge, the Y-model of grammar) and a number of non-standard assumptions. First, I assume that the Θ-criterion does not fully hold and that an argument may receive multiple θ-roles. Second, I assume that Merge operates freely, provided that there are two syntactic objects to be combined. finally, I assume that there is no operation

[1]This, of course, grossly oversimplifies the range of possibilities available for lexical variation. Restricting ourselves to categorizing heads, we can express the logical range of possibilities as in (i).

(i) For each category *cat*, a non-empty subset of $\{cat_{\emptyset}, cat_{F}\}$ is included in the lexicon.

Furthermore, the choice of lexicon will certainly affect the grammar in a variety of ways. This can be seen by comparing isolating languages such as Niuean, which seem to lack any morphological agreement, to languages such as Italian, which shows a great deal of agreement morphology.

Agree active in the Narrow Syntax. These assumptions, as I discuss in chapter 2, despite being non-standard, actually follow from the logic of the minimalist program.

In order to provide any answer to the question of how resultatives are related to bare stem compounds, we must have an idea of what adjectival resultatives are. That is, we must give a syntactic analysis of resultatives. Furthermore, we must provide what I call a parametric analysis—an analysis of how a parameter may be acquired and represented in the grammar. To that end, I discuss previous analyses in chapter 3 before offering my own in chapter 4. The syntactic analysis I offer, reproduced in figure 10.1, is one in which a result phrase is adjoined to the VP and a DP undergoes sideward movement between them. The parametric

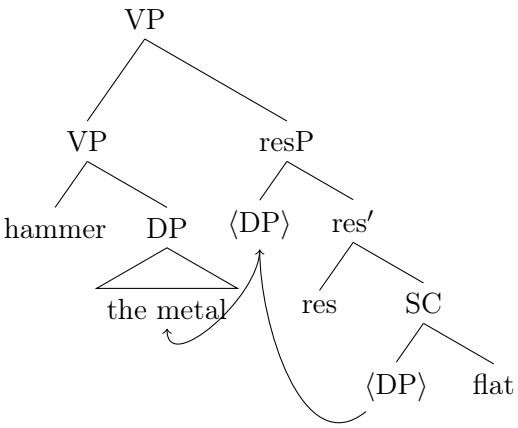


Figure 10.1: The structure of resultatives

analysis, I offer is based on a similar one by Kratzer (2005). According to this analysis, the presence of bare stem compounding in a child's PLD signals that the child's lexicon should admit categorizing heads without φ-features.

In order to show that resultatives depend on φ-less heads, we must show that a structure such as figure 10.1 can be derived only if the lexicon contains φ-less categorizing heads. I do so in chapter 6, but only after discussing the latest iteration (at least at the time of this thesis) of Chomsky's syntactic theory—label theory—in chapter 5. According to label theory, a syntactic derivation only converges if the structure it creates can be unambiguously labelled. In chapter 6, I show that the structure in figure 10.1 can be derived and labelled if the result adjective *flat* is categorized by a φ-less head $adj_{\emptyset}$. I then show that if *flat* is categorized by adj_{φ}, the derivation either fails or creates an unlabellable structure. Thus I have answered the question at hand.

In part II, I bring to the forefront the apparently loose theoretical ends left by Part I. Rather than tie these loose ends up with auxiliary or ad-hoc hypotheses, I investigate how they might inform our theory of the language faculty. In chapter 7, I argue that an apparent undergeneration problem of my proposed theory is actually due to the lack of a suitable theory of feature agreement. Such a theory, I propose, is one in which agreement occurs

postsyntactically.

In chapter 8, I point out an odd fact—that movement from [Spec, res] to [Comp, V] in figure 10.1 seems to be obligatory—and show that it seems to generalize to other cases of sideward movement—objects in the specifier of adjoined phrases must move to the host phrase. I argue that this fact is odd only if we make the standard (although often tacit) assumption that grammaticality is determined within the Narrow Syntax. Under an interface-based theory, such as label theory, this fact can be accounted for.

Finally, in chapter 9 I discuss a semantic question raised by my proposal. I assume that primary and secondary predicates compose via something like predicate modification. This mode of composition leads to what initially seem to be odd interpretations in which the events described by the two predicates are in fact the same event. I argue that despite this apparent oddness, there is no other principled way of interpreting the structures in question (*i.e.*, resultatives, depictives, and some direct perception reports), and furthermore, that the apparent oddness is only apparent. A closer look at both structures and the ontology of eventualities significantly diminishes this oddness, but a closer investigation will be needed to corroborate the predicted interpretation.

There are, of course, loose ends left by this thesis, which will have to be tied up in later investigations. My starting points for the thesis were works on adjectival resultatives by Snyder (1995, 2001, 2016) and Kratzer (2005) who drew a strong correlation between adjectival/nominal inflection and adjectival resultatives. As with all empirical generalizations, there are exceptions to this correlation. Exceptions, of course, are tricky things in any scientific inquiry. They can either strengthen a theory or destroy it, and there is no way to tell which they will do without a full analysis

For instance, Italian, which is one of the prototypical *resultative languages, does seem to generate a form of adjectival resultative, but only under fairly restrictive conditions. Napoli (1992), for instance, gives the following examples of Italian adjectival resultatives.

(10.1) a. Ha dipinto la macchina rossa.
"He painted the car red."

b. (i) Ho stirato la camicia piatta piatta.
"I ironed the shirt flat flat."

(ii) *Ho stirato la camicia piatta.

Folli and Ramchand (2005) add to this list the following cases, in which the result AP is intensified with *troppo*.

(10.2) a. Gianni ha cucito la camicia *(troppo) stretta.
"John sewed the dress *(too) tight."

b. Gianni ha sciolto il cioccolato *(troppo) liquido.

"John melted the chocolate *(too) liquid."

Napoli (1992) suggests a semantic/pragmatic analysis; namely, that Italian only allows resultatives when "the verb can be interpreted as focusing on the endpoint of its activity" (p75). Folli and Ramchand (2005), on the other hand, suggest a syntactic analysis; Italian only allows resultatives when the result AP is complex. Neither analysis is complete, though, and the case of Italian resultatives remains a puzzle.

If Napoli is correct, and the case of Italian resultatives is to be given a semantic/pragmatic analysis, then my syntactic explanation of the resultative parameter will face some difficulties. If, on the other hand, Folli and Ramchand are correct that this exception is to be given a syntactic analysis, then perhaps it will only strengthen my proposal.

Whelpton (2007) presents Icelandic as a possible counterexample to Kratzer's (2005) proposal. Recall that Kratzer's analysis was that resultatives could only be derived if the result adjective was uninflected, a proposal that is compatible with mine. Whelpton shows that, while Icelandic allows resultatives, it also seems to require inflectional morphology on result adjectives as in the following examples.

(10.3) a. Ég kýldi lögguna kalda.
I.Nom punched cop.the.FSgAcc cold.FSgAcc
"I punched the cop out cold."

b. Járnsmiðurinn hamraði álminn flatan.
blacksmith.the hammered metal.the.MSgAcc flat.MSgAcc
"the blacksmith hammered the metal flat."

c. Dóra æpti sig hása.
Dóra screamed herself.FSgAcc hoarse.FSgAcc
"Dóra screamed herself hoarse."

However, Whelpton also notes that Icelandic, unlike a prototypical *resultative language, does have bare stem compounding. Indeed, it has bare stem compounding that is interpreted as resultatives as in the following examples where bare result adjectives are compounded with deverbal adjectives.

(10.4) a. svart-litaður
black-coloured.mSgNom

b. þunnsneiddu sveppirnir
thin-cut.MPlNom mushrooms.the

c. fínmuldu piparkornin
fine-ground.NPlNom peppercorns.the

d. hreinskrúbbuðu pönnurnar
clean-scrubbed.FPlNom pans.the

e. mjúkbrædda súkkulaði

soft-meltedNSgNom chocolate

Whelpton presents this and other data as a rebuttal to Kratzer's (2005) analysis but offers no deep analysis or counter-proposal.[2] Without a deeper analysis, it is difficult to estimate the importance of his data as counterevidence to my proposal. Therefore I leave it to further investigation.

In addition to possible counterexamples, there are a number of phenomena related to adjectival resultatives which may be amenable to an analysis/explanation along the lines of what I propose here. First off, there is the case of directionalized locatives such as one of the (a) reading of (10.5).

(10.5) Kate kicked the ball between the posts.

a. ≈ Kate kicked the ball such that it passed/landed between the posts. (directionalized)

b. ≈ Kate stood between the posts and kicked the ball. (plain locative)

While these PPs are standardly assumed to be PathPs like PPs headed by, say, *through* or *around*, I argue elsewhere (Milway forthcoming) that such an assumption is unfounded. Rather, directionalized locatives are perhaps analyzable as PP resultatives based on their semantics. Furthermore, they show a parametric variation similar to that of adjectival resultatives. So, Germanic languages seem to have directionalized locatives, but Romance languages do not. There are, however, reports that certain varieties of Acadian French allow directionalized locatives. For instance, according to Ruth King and Yves Roberge (p.c. cited in Rooryck 1996, pp. 253–254) report that sentences like (10.6), while they only receive a plain locative reading in Metropolitan and Laurentian French, receive a directionalized locative reading in PEI French.

(10.6) La bouteille flottait [sous le pont].
The bottle floated under the bridge. (Rooryck 1996)

In previous work (Milway 2015), I hypothesized that this could be linked to the fact that, unlike Metropolitan and Laurentian French, Acadian French tends to allow P-stranding. So, for instance, the sentences in (10.7) are acceptable in PEI French but ungrammatical in most other varieties of French.

(10.7) a. Le ciment a été marché dedans.
the cement has been walked in
"The cement was walked in"

[2] On its face, Snyder's (2012) analysis of the resultative parameter seems to be able to account for Icelandic resultatives. This, however, is only true insofar as Snyder's analysis is theoretically permissible, a proposition that I dispute in chapter 3.

b. Oú il vient de?
where he comes from
"Where does he come from?" (Roberge and Rosen 2013)

A full analysis and explanation would require an in-depth empirical study, perhaps of the sort Snyder (1995) performed on adjectival resultatives. I leave such a study for future research.

Secondly, there are serial verb constructions (SVCs) of the type studied by Baker and Stewart (1999) and Stewart (2013). Consider, for example, the Edo SVC in (10.8).

(10.8) Òzó ghádìyán rè.
Ozo FUT buy yam eat
"Ozo will buy yams and eat them." (Baker and Stewart 1999)

SVCs and resultatives are similar in that both involve a single argument shared between two predicates which are related to each other by more than mere coincidence. So, in (10.8), the buying event is a prerequisite of the eating event, and the latter is, in some sense the goal of the former, and yams are the theme of both events. Also like resultatives, SVCs are parameterized, though they are rarer typologically than resultatives. Indeed, Stewart (2013) proposes that the two constructions are linked and that the SVC parameter may be a subparameter of the resultative parameter. Further research would be required to integrate my results with those of Baker and Stewart (1999) and Stewart (2013),

Finally, there is the case of Romanian bare noun resultatives[3] as discussed by Irimia (2012, pp. 220–224) and Farkas (2011). Romanian, like other Romance languages, disallows adjectival resultatives as shown in (10.9).

(10.9) *Femeia a curățat casa strălucitoare.[4]
Woman.the has cleaned.PstPrt house.the spotless.FSg
"The woman cleaned the house spotless."(Irimia 2012)

Unlike the other Romance languages, however, Romanian has a bare nominal resultative as shown in (10.10).

(10.10)Studentul s -a supărat foc.
student-the CL.3ReflAcc has get angry.Perf fire
"The student has got so angry that he became as red as fire."(Farkas 2011)

As the name suggests, the result nominal in a bare nominal resultative, despite the fact that Romanian allows nominal inflection.

[3]Irimia (2012) calls these "bare noun *pseudo*results."

[4]Irimia (2012) reports that this sentence is grammatical in Romanian, but only receives a depictive reading.

(10.11)a se supăra foc/*focul/*un foc/*focuri/*focurile
"to get angry fire/fire-the/a fire/fires/fires-the" (Farkas 2011)

Although this clashes with the generalization that Romance languages disallow resultatives, it is entirely consistent with my proposal. If we propose that the Romanian lexicon has the $n_{\emptyset}$ head but not the $adj_{\emptyset}$ head, then the bare noun resultative can be integrated with my analysis. This does, however, raise the question of why the bare noun resultative does not show up in other languages. I leave this question to future research.

The proposals made here are, of course, provisional as is the case for any scientific proposal. That is, they are subject to revisions, clarifications, and perhaps outright refutation. That said, I believe that with this thesis I have made two broad contributions to the ongoing study of the human language faculty. First, I have presented a template for the explanation of parametric variation, especially parametric semantic variation. Such variation can be explained by first finding a surface correlate of that parameter, and then showing how that correlate can be connected to the parameter. Second, I have incrementally developed the theory of the language faculty by identifying and fixing flaws in our understanding of such things as the syntax-semantics interface and adjunction. The flaws were found by applying the logic of the minimalist program to these domains, as were the proposed solutions to those flaws. I believe my solutions to be intriguing and suggestive, but they may, of course, be dead ends. The flaws, themselves, however, are more important; they represent domains that we previously thought we understood. Finding gaps in our understanding, such as these, is what makes scientific inquiry worth it. A failure of understanding is merely an opportunity to understand.

Bibliography

Arregi, Karlos and Andrew Nevins (2013). 'Contextual neutralization and the Elsewhere Principle'. In: *Distributed Morphology Today*. Ed. by O. Matushansky and A. Marantz. Cambridge: MIT Press.

Baker, Mark C. (1988). *Incorporation: A Theory of Grammatical Function Changing*. University of Chicago Press.

Baker, Mark C. and Osamuyimen T. Stewart (1999). *On Double-Headedness and the Anatomy of the Clause*. Rutgers University. unpublished.

Bjorkman, Bronwyn M., Elizabeth Cowper and Daniel Siddiqi (2018). *Unifying -ing without participles*. Poster presented at GLOW 41, Budapest, April 2018.

Bjorkman, Bronwyn M. and Hedde Zeijlstra (2014). 'Upward agree is superior'. Ms.

Bobaljik, Jonathan David (2008). 'Where's Phi? Agreement as a post-syntactic operation'. In: *Phi Theory: Phi-Features Across Modules and Interfaces: Phi-Features Across Modules and Interfaces*. Ed. by Daniel Harbour, David Adger and Susana Béjar. Oxford: OUP.

Bobaljik, Jonathan David and Susi Wurmbrand (2004). 'Anti-reconstruction effects are antireconstruction effects'. In: *Proceedings of the 2003 annual conference of the Canadian Linguistic Association*, pp. 13–25.

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

— (2005a). *In name only*. Vol. 1. Structuring Sense. Oxford University Press on Demand.

— (2005b). *The Normal Course of Events*. Vol. 2. Structuring Sense. Oxford University Press.

— (2013). *Taking Form*. Vol. 3. Structuring Sense. Oxford University Press.

Carey, Susan (1985). 'Constraints on semantic development'. In: *Neonate Cognition: Beyond the Blooming Buzzing Confusion*. Ed. by Jacques A. Mehler and Robin Fox. L. Erlbaum Associates, pp. 381–398.

Chametzky, Robert (1996). *A theory of phrase markers and the extended base*. Albany: SUNY Press.

Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

— (1981). *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris.

— (1995). *The Minimalist Program*. Cambridge: MIT Press.

Chomsky, Noam (2000a). 'Minimalist Inquiries: The Framework'. In: *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Ed. by Roger Martin, David Michaels and Juan Uriagereka. Cambridge: MIT Press, pp. 89–155.

— (2000b). *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.

— (2001). 'Derivation by phase'. In: *Ken Hale: a life in language*. Ed. by Michael Kenstowicz. MIT Press, pp. 1–52.

— (2004). 'Beyond Explanatory Adequacy'. In: *Structures and beyond. The cartography of syntactic structures*. Vol. 3. Oxford University Press, pp. 104–131.

— (2013). 'Problems of projection'. In: *Lingua* 130, pp. 33–49.

— (2015). 'Problems of projection: Extensions'. In: *Structures, Strategies and Beyond. Studies in honour of Adriana Belletti*. Ed. by Elisa Di Domenico, Cornelia Hamann and Simona Matteini. Amsterdam: John Benjamins, pp. 3–16.

Chomsky, Noam and Jerry A. Fodor (1980). 'Stating the Paradox'. In: *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*. Ed. by Massimo Piatelli-Palmarini. Cambridge, MA: Harvard University Press, pp. 259–261.

Cinque, Guglielmo (1996). 'The pseudo-relative and ACC-ing constructions after verbs of perception'. In: *Italian syntax and Universal Grammar*. Cambridge: Cambridge University Press, pp. 244–275.

Cinque, Guglielmo and Luigi Rizzi (2009). 'The Cartography of Syntactic Structures'. In: *The Oxford Handbook of Linguistic Analysis*. Ed. by Bernd Heine and Heiko Narrog. Oxford University Press.

Citko, Barbara (2005). 'On the nature of merge: External merge, internal merge, and parallel merge'. In: *Linguistic Inquiry* 36.4, pp. 475–496.

Collins, Chris and Edward Stabler (2016). 'A Formalization of Minimalist Syntax'. In: *Syntax* 19.1, pp. 43–78. ISSN: 1467-9612. DOI: 10.1111/synt.12117. URL: http://dx.doi.org/10.1111/synt.12117.

Epstein, Samuel and T. Daniel Seely (2006). *Derivations in Minimalism*. Cambridge: Cambridge University Press.

Farkas, Imola Ágnes (2011). 'Predicative Bare Nouns in Romanian Resultatives'. In: *Revue roumaine de linguistique* 56.3, pp. 277–292.

Fodor, Jerry A. (1975). *The Language of Thought*. Cambridge: Harvard University Press.

— (1980). 'Fixation of Belief and Concept Acquisition'. In: *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*. Ed. by Massimo Piatelli-Palmarini. Cambridge, MA: Harvard University Press, pp. 143–149.

— (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. A Bradford book. MIT Press. ISBN: 9780262560252.

Fodor, Jerry A. (2001). *The Mind Doesn't Work that Way: The Scope and Limits of Computational Psychology*. Bradford book. MIT Press. ISBN: 9780262561464.

— (2010). *LOT 2: The Language of Thought Revisited*. Oxford: Oxford University Press.

Folli, Raffaella and Gillian Ramchand (2005). 'Prepositions and results in Italian and English: An analysis from event decomposition'. In: *Perspectives on aspect*. Ed. by Henk J. Verkuyl, Henriette de Swart and Angeliek van Hout. Dordrecht: Springer, pp. 81–105.

Frampton, John and Sam Gutmann (2008). 'Crash-proof syntax'. In: *Derivation and Explanation in the Minimalist Program*. Ed. by S. Epstein and T.D. Seely. Generative Syntax. Wiley, pp. 90–105. ISBN: 9780470754696.

Gallego, Ángel J. (2017). 'Remark on the EPP in Labeling Theory: Evidence from Romance'. In: *Syntax* 20.4, pp. 384–399. DOI: 10.1111/synt.12139. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/synt.12139.

Gleitman, Lila (1990). 'The structural sources of verb meanings'. In: *Language acquisition* 1.1, pp. 3–55.

Harley, Heidi (2009). 'Compounding in Distributed Morphology'. In: *The Oxford Handbook of Compounding*. Ed. by Rochelle Lieber and Pavol Štekauer. Oxford: Oxford University Press, pp. 129–144.

Harwood, William (2015). 'Being progressive is just a phase: celebrating the uniqueness of progressive aspect under a phase-based analysis'. In: *Natural Language & Linguistic Theory* 33.2, pp. 523–573.

Heim, Irene and Angelika Kratzer (1998). *Semantics in Generative Grammar*. Malden, MA: Blackwell.

Hornstein, Norbert (1999). 'Movement and Control'. In: *Linguistic Inquiry* 30.1, pp. 69–96.

— (2009). *A theory of syntax: minimal operations and universal grammar*. Cambridge: Cambridge University Press.

Huang, C-T James (1984). 'On the distribution and reference of empty pronouns'. In: *Linguistic inquiry*, pp. 531–574.

Irimia, Monica-Alexandrina (2012). 'Secondary predicates'. Doctoral dissertation. University of Toronto.

Jaeggli, Osvaldo A. (1980). 'Remarks on *To* Contraction'. In: *Linguistic Inquiry* 11, pp. 239–246.

Kayne, Richard S. (1994). *The Antisymmetry of Syntax*. Cambridge: MIT Press.

Kentner, Ashley and Ronnie B. Wilbur (2018). 'Wh-clefts as evidence of resultatives in ASL'. In: *Sign Language & Linguistics* 21.1, pp. 77–144.

Klein, W. (1994). *Time in Language*. Germanic linguistics. London: Routledge. ISBN: 9780415104128.

Kratzer, Angelika (1996). 'Severing the external argument from its verb'. In: *Phrase structure and the lexicon*. Ed. by J. Rooryck and L. Zaring. Dordrecht: Kluwer, pp. 109–137.

Kratzer, Angelika (2005). 'Building Resultatives'. In: *Event Arguments: Foundations and Applications*. Ed. by Claudia Maienborn and Angelika Wöllstein.

Lasnik, Howard (1999a). 'Last Resort'. In: *Minimalist Analysis*. Malden, MA: Blackwell, pp. 120–150.

— (1999b). *Minimalist Analysis*. Malden, MA: Blackwell.

— (1999c). 'Verbal Morphology: *Syntactic Structures* meets the Minimalist Program'. In: *Minimalist Analysis*. Malden, MA: Blackwell, pp. 97–119.

Lightfoot, David (1976). 'Trace Theory and Twice Moved NPs'. In: *Linguistic Inquiry* 7, pp. 559–582.

Manzini, M. Rita and Kenneth Wexler (1987). 'Parameters, Binding Theory, and Learnability'. In: *Linguistic Inquiry* 18.3, pp. 413–444.

Marantz, Alec (1997). 'No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon'. In: *University of Pennsylvania working papers in linguistics* 4.2, p. 14.

Milway, Daniel (2015). *Directionalized locatives: A label theoretic account*. Unpublished generals paper. University of Toronto.

— (forthcoming). 'Modifying Spatial P: A remark on Svenonius (2010)'. In: *Canadian Journal of Linguistics/Revue canadienne de linguistique*.

Moro, A. (2000). *Dynamic Antisymmetry*. Linguistic Inquiry monographs. Cambridge: MIT Press. ISBN: 9780262632010.

Napoli, Donna Jo (1992). 'Secondary resultative predicates in Italian'. In: *Journal of Linguistics* 28, pp. 53–90.

Nunes, Jairo (1995). 'The Copy Theory of Movement and Linearization of Chains In The Minimalist Program'. Doctoral dissertation. University of Maryland.

— (2001). 'Sideward movement'. In: *Linguistic Inquiry* 32.2, pp. 303–344.

Pesetsky, David and Esther Torrego (2006). 'Probes, goals and syntactic categories'. In: *Proceedings of the 7th annual Tokyo Conference on Psycholinguistics*, pp. 25–60.

Piatelli-Palmarini, Massimo, ed. (1980). *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*. Cambridge, MA: Harvard University Press.

Pietroski, Paul M. (2005). *Events and semantic architecture*. Oxford: Oxford University Press.

Popper, Karl (1959). *The Logic of Scientific Discovery*. Harper Torchbooks. HarperCollins Canada, Limited. URL: `https://books.google.ca/books?id=-UbuAAAAMAAJ`.

— (2014). *Conjectures and refutations: The growth of scientific knowledge*. Routledge.

Preminger, Omer (2013). 'That's not how you agree: A reply to Zeijlstra'. In: *The Linguistic Review* 30.3, pp. 491–500.

Pustejovsky, James (1998). *The Generative Lexicon*. Bradford Books. MIT Press. ISBN: 9780262661409.

Pylkkänen, Liina (2008). *Introducing Arguments*. Cambridge: MIT Press.

Quine, Willard Van Orman (1960). *Word and Object*. Cambridge, MA: MIT Press.

Ramchand, Gillian (2008). *Verb Meaning and the Lexicon*. Cambridge, UK: Cambridge University Press.

Roberge, Yves and Nicole Rosen (2013). 'Preposition Stranding and *que*-deletion in Varieties of North American French.' In: *Linguistica atlantica* 21, pp. 153–168.

Rooryck, Johan (1996). 'Prepositions and minimalist case marking'. In: *Studies in comparative Germanic syntax*. Ed. by Höskuldur Thráinsson, Samuel David Epstein and Steve Peter. Vol. 2. Dordrecht: Kluwer, pp. 226–256.

Snyder, William (1995). 'Language acquisition and language variation: The role of morphology'. Doctoral dissertation. Massachusetts Institute of Technology.

— (2001). 'On the nature of syntactic variation: Evidence from complex predicates and complex word-formation'. In: *Language* 77.2, pp. 324–342.

— (2012). 'Parameter theory and motion predicates'. In: *Telicity, change, and state: A cross-categorial view of event structure*. Ed. by Violeta Demonte and Louise McNally. Vol. 39. Oxford University Press, pp. 279–299.

— (2016). 'Compound word formation'. In: *The Oxford Handbook of Developmental Linguistics*. Ed. by Jeffrey Lidz, William Snyder and Joseph Pater. Oxford: Oxford University Press, pp. 89–110.

Son, Minjeong and Peter Svenonius (2008). 'Microparameters of Cross-Linguistic Variation: Directed Motion and Resultatives'. In: *Proceedings of the 27th West Coast Conference on Formal Linguistics*. Ed. by Natasha Abner and Jason Bishop, pp. 388–396.

Stepanov, Arthur (2001). 'Late adjunction and minimalist phrase structure'. In: *Syntax* 4.2, pp. 94–125.

Stewart, Osamuyimen T. (2013). *The serial verb construction parameter*. Routledge.

Whelpton, Matthew (2007). 'Building resultatives in Icelandic'. In: *Proceedings of the 34th Western Conference on Linguistics*.

Williams, Edwin S (1983). 'Against small clauses'. In: *Linguistic inquiry* 14.2, pp. 287–308.

Wurmbrand, Susi (2014). 'The Merge Condition'. In: *Minimalism and beyond: radicalizing the interfaces*. Ed. by Peter Kosta et al. Amsterdam: John Benjamins, pp. 130–167.

Yokoyama, Tomohiro (2015). 'Features Wearing Two Hats: Derivation of Object Marked Verbs in Kinyarwanda'. In: *Proceedings of the 45th Annual Meeting of the North East Linguistic Society*. Vol. 3, pp. 241–250.

Zeijlstra, Hedde (2012). 'There is only one way to agree'. In: *The Linguistic Review* 29.3, pp. 491–539.