The domain of Finiteness
Anchoring without Tense in copular amalgam sentences

by

Teresa O'Neill

This manuscript has been read and accepted for the Graduate Faculty in Linguistics in satisfaction of the dissertation requirements for the degree of Doctor of Philosophy.

Marcel den Dikken

_____

Date

_____

Chair of Examining Committee

Gita Martohardjono

_____

Date

_____

Executive Officer

Marcel den Dikken

_____

William Haddican

_____

William McClure

_____

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

The domain of Finiteness
Anchoring without Tense in copular amalgam sentences

by

Teresa O'Neill

Advisor: Marcel den Dikken

The central thesis of this work is that a clause consisting only of left-peripheral functional structure can be fully finite. Generative models of clause structure typically assume that a finite clause must be tensed, including a projection of T and a temporal relation between the proposition and the utterance context. In light of evidence from tenseless languages, this assumption has come under scrutiny in recent years. This dissertation offers a new body of evidence from English, a tensed language, in support of the claim that finite clauses can lack the projection of T.

Drawing on the results of formal acceptability experiments, this dissertation presents an original investigation of the understudied family of specificational copular amalgam sentences (e.g., *She wrote about finiteness is what she did*), which differs from canonical specificational copular sentences with respect to a number of syntactic and semantic properties. The most salient of these properties is the occurrence of a root clause in the role of logical and structural subject. I propose that copular amalgam sentences are finite, but their functional spine consists only of the C-domain, lacking projections of T and V. Since C-domain heads can project in the absence of T and V, there can be no implicational relation between higher and lower heads in the functional sequence.

Copular amalgams show that finiteness can be reduced to phenomena originating in the left periphery of the clause. These phenomena include [T] and [φ] inflection, the licensing of an independently referential subject, and independent anchoring of the proposition to the utterance context. Independent anchoring, which is typically conflated with temporal anchoring in the T domain, obtains via deixis to the utterance context in finite clauses that lack T. This dissertation has two main contributions: to catalogue the properties of a typologically rare, yet understudied construction, and to challenge the Extended Projections model of the clause.

Acknowledgements

When the end of this journey seemed too far away, I used to read the acknowledgements in dissertations I admired, and (in addition to becoming a little misty-eyed), I would feel reinvigorated and inspired to pick up my writing again. Finally, here I am, writing my own acknowledgements section. Of course, I owe an enormous debt of gratitude to so many people for helping me academically, materially, and emotionally that there is no way I can thank everyone adequately. This is just a start.

First of all, I am deeply grateful to my committee members for supporting this project, and for their excellent and patient feedback. Since my first day of Syntax 1 so many years ago, I have been in awe of Marcel den Dikken's brilliant mind, encyclopedic knowledge (of perhaps everything), and crystal clear teaching style. His incredibly thorough and astute responses to questions, musings, and work in progress keep me focused and restrained. The sheer speed with which he crafts substantive and insightful replies to emails is unparalleled (and a little eerie). It is impossible to be idle under his guidance. I must also thank Marcel for introducing me to specificational sentences through his seminar on cleft sentences at the CUNY Graduate Center. I may never have stumbled upon these marvelous beasts otherwise!

Christina Tortora has been an invaluable source of support throughout this project. I thank Christina for teaching me to slow down and to be patient with the literature, the data, and myself. Her mentoring style wisely balanced free-rein to explore unorthodox ideas and a healthy amount

of restraint. Her insistence on clarity in thinking and writing has transformed my approach to scholarship and to teaching. I admire her deep respect for non-normativized linguistic phenomena and their users, which adds joy and purpose to the work of the syntactician. I will always be grateful to Christina for her confidence in me.

I thank Bill McClure for his help over the years with the semantic ingredients of my proposal, and for introducing me to the tools of model-theoretic semantics. His openness to the data and the analytical approach helped me to think freely about the puzzles I wrestled with, and his skepticism encouraged me to pursue unorthodox solutions.

Although Bill Haddican joined my committee more recently, his guidance has transformed this project. I got into this for the data, and stayed for the theory, but we all know the adage about the house built on sand! I remember my first meeting with Bill, where I explained my naïve approach to constructing surveys. Bill said, very gently, that I should probably throw it out and start over. I am so glad that he did! The methods and tools he introduced me to have bolstered the credibility and rigor of this work immensely, and have left me with a growing set of skills that I otherwise may never have developed.

Many other teachers have helped to form me as a scholar, a teacher, and a human being. At CUNY, I learned from Diane Bradley, Ricardo Otheguy, and Janet Fodor, among many other things, to write like a linguist. Gita Martohardjono introduced me to the process of doing research in her Second Language Acquisition Laboratory. I also thank Gita for taking a chance on an International Studies major and offering me the opportunity to learn linguistics at the Graduate Center. In more recent months, Samer Al-Khatib has offered indispensable guidance as I delve deeper into the semantics side of the field. Sam has been a patient and generous teacher. I am also grateful to Daniel Kaufman and Chris Collins for introducing me to linguistic fieldwork, my other passion

in linguistics (besides wacky English sentences). When I encountered Martina Wiltschko's work on endangered First Nations languages, my eyes were opened. I am grateful to Martina, not only for sharing her work on Halkomelem and Blackfoot, but also for her generous feedback and astute criticism on another project, which resonate with me still as I develop as a researcher and a writer. Finally, I thank my students, who throughout the years have taught me so much about linguistics, languages, humanity, and myself. CUNY students remain the hardest-working people I have ever met.

Without my linguist-friends, finishing a dissertation would have been far too daunting—and boring. I am lucky to have passed my time at the Graduate Center with such a brilliant, skeptical, fun-loving, supportive, and hard-working group of people. For good times, emotional support, sounding-boarding, Kraftwerk, play-dates, advice on teaching, upstate retreats, and many vigorous hours of kvetching, I thank, in no particular order, Marisa Nagano, Emily Zane, Emily Olsen, Alex Funk, Frances Blanchette, Sonya Elliot, Chris Warnasch, Katie Hawkland, Elizabeth Pratt, and Emily Wilson. Consortium classes at NYU introduced me to many more wonderful classmates, in particular, Jim Wood, Neil Myler, and Tricia Irwin. I am grateful to them for many stimulating discussions and advice on post-post-graduate life. I thank my non-linguist friends for keeping me grounded, for reminding me that there's a lot more to life than dissertating, and for providing me with a wealth of data (especially Rachel, Katie, Ben, and Sara).

I am grateful to Nishi Bissoondial for keeping our program running, for navigating the immense CUNY bureaucracy, for patiently helping me with many last-minute urgent requests, and for emotional support over the years. Likewise, I thank Maria Cadme at Queens College for her hard work, skill, and kindness.

Finally, without my family, I would never have come this far. My New York family has taught

me what generosity really looks like. The gifts of support I have received from them—financial, temporal, cultural, emotional—are immeasurable. I thank Grammy and Warren for their gifts and their love for our little family. To Denise and Jimmy, I will never be able to adequately express my gratitude. Their devotion to the wellbeing of four generations of a family is an inspiration, and a comfort during these hectic months of writing. I am grateful to my parents and my siblings (and siblings-in-law) for *everything*—but especially for their gift of beautiful and peaceful spaces to write in, for taking care of Mimi, for their faith in me, and for their love. A special thanks to Mimi for reminding me to focus on the things that matter, and for showing me how to find joy in the small things.

There should be a special award for dissertation spouses. Throughout this process, Edmund has been a champion. He has lived through the last few hectic months of writing with cheerfulness, patience, and grace. As generous breadwinner, skilled homemaker, devoted father, and loving husband, he has kept me whole throughout this often stressful time. I cannot thank him enough. I am happy to be home again.

For Edmund and Miriam

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1   Introduction

The aim of this work is to bring a new empirical domain to bear on a very old question: what makes a sentence a sentence? I address this question from the perspective of generative syntax. What are the basic *formal* characteristics of an independent clause? The two *functional* ingredients of independence correspond quite conveniently to the labels of the left-peripheral functional heads that encode them in some Minimalist models of clausal architecture: Fin(iteness) and Force. Roughly, a clause is *finite* if it includes a direct relation (associated with Fin) between an eventuality under discussion and an utterance context; a clause is independent if it includes a relation (associated with Force) between a proposition, an utterance context, and a speaker. A clause with both of these properties, I will call "fully finite".

These two properties are necessary and sufficient to make a clause independent. Perhaps conspicuously absent from the set of criteria for full finiteness is Tense/tense. The concepts of tense and finiteness are often conflated, with the syntactic projection of T(ense) and/or the semantic contribution of tense treated as a prerequisite for full finiteness. Standard generative models of the clause likewise assume an implicational relation between Fin and T (e.g., Grimshaw 1991, 2000; Chomsky 1995, 2001, 2008; Rizzi 1997, 2004; Cinque 1999). I argue in this work that a fully finite

clause can in fact be projected *directly from Fin*.

Evidence that a fully finite clause can lack the functional structure below Fin is somewhat elusive. The assumption that clause structure requires Tense has been called into question by the existence of languages that lack grammatical tense distinctions (e.g., Yucatec Maya, Bohnemeyer 2002; Kalaallisut, Shaer 2003; Bittner 2005; Mandarin Chinese, Lin 2010; Guraní, Tonhauser 2011; Halkomelem Salish and Blackfoot, Ritter and Wiltschko 2014, etc.). These languages manage to generate fully finite clauses without grammatical tense. Instead, they either substitute some other category (Ritter and Wiltschko 2014), rely on context (Tonhauser 2011), or perhaps have a morphologically unrealized Tense head in their structure (Matthewson 2006). In each of these analyses, finiteness depends on some other structural and semantic component. The minimal amount of structure constituting a fully finite clause seems to include Force, Fin and some other category. How can the contribution of the Finiteness domain be isolated?

Tensed languages like English clearly have clause types that are not inflected for tense. In imperatives, for example, the verb form must be bare, and the temporal interpretation of the proposition is restricted to the future.

(1)   a.   Be nice!

      b.   *Are nice!

      c.   *Be nice yesterday!

The presence of T can nevertheless be detected by functional elements like negation. The special morphological realization of T in imperatives is conditioned by the features associated with Force. Once again, the contribution of Finiteness cannot be severed from the contribution of Tense. To isolate the contribution of Finiteness, we must identify a *declarative* clause type in a tensed language like English that lacks the properties associated with the T(ense)-domain of the clause, but

satisfies the Fin and Force-based criteria for full finiteness.

The copular "amalgam" is just such a clause type.[1] The most common instantiations of the copular amalgam are illustrated in (2)–(4), below.

(2) a. She's demonstrating the copular amalgam is what she's doing.

b. She's demonstrating the copular amalgam is the idea.

(3) a. What she's doing is she's demonstrating the copular amalgam.

b. The idea is she's demonstrating the copular amalgam.

(4) That's what she's doing is she's demonstrating the copular amalgam.

Copular amalgams have the peculiar property of featuring a root-like clause in the semantic and syntactic environments where a "canonical" clause features a nominal expression. In (2) and (4), a root clause serves as the apparent structural subject of the finite copula, which is impossible in other contexts, e.g., (5).

(5) a. *They finished their homework is surprising.

b. *They finished their homework surprised me.

Although they are abundant in spoken English, copular amalgams have gone relatively unnoticed in contemporary generative linguistics.[2] While some copular amalgam sentence types are mentioned in previous literature, sometimes in conjunction with the claim that they are ungrammatical, they have never received a unified syntactic analysis. I account for a collection of unique syntactic properties of copular amalgam sentences by proposing that their functional structure

---

[1]The label "amalgam" for these structures, which I will use here, is due to Declerck (1988) (and subsequent discussion in Lambrecht 2001, and Lambrecht and Ross-Hagebaum 2006). While I will argue that these sentences are *not* syntactic amalgamations, I use the term to maintain a connection to this earlier literature.

[2]Exceptions include Ross (1972, 2000), Massam (1999, 2013), den Dikken et al. (2000), and Schlenker (2003). In the discourse literature, particularly the work of Lambrecht (2001), they have been examined more closely.

consists entirely of the Finiteness domain of the clause (CP), with no lower projections of T or V. Copular amalgams thus provide the missing structural evidence that Fin does not depend on T.

The absence of T is not without consequences: its ability to "survive" as fully finite without T is closely related to the marked property of having a root clause as a subject. The structure can forego the T projection because its root clause subject does not enter the (nominal) grammatical subject licensing relation, and its proposition is anchored via a deictic relationship to the utterance context. A close look at precisely *how* copular amalgams instantiate the criteria for finiteness in the absence of T gives us a clearer picture of the function of Fin in clause structure.

1.2 Copular sentences: back to basics

Copular sentences provide an excellent laboratory for questions about functional structure. A lexical verb introduces complications from argument structure and aspect, but a copular sentence is stripped down to essentials: subject, predicate, and functional structure. All but the most fundamental components of functional structure can be filtered out.

A copular clause like (6) satisfies the criteria for full finiteness in the absence of a Verb, but it does require T.

(6)   The children were happy.

In this sentence, T is a key player in establishing the anchoring relation. The eventuality (a state) under discussion is located in the past with respect to the reference context. The T-domain of the clause also encodes the grammatical subject relation, marking [φ]-feature agreement and providing a landing site for A-movement. This inflectional information in turn depends on the presence of a positively specified Fin head, which introduces a deictic eventuality corresponding to the utterance context, which determines the reference time. The speaker at the utterance context commits herself

to its truth (via an operator in Force) (for different formalizations, see Klein 1998; Meinunger 2006; Krifka 2014).

Copular sentences come in a variety of flavors. Predicational sentences ascribe properties to their subjects, and the copula itself makes no semantic contribution. Its presence is motivated purely for formal reasons. For instance, when grammatical relations and finiteness functions are encoded in a higher clause, the predication relation can be expressed with no overt functional material at all.

(7)   I consider [the children happy].

The copula also occurs in equatives, where two expressions stand in an identity relation.

(8)   His favorite city is your favorite city.

Specificational copular sentences, e.g., (9), which are the focus of this work, share some properties with both predicational and equative sentences.

(9)   His favorite city is Paris.

The function of the specificational sentence is to *specify* the content of an expression. A less referential expression, *his favorite city*, receives a value from a referential expression, *Paris*. I will adopt a model of specificational copular sentences that treats them as (inverse) predications (Heggie 1988; Moro 1997, 2000; Mikkelsen 2005; den Dikken 2006). They share with equatives the property that their less referential expression is intensional (Schlenker 2003; Romero 2004, 2005; Comorovski 2008). The copular element adds no encyclopedic meaning to the copular sentence; its presence is required in certain contexts for formal reasons.

The copular amalgam presents both a puzzle and an opportunity. It exhibits the same basic

configuration and interpretation as the *canonical* specificational copular sentence, e.g., (9), but its syntactic and semantic components are quite different. Where the canonical sentence features a DP (often a free relative), the amalgam features a root-like clause. Where the canonical clause's logical subject is a referential expression, the amalgam clause's subject is a proposition.

(10)   a.   What he likes is [craft beer].

         b.   What he likes is [he likes craft beer].

The subject-initial specificational sentence has a close structural counterpart whose post-copular expression is an ordinary one-place predicate, (11); the amalgam specificational sentence does not, (12).

(11)   a.   Craft beer is [what he likes].

         b.   Craft beer is [flavorful / pretentious].

(12)   a.   He likes craft beer is [what he likes].

         b.   *He likes craft beer is [flavorful / pretentious].

Given that bare finite clauses cannot function as ordinary structural subjects, what is the syntax of the copular amalgam sentence?

1.3   Proposal

The proposed structure for copular amalgams takes their surface bisentential form seriously. The copula in amalgams links a root clause and a predicate of propositions (an indirect/concealed question *wh*-clause or DP). These types fit together naturally, like the nominal subject and predicate of individuals in a canonical clause. A root clause subject lacks the requisite feature composition to Agree with T and displace to a derived subject position, however, so it composes with its predicate

in situ: in a small clause. Unlike an ordinary small clause, however, the copular amalgam clause is inflected for finiteness ([tense] and [ɸ]).

(13) a. [He likes craft beer] is [what he likes].

  b. *[He likes craft beer] be [what he likes].

  c. *[He likes craft beer] ∅ [what he likes].

Also unlike a small clause, the copular amalgam is fully finite: the speaker commits herself at the utterance time to the truth of the proposition: that the root clause specifies the content of (or answer to) the *wh*-clause at utterance time. The copular clause, with the basic structure in (14), manifests the functions of Fin and Force.

(14)

ForceP
├── Op-ASSERT
└── Force′
    ├── Force
    └── FinP
        ├── CP
        │   └── He likes craft beer
        └── Fin′
            ├── Fin
            │   └── is
            └── CP
                └── what he likes
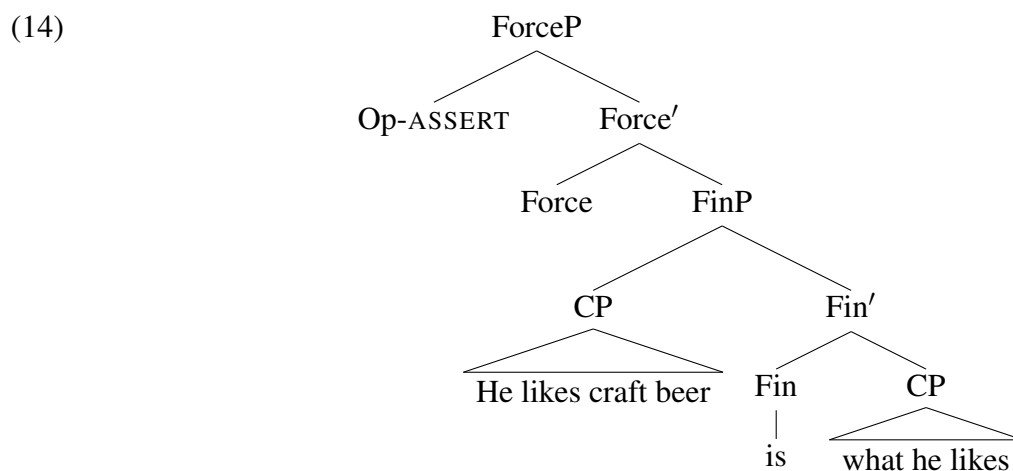
  The absence of the T-domain is motivated not only by theory-internal considerations about Agree and structural subject-hood, but also on empirical grounds. The copula in amalgams cannot combine with any material associated with the T-domain, other than simple [tense] and [ɸ] inflection, which actually originate in the C-domain. The major constituents of the copular amalgam cannot undergo ordinary A′-movement, since they are base-generated in information structure-sensitive A′-positions. Finally, the copular amalgam clause fails to express the range of temporal interpretations the T-domain makes available. Regardless of the tense *form* of the copula, the

proposition is anchored deictically to the utterance context.

I propose that the eventuality under discussion in the copular amalgam is not a state anchored with respect to the utterance context, but rather, the utterance context itself. In a clause whose sole participants are discourse units and utterance contexts, only the structural domain of discourse relations and the utterance context is projected. The anchoring relation identifies the utterance context of the embedded (subject) root clause with the utterance context of the matrix (copular) clause. From a pragmatic perspective, this structure makes copular amalgams particularly useful for "talking about talking".

1.4   Roadmap

This thesis is divided into three parts.

1.4.1   Part 1

Part 1 has an empirical focus. Chapter 2 introduces the form and function of specificational copular sentences. I summarize the predicate inversion approach to specificational copular sentences, which treats the copula itself as devoid of lexical content. Chapter 3 then describes my methodological approach to amalgam copular sentences. Since amalgams are prescriptively stigmatized and exhibit robust inter-speaker and intra-speaker variation, sources of data for this thesis include traditional acceptability judgment consultations, corpora, and crowd-sourced, controlled acceptability judgment surveys.

1.4.2   Part 2

Part 2 turns to the analysis of copular amalgam sentences, particularly in comparison to their canonical counterparts. Chapter 4 describes the core syntactic and semantic properties of the copular amalgam. This chapter presents the findings of a series of acceptability judgment surveys that

illustrate the crucial contrasts between the canonical and amalgam specificational sentence types. An additional empirical contribution of this chapter is a clarification of the empirical status of canonical pseudoclefts in a variety of syntactic environments. Chapter 4 also discusses a semantic contrast between the canonical and amalgam sentence types. The canonical type expresses a relation between an entity-type expression and an intensional predicate of individuals, while the amalgam expresses a relation between a proposition-type expression and an intensional predicate of propositions, i.e., a question.

What sort of syntax underlies a clause relating a question and its answer? Chapter 5 develops an analysis of amalgam clauses, based on the findings of Chapter 4. The analysis relies on a model of the clause that is partitioned into three (or four) domains, each of which is dedicated to a distinct abstract function (e.g., Hale 1986; Chomsky 1986, 2001; Abney 1987; Pollock 1989; Grimshaw 1991; Grohmann 2003; Sigurðsson 2004; Borer 2005; Wiltschko 2014, among others). These functions and their formal realizations in English diagnose the base-merged and spell-out position of functional elements like the copula. The copula in canonical specificational sentences associates with the domain dedicated to grammatical relations, agreement, and (temporal) displacement. The copula in amalgams, however, associates with the domain dedicated to discourse relations, force/clause-typing, and deixis to the utterance context. The unusual properties of the amalgam clause follow from the absence of the T-domain found in ordinary clauses.

1.4.3 Part 3

Part 3 examines the implications of the proposed clause type for the theory of finiteness. Chapter 6 introduces the phenomena associated with finiteness and gives a brief history of finiteness in generative grammar. It lays out the main insights surrounding the so-called tense-case connection, the restriction of overt nominative case-marked subjects to tensed clauses. The traditional tense-case

connection has fallen under renewed scrutiny in recent decades, as the empirical landscape turns out to be more complex than what earlier theories can handle. If finiteness does not entail tense, and nominative subjects are not restricted to tensed clauses, then what natural class of properties characterizes finiteness cross-linguistically? It turns out that finiteness cannot be given a mono-lithic characterization. Finiteness is *gradient* from both a bottom-up and a top-down perspective. A *maximally* finite independent clause includes Force, Fin, and T. Clauses can decrease in finite-ness in *either* direction: a clause may be finite in the traditional sense, but lack independent Force; or a clause may be independent, but lack distinct temporal reference.

Chapter 7 takes up the puzzle of what allows a bare finite clause to serve as a subject just in the amalgam copular sentence type and nowhere else. The proposal is that unlike an embedded clause, which features an element of T moved to C, the highest structural domain of the root clausal subject lacks any element belonging to the T-domain. The root clause therefore lacks the dependent form that requires other arguments to be grammatically licensed. Moreover, since the root clause lacks [φ]-features, it cannot enter an agree relation with T/Subj. The root clausal subject of an amalgam consists of a fully specified, deictic left-peripheral domain, like the matrix Fin head it agrees with.

Finally, Chapter 8 argues that the copular amalgam clause lacks not only the syntactic projec-tion of T, but also semantic tense. Since the copular amalgam clause consists only of the deictic domain, it lacks the temporal argument structure necessary for temporal displacement. Because the tense properties of specificational copular sentences are puzzling in their own right, this chap-ter develops an analysis of temporal semantics in canonical pseudoclefts. This analysis provides a baseline for analyzing the tense properties of amalgams. Then, a comparison of the two sen-tence types, including an interpretation experiment, shows that the copular amalgam clause lacks a time argument. The copular amalgam clause is instead anchored deictically to an indexical el-

ement of the utterance context. Deictic anchoring is precisely what gives amalgams their curious

speech-act-within-a-speech-act interpretation.

PART 1: The Data

It is what it is, is what it is.

Chapter 2

Overview of copular amalgams

2.1   Variation in copular clauses

Copular sentences have long fascinated linguists, traditional grammarians, and philosophers of language. In a copular sentence, the main predicate is not headed by a Verb, but by some other category, typically an Adjective, Noun, or Preposition.

(1)   a.   That girl is tall.

       b.   That girl is a student.

       c.   That girl is in my office.

Copular sentences raise fundamental questions about the nature of lexical categories (e.g., what makes a verb a verb?) and the role of functional elements in grammar (e.g., why do they exist at all?). The copula itself is at the center of these questions. While copular sentences express a cross-linguistically stable set of meanings, copular elements vary in form (see Pustet 2003 for a detailed cross-linguistic survey). They are sometimes covert, sometimes overt, sometimes verbal, sometimes nominal, sometimes inflected, sometimes uninflected. The first part of this dissertation locates the amalgam copula in this diverse landscape.

While Standard American English requires the copula to be expressed overtly in finite clauses, many languages do not. For example, Chol (Mayan) relates subject and predicate directly, without the mediation of an overt copula.

(2)   a.   Chañn jiñi wiñik.
   tall   DET man
   'The man is tall.'

   b.   Maystraj jiñi wiñik.
    teacher   DET man
    'The man is a teacher.'   Chol

    (Coon 2013: (4))

Russian (Partee 2000; Geist 2007), Hebrew (Rapoport 1987; Doron 1986), Arabic (Benmamoun 2008), African American Vernacular English (e.g., Labov 1995), and many other languages have a zero copula only in present tense predicational sentences.

(3)   a.   On student.
   he student
   'He is/*was a student.'

   b.   On byl student.
    he was student
    'He was a student.'   Russian

(4)   a.   She a student.
   'She is/*was a student.'

   b.   She was a student.
    'She was a student.'   AAVE

Semantic differences, like individual vs. stage-level predication (Spanish) or locative vs. adjectival or nominal predication (Bantu; McWhorter 2012) can also condition the form of the copula.

Another common point of language-internal variation is the category of the copular element. In Hebrew, Russian, and Irish (Carnie 1995, 1997), for example, a pronominal element serves as the copula in some contexts. These languages also share with Haitian Creole (DeGraff 1992; Déprez 2003) the property that the distribution of an overt copula is conditioned by the referentiality of the predicate (e.g., definite, pronominal, proper name, etc.).

(5)   a. Dani (hu) ha-more.
           Dani (3SG) the-teacher
           'Dani is the teacher.'

        b. Ha-more   *(hu) Dani.
           the-teacher *(3SG) Dani
           'The teacher is Dani.'   Hebrew

           (Doron 1986)

(6)   a. M. Twain (–*eto) pisatel'.
           M. Twain (*this) writer.NOM
           'M. Twain is a writer.'

        b. M. Twain *(–eto) S. Clemens.
           M. Twain *(this) S. Clemens
           'M. Twain is S. Clemens.' Russian

           (Geist 2007: (31)–(32))

(7)   a. Is captaen (é)   Seamus.
           COP captain (3SG) Seamus
           'Seamus is a captain.'

        b. Is *(é)   captaen Seamus.
           COP *(3SG) captain Seamus
           'The captain is Seamus.'   Irish

           (Carnie 1997:15)

(8)   a. Jan (*se) malad.
           Jan (*COP) doctor
           'Jan's a doctor.'

        b. Jan *(se)   yon dokte.
           Jan *(COP) the doctor
           'Jan is the doctor.' Haitian Creole

Relevant to the present work is the fact that there is often a diachronic or synchronic relationship between copulas and focus particles. For example, in Haitian Creole and Garifuna, the same form serves as a focus marker and as a copula, which is obligatory in specificational contexts.[1]

(9)   a. Úruwei Obama.
           ruler   Obama
           'Obama is the president.'

        b. Obama úruwei   *(ba-i).
           Obama ruler-FOC 3SG.M
           'The president is Obama.' Garifuna

The formal variation among copulas, both within and across languages, provides a useful testing ground for hypotheses about clause structure. When a sentence lacks a lexical verb, how do subject and predicate combine? What functional structure must be projected? By examining the behavior of the copula in different environments, we hone our understanding of the functional structure of the clause.

---

[1]Unless otherwise noted, Garifuna examples are from my own fieldwork.

The present thesis tackles a relatively small piece of the puzzle, but it is a piece that has received very little attention, even though it occurs in English, the most exhaustively analyzed language in modern linguistics. It examines the copula as a relator of propositions, (10).

(10)   [That's what I did] is [I called him right back].

When the clause is stripped down to its most basic parts, the formal properties of the copula reveal the structure holding the clause together.

2.2   Classifying copular sentences

Given the formal variation in copular clauses, taxonomical questions figure prominently in generative research on copular sentences: (i) How many basic types of copular sentence are there? (ii) What distinguishes these types? (iii) To what extent are they derivationally related? and (iv) How many copulas are there? Classifying copular sentence types is not a straightforward task, since different natural classes emerge depending on whether the focus is on syntactic behavior, semantic composition, or information function (Higgins 1979; Akmajian 1979; Williams 1983; Partee 1986, Partee 2000; Heggie 1988; Declerck 1988; Heycock 1995, 2012; Moro 1997; Heycock and Kroch 1999; Sharvit 1999, Sharvit 2003; den Dikken et al. 2000; Ross 2000; Lambrecht 2001; Schlenker 2003; Romero 2005; Mikkelsen 2005, Mikkelsen 2011; den Dikken 2006, among many others). In this thesis, I adopt a syntactic-semantic classification that derives specificational sentences from predicational small clauses.

Let us begin with the four-way classification proposed by Higgins (1979).

(11)   a.   John is a pleasant guy.   predicational

         b.   The teacher is John.   specificational

      c.   Your teacher is my teacher.   equative

      d.   That (man) is my teacher.   identificational

The predicational sentence type (11a) is the most basic: it simply ascribes some property to its subject. A specificational sentence (11b), on the other hand, is not a property-ascription. It resolves anaphora, presenting a gap in one constituent and filling it in the other. Equatives (11c) and identificational copular sentences (11d) are both identity statements, as their label suggests. In an equative, the two expressions are referential, while in an identificational sentence, one of the terms includes a demonstrative element.

At the center of the taxonomical enterprise is the specificational copular sentence, because it bridges the typological gap between the more superficially distinct predicational and equative sentence types. Specificational copular sentences are characterized by two constituents: one introducing a variable, and the other providing a value for that variable. (For the purpose of the present exposition, I will use the atheoretical labels XP and YP to refer to the precopular and postcopular constituent, respectively.)

In the prototypical specificational sentence, both XP and YP are DPs.

(12)   a.   My favorite drink is coffee.

         b.   The tallest girl in school is Mary.

In pseudoclefts, a widely studied type of specificational sentence, the expression with the gap is a *wh*-clause (13), and the constituent on the other side of the copula can be of any category.[2]

(13)   a.   What I need is a coffee.   DP

         b.   What she is is kind.   AP

---

[2]DP is the most common; there is some variation in the acceptability of different categories.

c.	What I did is leave early.	VP
d.	Where he went is to the store.	PP
e.	How he did it is quickly.	AdvP
f.	What he meant was that he wanted to leave.	CP
g.	What they wanted was for her to leave.	CP
h.	What I want is you fired.	Small Clause

The form of the pseudocleft suggests a similarity to question-answer pairs. Just as in a question-answer discourse, one constituent is a *wh*-clause, and the other is a focus supplying the value for the *wh*-bound variable. Many authors (e.g., Faraci 1971; Huddleston 1971; Ross 1972; Higgins 1979; Akmajian 1979; Declerck 1988; den Dikken et al. 2000; Schlenker 2003; Romero 2005, etc.) have explored the self-answering question function of the specificational pseudocleft.

The question-answer pair analogy will turn out to be central to the structure I propose for *amalgam* specificational copular sentences, e.g., (14).

(14)   What I need is I need a coffee.

Copular amalgams are not only functionally similar to question-answer pairs; they are actually composed of sentential constituents, like a question-answer discourse.

(15)   What do I need?

          I need a coffee.

2.2.1   Specification and predication

In order to characterize the specificational copular sentence in its canonical and amalgam guises, it is important to establish its relationship to the more basic predicational sentence type. Spec-

ificational and predicational copular sentences share a number of properties. Both consist of an

asymmetry between a referential constituent and a constituent that is less referential.[3]

(16)   John is nice.   predicational

(17)   The tallest guy is John.   specificational

In predicational copular sentences like (16), there is an obvious asymmetry between the XP and

YP with respect to their semantic type: YP, the predicate *nice*, ascribes a property to XP, the subject

John. Specificational sentences, exemplified in (17), exhibit the same asymmetry, albeit somewhat

more subtly and in reverse linear order: YP is referential, and XP is predicative. Because of the

reversal in the linear order of the major semantic constituents, I will henceforth use labels that refer

not to their linear position, but to their role as either (logical) subject (*John*) or predicate (*the tallest*

guy) (Williams 1983; Partee 1986; Heggie 1988; Mikkelsen 2005; den Dikken 2006).

2.2.1.1   Reversibility

Like ordinary verbal clauses, predicational clauses display an unmarked subject-predicate order.[4]

Specificational copular sentences, by contrast, stand out precisely because of their reversibility. In

their prototypical form, they show the marked predicate-initial order.

---

[3]The precise way in which one constituent in a specificational sentence is "less referential" is open to debate (Akmajian 1979; Higgins 1979; Declerck 1988; Heycock and Kroch 1999; Schlenker 2003, etc.). These varied approaches agree that there is a referentiality asymmetry between the two constituents, but they characterize its semantics differently. I follow Schlenker (2003), Romero (2004, 2005), Comorovski (2008) in formalizing the intuition in terms of intensionality—the less referential expression in a specificational sentence is world- and time-dependent. This property will be discussed further in Chapter 4 and Chapter 8.

[4]While it is possible to reverse the linear order of the subject and predicate in a predicational copular sentence, this process is extremely marked, and can occur only in special contexts. First of all, the predicate must be an indefinite description, and second of all, it must be interpreted as a continuing topic, as in (i). It is unclear whether reverse predications like (i) should be treated as a type of specificational sentence, where the postcopular expression gives an example of the precopular description. I will not discuss the restrictions on reversibility in simple predicational copular sentences further here.

(i)   a.   Can you give me an example of a doctor?
        b.   Yes, a doctor is John.

(18)   The doctor is John.

When the order of subject and predicate is reversed, the specificational sentence is linearly indistinguishable from a predicational sentence.

(19)   John is the teacher.

What makes (19) "specificational" is its information function. *John* is focused, and provides a value for a variable in the predicate. If *John* serves as a topic, and the sentence's function is to ascribe some property to John, then it is predicational.[5] This ambiguity ultimately stems from the semantic type of the predicate: if it is a simple one-place predicate, the sentence is predicational; if it is an intensional predicate, the sentence is specificational.

(20)   a.   John is the teacher$_{<e,t>}$.

         b.   John$_F$ is the teacher$_{<s,<e,t>>}$.

I will refer to specificational sentences in subject-initial order (20a) as "reverse" specificational sentences. Even though subject-initial order is unmarked in other contexts, the predicate-initial order is more basic in specificational contexts. This is a surface description, implying nothing about the sentences' syntactic derivation.

Even without reversing the positions of the predicate and subject in a specificational sentence, ambiguity often occurs (Akmajian 1979; Higgins 1979). Consider the ambiguous pseudocleft in (21).

(21)   What she saw was a man and a woman.

---

[5]I adopt the following typographical convention: "topic" with a lowercase 't' refers to an information structural status; "Topic" with an uppercase 'T' refers to a syntactic position above Fin in the A′-domain.

Under a predicational "hermaphrodite" interpretation of (21), the precopular constituent is a referential expression, which is ascribed the property of being both male and female. On the specificational reading, *what she saw* is a predicate, and *a man and a woman* specifies its members at the evaluation world. The syntactic ambiguity rests on which of the two constituents is construed as the underlying subject, and which as the predicate.

2.2.1.2 Connectivity effects

The well-known examples in (23), observed in Higgins's (1979) seminal work, illustrate a notable property of specificational sentences that distinguishes them from predicational sentences: so-called "connectivity effects".[6]

(22) a. What John$_i$ is is important to him$_i$. predicational

    b. What John$_i$ is is important to himself$_i$. specificational

The value in specificational sentences, particularly pseudoclefts, enters a range of unexpected syntactic and semantic dependencies with an element inside the predicate. These dependencies are unexpected because they obtain in the absence of a c-command relationship: precisely those dependences that would obtain straightforwardly in the simple-clause counterpart of the specificational sentence.

(22a), which does not exhibit bound anaphor connectivity, is predicational: 'the thing that John is (e.g., his profession) has the property of being important to him.' In (21b), however, the anaphor

---

[6]See discussion of "connectedness" in Higgins (1979) and Akmajian (1979). Connectivity effects are discussed widely throughout the copular sentence literature, and lie at the heart of many of the debates about the syntax and semantics of specificational sentences (e.g., Ross 1972; Heggie 1988; Jacobson 1994; Bošković 1997a; Heycock and Kroch 1999; Sharvit 1999; den Dikken et al. 2000; Schlenker 2003; Caponigro and Heller 2007; Sharvit 2009). Connectivity effects pose no mystery in copular amalgam sentences, however, since their logical subject is a complete clause, where any dependent element is locally licensed (den Dikken et al. 2000).

(i) a. What John is is John is important to himself.
    b. John is important to himself is what John is.

himself is bound, even though it is not locally c-commanded by its binder, *John*. (22b) is specifi-

cational, meaning: 'the property x such that John has x is the following: important to himself.' The

effect is even more striking in the reverse pseudocleft.

(23)   Important to himself$_i$ is what John$_i$ is.

2.2.1.3   Optionality of the non-finite copula

Predicational and specificational copular sentences differ in some non-finite contexts with respect

to the obligatoriness of the copular element. The predicate and the subject in simple predicational

sentences can compose directly, via function application. In non-finite contexts, where the small

clause combines with a superordinate verb, an overt copula is often not required. (Of course, Stan-

dard English requires an overt copula in finite contexts.)

(24)   a.   I find [John interesting].

         b.   That would make [John sad].

         c.   I wouldn't want to go, with [John in such a bad mood].

Specificational sentences, especially in their subject-initial order, similarly allow a zero-copula in

some non-finite environments.

(25)   a.   I consider [John the best man for the job].

         b.   That would make [John the murderer].

In predicate-initial order, however, a non-finite form of the copula is typically required.

(26)   a.   I find [the best man for the job *(to be) John].

         b.   I wouldn't want to apply, with [the best man for the job *(being) John].

2.2.1.4 Extraction

Another difference between the two copular sentence types is the availability of A'-movement. Predicational sentences are quite free in this respect.

(27) a. Who$_i$ did she say t$_i$ was nice?

  b. Nice$_i$, Mary was t$_i$; intelligent$_i$, she was not t$_i$.

Specificational sentences, by contrast, are much more restricted, once again particularly in their predicate-initial order.

(28) a. *[What kind of candidate]$_i$ did she say t$_i$ was John?

  b. *John$_i$, the best man for the job certainly was t$_i$.

2.2.1.5 Information structure

Specificational sentences are also more restricted with respect to their information structure. In unmarked contexts, the subject in English serves as the topic by default, but it can also be a focus. The predicational copular sentence structure imposes no particular information structural constraints.

(29) A: Give me an example of a great guy.

  B: [JOHN]$_F$ is a great guy.

(30) A: Phil is a great guy.

  B: No, [JOHN]$_F$ is a great guy.

(31) A: What can you tell me about John?

  B: John is [a great GUY]$_F$.

(32) A: I think John is nice.

B: Yes, John is [a great GUY]$_F$.

Specificational sentences have a fixed information structure, on the other hand. The logical subject is always narrowly focused and the predicate is always a topic (Declerck 1988; Heggie 1988; Heycock and Kroch 1999; Lambrecht 2001; Mikkelsen 2005; den Dikken 2006). Mikkelsen (2005) argues that the marked predicate-initial order of a specificational sentence is licit, because it preserves the unmarked Topic-Comment information structure of ordinary clauses.

(33)   A: Who's the teacher?

    B: The teacher is [JOHN]$_F$.

(34)   A: What can you tell me about John?

    B: #The teacher is [JOHN]$_F$.

2.2.2   Specification and equation

The reversibility of specificational sentences hints that they are formally symmetrical to an extent that distinguishes them from predicational sentences (see Heycock and Kroch 1999). Consider the syntactic category and logical type of their subject and predicate. As we have seen, it is common for both of the expressions flanking the copula in specificational sentences to be DPs. The meaning of specification entails that the referents of these two expressions are the same. These properties suggest a similarity between specificational copular sentences and the more overtly symmetrical types: equatives and identificational copular sentences.

In equatives, both constituents seem to refer directly to individuals. At a superficial level of analysis, there is no syntactic or semantic asymmetry between the two expressions. Equatives can even be tautologies.

(35)   a.   Cicero is Tully.

        b.   My favorite city is your favorite city.

        c.   Honest is honest.

Rothstein (1995) and Heycock and Kroch (1999) point out that both terms in an equative-like spec-

ificational sentence can undergo non-restrictive modification, which is compatible with referential

expressions, but generally not with predicates:[7]

(36)   a.   The duty nurse, who is very efficient, is Rina, who I am very fond of.   (Rothstein

              1995: (45))

        b.   *I consider Rina the duty nurse, who is very efficient. (Heycock and Kroch 1999:374

              (32a))

Like specificational sentences, equatives are reversible.

    Equative sentences show similar syntactic restrictions to specificational sentences. For instance,

when embedded under ECM predicates like *consider*, the overt non-finite copula is required.

(37)   I expect your favorite city *(to be) my favorite city.

(38)   I expect the best man for the job *(to be) John.

Extraction out of both specificational and equative sentences is highly restricted.

---

[7]I will argue in Chapter 4 that the predicative expression in a specificational sentence is intensional; that is, I do
not treat specificational sentences as equative relations between two referential individuals. While I will not discuss
the modification test in detail here, I am not convinced that it shows that specificational sentences must be equative.
Consider, for example, the reverse version of (36a), which is specificational in (i).

(i)   A: Do you know who the duty nurse is?
         B: *Yes, [RINA]$_F$, who I am very fond of, is the duty nurse, who is very efficient.

If specificational sentences are equative, then *the duty nurse* should be referential, so the unacceptability of (i) is
unexpected. The asymmetry between the two orders argues against treating specificational sentences as equatives.

(39)   A: *Who/what$_i$ did he say t$_i$ was John?   specificational

         B: —The best man for the job.

(40)   A: *What$_i$ did he say t$_i$ was his favorite city?   equative

         B: —Her favorite city.

These examples illustrate that extracting the precopular constituent fails—in the specificational

case, this is an illicit extraction of a topical predicate.

    The two sentence types do not pattern entirely together, however. In equatives, either term can

serve as the focus, depending on context, while in specificational sentences (recall (33) and (34),

above), the logical subject must be focused.

(41)   A: Did you know that the north star is the (same as the) morning star?

         B: No, the morning star is the EVENING star.

         B′: No, the EVENING star is the morning star.

(42)   A: Did you know that the north star is the (same as the) evening star?

         B: No, the MORNING star is the evening star.

         B′: No, the evening star is the MORNING star.

2.2.3   Derivational relationships between copular sentence types

The preceding sketch showed that specificational sentences share some properties with predica-

tional sentences, and some with equatives (Table 2.1).

    The split in the behavior of specificational sentences has inspired three main approaches to

analyzing their syntax.[8] In the first, specificational sentences are treated as a type of predicational

---

[8]For a rich review of the literature on specificational copular sentences, see den Dikken (2005a). Since my present
purpose is not to reanalyze the canonical specificational copular sentence, but to establish a baseline for analyzing the
understudied copular amalgam, the review of this literature given here is brief.

	Predicational	Specificational	Equative
Reversible	✗	✓	✓
XP and YP same category	✗	✓	✓
ECM+zero copula	✓	✗	✗
A′-movement unrestricted	✓	✗	✗
Flexible information structure	✓	✗	✓

Table 2.1: Properties of specificational copular sentences and their relatives

sentence where the two major constituents are either base-generated in reverse order (e.g., Williams 1983; Mikkelsen 2005) or invert in the course of the derivation (e.g., Heggie 1988; Moro 1997; den Dikken 2006). In the second, specificational sentences are treated as a type of equative (Heycock and Kroch 1999 give the most fleshed out version of this proposal). In the third, each type of copular sentence has a distinct structure (Higgins 1979; Akmajian 1979). In what follows, I summarize arguments supporting the first approach, since it provides the main underpinnings for my analysis of copular amalgam clauses.

2.2.3.1 Predicate inversion

There is no need to maintain a deep distinction between predicational and specificational sentences. The differences between the two sentence types have been fruitfully analyzed as following from a conspiracy of syntax and information structure (Heggie 1988; Moro 1997, 2000; Mikkelsen 2005; den Dikken 2006). These authors argue that specificational copular sentences are predicational sentences with a marked syntax in which the predicate occupies the precopular position. This markedness is behind the increased restrictions on specificational sentences, especially when it comes to A′-extraction and embedding. These authors reach different conclusions about the precise derivation of copular sentences, so I will briefly present their main points.

Heggie's (1988) analysis derives specificational sentences from predicational small clauses, via inversion of the predicate into an A′ (Topic) position. Since the predicate occupies a Topic position,

specificational sentences exhibit frozenness. A-movement is impossible, and A′-movement has no potential landing site above the Topic position that would not give rise to an information structure clash.

For Moro (1997, 2000), both predicational and specificational sentences originate as symmetrical small clauses. Because this symmetrical structure cannot be linearized, assuming an algorithm that depends on antisymmetric relationships (Kayne 1994), either the predicate or the subject must dislocate to create an asymmetrical configuration. This movement lands it in the structural subject position of the clause.

Mikkelsen (2005) and den Dikken (2006) side with Moro in proposing that the predicate occupies an A-position in predicate-initial specificational sentences. In general, Heggie's empirical database is questionable. Several of the claims supporting the A′-movement approach to predicate inversion (unavailability of specificational clauses in ECM environments, predicate raising to subject, T-to-C movement of the copula, adverbial modification to the right of the copula, etc.) are inaccurate (see Chapter 4 for controlled judgment surveys of these controversial features). Mikkelsen presents additional data from word order in Danish which distinguish predicate inversion in specificational sentences from true instances of predicate topicalization. She shows that predicate inversion patterns with A-movement, while predicate topicalization patterns with A′-movement. Mikkelsen's proposal treats the alignment between syntax and information structure as essential to the special properties of specificational sentences. Since the subject position is the unmarked topic position, a topical predicate, as in a specificational sentence, can serve as a structural subject.

Den Dikken also appeals to information structure in his predicate inversion analysis of specificational sentences. Contrary to Moro, den Dikken defends an asymmetrical model of the predica-

tional small clause. Similar to Bowers (1993), he argues for the structure in (43), where a functional head belonging to an abstract category (Relator) mediates between subject and predicate.

(43)

```
                        RP
                  ┌──────┴──────┐
                 XP            R′
              ┌───┴───┐    ┌────┴────┐
           Subject    R        YP
                             ┌────┴────┐
                              Predicate
```
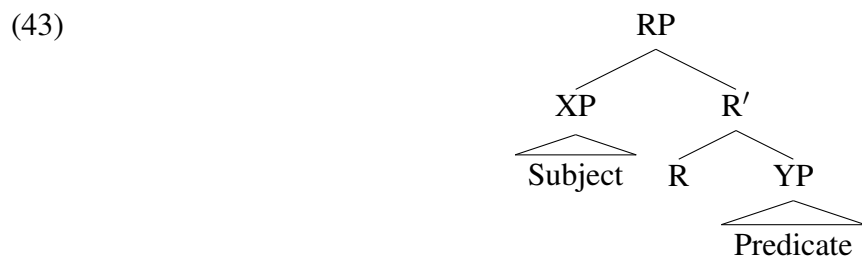
The Relator can be spelled out by a verbal copula. Since Relator is a meta-category, it also rears its head in nominal environments, where it is spelled out by a prepositional element like *as*, *for*, or *of*, and in A′-environments, where it marks Focus and Topic. The small clauses in (44) therefore have the same predicational syntactic core, although they are headed by different elements.

(44) a. [With John **as** our champion], we can't lose.

 b. [With John **being** our champion], we can't lose.

 c. [With John **for** our champion], we can't lose.

 d. [With John ∅ on our side] / [%∅ our champion], we can't lose.

Typical specificational copular sentences result from inverting the RP structure. The predicate is able to raise around its subject in an apparent violation of Relativized Minimality (Rizzi 1990) because the Relator remerges with its own projection—an instance of domain-extending head movement. With the small clause domain extended to include a new outer specifier, the predicate can raise locally around its subject, and gain access to an A-position in the T-domain.

(45)

                              LP
                    ╱   ╲
                 YP   L′
               ╱╲   ╱   ╲
           Predicate   L+R   RP
                               │   ╱   ╲
                             to-be   XP   R′
                                      ╱╲   ╱ │
                                  Subject   R   t$_{YP}$
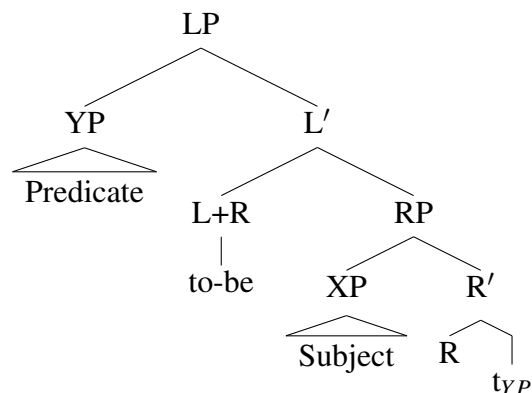
The presence of the additional projection is signaled by a "LINKER", an overt form of the copula.

This is why small clauses in subject-initial order are freer with respect to the optionality of the

overt copula, according to den Dikken (2006). In the predicate-initial order, where the LINKER is

projected, the head of the small clause must be spelled out overtly. In the non-finite embedded

clauses in (46), the non-finite copula *to be* spells out the LINKER.

(46)   a.   She considers [the best man for the job *(to be) John].

        b.   She considers [John (to be) the best man for the job].

Since equatives pattern with predicate-initial specificational sentences in requiring an overt

LINKER element ((37) and (47b)), den Dikken argues that there is in fact no separate category of

equatives: equatives are inverse predications, like specificational sentences. Heycock and Kroch

(1999) take the opposite tack, and propose that specificational sentences are equatives, but their

proposal, similarly, relies on the presence of a functional head that is absent in simple predicational

contexts.

(47)   a.   Cicero is Tully.

        b.   I expected Cicero *(to be) Tully.

While the distribution of non-finite *to be* in complement clauses is a useful test for predicate

inversion, it is not always conclusive, and its distribution depends on the properties of the predicate selecting the small clause. For example, some non-inverted predicational clauses require an overt non-finite copula.

(48)   a.   I believe [John *(to be) friendly/the best candidate].

        b.   I take [John *(to be) friendly/the best candidate].

Furthermore, the overt non-finite copula (with or without infinitival *to*) must not occur in a small clause embedded under *make* (and some other predicates), regardless of the order of subject and predicate.

(49)   a.   That makes [John (*be) / (*to be) a nice guy].[9]

        b.   That makes [John (*be) / (*to be) the teacher].

        c.   That makes [the teacher (*be) / (*to be) John].

The distribution of the overt non-finite copula does not depend on the presence of the LINKER alone. Elements other than *be* may realize the LINKER, and *be* may realize elements other than the LINKER (e.g., as an aspectual auxiliary). Because the properties conditioning the presence of the overt LINKER are complex (see Heycock 1991, 1994, 1995; Rothstein 1995, 1999; Heycock and Kroch 1999; den Dikken 2006; Geist 2007), I will set the issue aside, relying on the distribution of the non-finite copula as a diagnostic for predicate inversion only in well understood environments, like the complement of *consider*.

Another central component of den Dikken's (2006) proposal, similar in spirit to Mikkelsen's, is that predicate inversion feeds information structure, which in turn accounts for several well known restrictions on the syntax of specificational copular sentences. The Phase Impenetrability

---

[9](49a) is acceptable on the agentive/activity reading of *be*, which can be paraphrased as, 'That makes John behave like a nice guy.'

Constraint (Chomsky 2001) in conjunction with Relativized Minimality (Rizzi 1990) requires that the small clause subject remain "frozen" in its base position: when the Relator reprojects, the subject is trapped within the RP phase. Since the subject is the most deeply embedded overt element of the clause, it receives a narrow focus interpretation by default (Selkirk 1995); and since it is syntactically frozen in place, den Dikken (2006) proposes that focus fails to project up the tree, as it can in other contexts. As a reflex of the subject's focus interpretation, the raised predicate must be a topic.

(50)   Is John the principal?

      a.   No, the principal is JOE_F.

      b.   #No, the TEACHER_F is John.

Since the in-situ subject must be the focus, the predicate is barred from undergoing any further syntactic operations that would land it in a focus position. This is why specificational sentences are famously restricted when it comes to A'-movement operations.

(51)   *Which staff member did you say t_i is John t_i?

Arguing against Heggie's (1988) A'-fronting model, den Dikken shows that the predicate is not completely barred from moving, which would be expected if it were in a dedicated Topic position (a criterial position, in Rizzi's 2006 terms). In fact, long topicalization of the fronted predicate is possible, which shows that the moved element originated in an A-position.

(52)   The principal$_i$, nobody said t_i is JOHN.

The fact that some "frozenness" can be observed in subject-initial specificational sentences casts doubt on the derivational approach to the information structure of specificational sentences.

Consider the following examples. (53) shows a grammatical instance of predicate topicalization in a predicational sentence. (54) shows its ungrammatical counterpart in a subject-initial specificational sentence.

(53)   A: I heard that you are very proud of yourself.

   B: Yes indeed; [proud of myself]$_i$, I certainly am t$_i$.

(54)   A: I heard that important to himself is what John is.

   B: *Yes indeed; [what John is]$_i$, important to himself certainly is t$_i$.

Since no inversion has taken place in the specificational sentence in (54), there should be no ban on predicate topicalization. Moreover, the predicate is assigned a topic interpretation in both positions, so the problem does not arise from an information structure clash.

2.2.3.2 Semantic relationships between copular sentence types

The lack of consensus about the structure of copular sentence types cannot be disentangled from the debate about the meaning of these sentence types. In particular, the semantic type of the "predicate" in a specificational sentence and the semantic contribution of the copula (if any) in all sentence types, remains an open question. A perspective is offered in Chapter 4, section 6.1.

If the specificational predicate is simply a predicate in an unusual structural position (Williams 1983; Partee 1986; Heggie 1988; Moro 1997; Mikkelsen 2005; den Dikken 2006), then the specificational sentence type should be collapsed with the predicational type. If instead the putative predicate has the same logical type as the subject (Jacobson 1994; Heycock and Kroch 1999; Sharvit 1999; Schlenker 2003), then it should be collapsed with the equative category. Table 2.2 shows the types of "subject" and "predicate" in the three sentence types, highlighting the controversial status of the less referential constituent of the specificational sentence. (The table assumes

the prototypical Topic-Comment order for the specificational sentence.)

		Precopular	Postcopular
Predicational	uncontroversial	X	$\langle X,t\rangle$
Specificational	Higgins (1979), Williams (1983), Partee (1986), Heggie (1988), Mikkelsen (2005), den Dikken (2006)	$\langle e,t\rangle$	e
	Heycock and Kroch (1999)	e	e
	Schlenker (2003)	$\langle s\langle s,t\rangle\rangle$	$\langle s,t\rangle$
	Jacobson 1994, Sharvit 1999	$\langle e,t\rangle$	$\langle e,t\rangle$
	Romero (2005), Comorovski 2008	$\langle s,e\rangle$	e
Equative	uncontroversial	X	X

Table 2.2: Logical types of XP and YP in copular sentence types

2.2.3.2.1 The contribution of the copula

Closely related to the choice of analysis of the predicate in a specificational sentence is the analysis of the semantic contribution of the copula. In the predicational copular sentence, the copula is vacuous: it simply passes up the denotation of the predicate, which combines directly with the subject.[10]

(55) $\lambda P \lambda x[P(x)]$

The predication relation can obtain in the absence of an overt copula, e.g., in a bare small clause, or with an overt copula, e.g., in a tensed clause, with no discernible difference in the meaning of the relation. Williams (1983) proposes that in inverse (specificational) sentences, the vacuous copula simply takes its P and x arguments in the opposite order.

Now consider equatives, which express identity statements. It is certainly the case that predication (i.e., function application) and identity have different meanings, but it does not follow that two homophonous versions of the copula contribute these meanings. There is some debate as to

---

[10]The vacuous analysis of the copula was espoused even by some early philosophers of language and logicians.

whether the copula itself contributes the meaning of identity (Russell 1919; Higgins 1979; Akma-

jian 1979; Jacobson 1994; Sharvit 1999; Schlenker 2003; Mikkelsen 2005; Romero 2005, etc.), or

whether this meaning is associated with an abstract functor that applies either to the denotation of

an individual to yield a predicate (Partee 1986, 2000), to the predicational small clause (Heycock

and Kroch 1999), or to the copula itself to generate an identity function from a vacuous predication

function (Geist 2007).

Partee (1986) proposes that the copula always has the denotation in (55). In equatives, where

two referential expressions must combine via the copula, the type-shifter IDENT resolves the type

mismatch by mapping one of those referential expressions onto the singleton set that contains it.

(56)   IDENT: $\lambda x \lambda y [x = y]$

For example, IDENT, applied to the individual denotation *Tully*, yields a predicate denotation: the

set of individuals that are equal to Tully.

(57)   IDENT(Tully): $\lambda y [y = Tully]$

This type-shift enables the two expressions in an equative to combine by function application. The

type-shifting solution works well, but ideally, we would find overt evidence to support it.

Geist (2007) proposes that the distribution of the pronominal copula in Russian equatives of-

fers precisely this evidence. She applies a different version of Partee's proposal, arguing that IDENT

composes with the denotation of the vacuous predicational copula itself, to yield an identity pred-

icate. The application of predicate composition is signaled by *eto* 'this', a pronominal copula that

occurs only in equatives, but never in specificational or predicational sentences (recall (6), above).

Parallel evidence is not available in English, which shows no formal difference between the copula

of predicational, specificational, or equative sentences—at least in finite clauses. Geist takes the

obligatoriness of non-finite *be* in English non-finite equative and specificational clauses to support her type-shifting approach, but once again, the absence of *be* under predicates like *make* is problematic.

Unlike in predicational clauses, specificational and equative clauses allow elements of any syntactic category to be subjects. Do these elements all have the same semantic type, or is the specificational/equative relationship blind to category? Partee (1986, 2000) proposes that the copula is polymorphous: for any type X, it relates X and $<X,t>$. Type-driven translation enriched with type-shifting operations will allow the two expressions to combine.

2.2.3.2.2 The type of the predicate

The predicate-inversion models of specificational sentences generally assume that the predicate is an expression of type $<e,t>$, familiar from ordinary contexts (perhaps derived by application of IDENT). Although it is has the form of a referential definite DP, Mikkelsen (2005) argues that it is not referential. If it were, it should be able to bind an animate pronoun in a tag question, rather than an inanimate one, contrary to fact, (58).

(58)   The tallest girl in the class is Mary, isn't *she / it?

Just because the predicate is non-referential, however, doesn't mean that it is an $<e,t>$-type predicate. Mikkelsen observes that only expressions that can denote both properties and individuals can serve as specificational predicates. Interestingly, Mikkelsen's tag question test clashes with Heycock and Kroch's (1999) non-restrictive modification test.

(59)   The tallest girl in the class, who is very studious, is Mary, isn't she / *it?

This is the now-familiar puzzle: are specificational sentences equative or are they predicational?

Schlenker (2003), Romero (2005), and Comorovski (2008) take seriously the parallels between specificational sentences and question-answer pairs, and propose that the predicate in a specificational sentence is an intensional object: a question, concealed question, or individual concept, respectively. For Schlenker, this means that the specificational predicate denotes a proposition intension.

Because I maintain a one-*be* approach to the copula, I will marry the two sets of approaches. The predicate of a specificational sentence is an intensional predicate ($<$s,$<$X,t$>>$). It is not a referential expression. The copula's function is vacuous. It happily combines any two expressions where one is of type X and the other of type $<$X,t$>$, whether X is an entity, property, or proposition. An additional world variable is needed to account for world and time-dependence in the interpretation of specificational copular sentences (see Chapter 8 for detailed discussion).

2.2.4   Summary

The preceding sketch of specificational copular sentences shows that an intricate typology—constructed from syntactic or semantic criteria—is not needed. Specificational sentences can be derived from predicational small clauses via predicate inversion, where the underlying predicate lands in the structural subject position.

On the semantics side, the most parsimonious analysis of the copula is that it is vacuous. Type-shifting operations mediate the relation between the subject and predicate if they cannot compose directly. The predicate is world-time dependent, so it is an intensional object that takes a world-time argument (s) and an argument of the same type as the subject, whether that be an entity, a property, or a proposition.

2.3 The amalgam specificational copular sentence

Let us now turn to the sentence type at the center of this thesis: the *amalgam* specificational copular sentence. The main types of copular amalgam sentence are illustrated in (60)–(67) below, with the amalgam copula in boldface.

(60)   What he needs **is** (he needs) a break.   *wh*-initial amalgam pseudocleft

(61)   He needs a break **is** what he needs.   Reverse amalgam pseudocleft

(62)   The problem **is** he needs a break.   DP-weight amalgam

(63)   He needs a break **is** the problem.   Reverse DP-weight amalgam

(64)   That's what he needs **is** (he needs) a break.   *That's x is y*

(65)   You know what he needs? **is** (he needs) a break.   Question-answer amalgam

(66)   He really needs that if you ask me, **is** (he needs) a break.   Free-*be*

(67)   The problem **is**, **is** he needs a break.   Double-*is*

The copular amalgam sentences exhibit the same basic configuration and interpretation as *canonical* specificational copular sentences. The copula mediates between two constituents in an asymmetrical relationship. One of the constituents is more referential and provides a value for a variable in the other, which is less referential. The two sentence types are quite different, however, when it comes to their syntactic and semantic components.

The copular amalgam sentences form a natural class with respect to two interrelated properties. First, their logical subject is a bare, root-like clause. Secondly, they include an occurrence of the finite copula in an unusual role: relating propositions. These two properties present a puzzle. Given

that bare finite clauses cannot function as ordinary logical or structural subjects, what is the syntax of the copular amalgam sentence?

To demonstrate what is at stake in answering this question, the rest of this chapter will lay out some of the basic properties of the copular amalgam sentence type.

2.3.1 The term "amalgam"

The term "amalgam" (pseudo)cleft originates in Declerck (1988). Copular amalgams are understudied in formal linguistics, but they have not gone wholly unnoticed. Akmajian (1979) and Higgins (1979) point out amalgam pseudoclefts, but judge them as ill-formed. Ross (1972, 2000) gives them a fighting chance, and includes them in larger surveys of pseudocleft sentences. Later, den Dikken et al. (2000) explicitly distinguish the weight-initial amalgam cleft ("Type A specificational pseudocleft") from the canonical pseudocleft. Lambrecht (2001) includes the amalgam cleft in a profile of the discourse properties of cleft sentences. In the same framework, Ross-Hagebaum (2004) and Lambrecht and Ross-Hagebaum (2006) provide a corpus-based analysis of *that's x is y* and other apparently amalgamated sentence types.

The reason these sentences have been labeled "amalgams" is that they seem to consist of two sentences overlapping on a shared constituent. The areas of apparent overlap in the amalgam pseudocleft examples below are identified with angle brackets.

(68)   a.   What I want is I want <a vacation>.

       b.   I want <a vacation> is what I want.

They are reminiscent of the Horn-amalgam, like (69) (Lakoff 1974, attributed to Larry Horn):

(69)   John is going to, I think it's Chicago next Sunday.

Amalgam sentences present a bracketing problem. (69) seems to be composed of an amalgamation of the two sentences in (70).

(70)   a.   John is going to Chicago next Sunday.

         b.   I think it's Chicago.

Van Riemsdijk (2006) argues that the canonical formulation of merge straightforwardly allows amalgams like (69), where one terminal node is dominated by two different root notes. He labels this type of merge "graft".[11] Multi-dominance approaches to amalgams are also proposed by de Vries (2009), Kluck (2011), and Heringa (2012).

Nevertheless, such approaches are unsuitable for copular amalgams (see section 2.5.1.1 for discussion). Although all the copular amalgam sentence types introduced in the previous subsection share the same bracketing problem, I do not analyze any of them as actual syntactic amalgamations. Instead, the copula serves to relate different categories of constituent than it usually does: propositions.

(71)   a.   [What I want] is [I want a vacation].

         b.   [I want a vacation] is [what I want].

2.3.2   Some more terminology

To facilitate the discussion of copular amalgam sentences, I will briefly establish some terminology. The major constituents of canonical specificational sentences go by several different labels in the existing literature, so it is important to be explicit. This paper will use the following terms:

---

[11]Grafting is also responsible for the formation of much less exotic free relative structures and "transparent" free relatives, like (i).

(i)   I saw what you might call a mime performing in Times Square.

(72)   **anchor**: the immediate constituent including the variable and its binder

        value: the focus identifying the variable

        weight: the dependent of the copula that includes the anchor

        counterweight: the dependent of the copula that includes the value

Consider the question-answer amalgam below.

(73)   [You know [what$_i$ he needs e$_i$]$_{ANCHOR}$]$_{WEIGHT}$ is [[he needs a break$_i$]$_{VALUE}$]$_{COUNTERWEIGHT}$

The anchor here is *what he needs*. It includes a variable, the internal argument of *needs*, and the *wh*-operator *what* binding that variable. The anchor is embedded within a fully sentential weight clause, *you know what he needs*. On the other side of the copula is a value for the variable: *a break*. The value is embedded within a fully propositional counterweight clause: *he needs a break*.

We might conceptualize the relevant pieces of this sentence in two ways. In one respect, the anchor and the value are interpreted as standing in a specificational relation:

(74)   You know [what he needs] is he needs [a break].

The main syntactic constituents of the sentence, however, are the weight and the counterweight:

(75)   [You know what he needs] is [he needs a break].

Now consider the amalgam pseudocleft, exemplified in (76).

(76)   [He needs [a break$_i$]$_{VALUE}$]$_{COUNTERWEIGHT}$ is [[what$_i$ he needs e$_i$]]$_{ANCHOR/WEIGHT}$

The weight in (75) happens to be the same as the anchor: the *wh*-clause including the relevant variable is not included in any larger constituent. In postcopular position is the counterweight clause, *I need a break*, which contains the focus. The value in (76) is a subconstituent of the

counterweight clause.

If the anchor in the weight clause has a propositional denotation, the entire counterweight

clause can serve as the value, as in (77).

(77)   [[He needs a break]]_{VALUE/COUNTERWEIGHT} is [Op_{ANCHOR} his biggest problem e]_{WEIGHT}

I distinguish between amalgam specificational copular sentences, where the value/counterweight is

a bare finite clause, and canonical specificational sentences like (78), where the value is a formally

embedded CP:

(78)   a.   [That he needs a break] is his biggest problem.

          b.   [He needs a break] is his biggest problem.

I will refer to the major constituents of canonical sentences like these as the *anchor* and the *value*,

reserving *counterweight* for amalgams.

2.3.3   Ingredients of copular amalgam sentences

Let us take a brief tour through these ingredients, as they occur in the different amalgam sentence

types.

2.3.3.1   The anchor

The anchor expression in an amalgam specificational copular sentences is always a CP or DP with

a propositional denotation (79).

(79)   a.   [What he studies] is he studies biology.

          b.   He really loves Paris is [my impression].

          c.   That's [what we saw] is we saw *The Godfather*.

          d.   You know [who she invited]? is (she invited) John.

    e.   [The funny thing] is, is he didn't even like it.

The CP anchors occurring in amalgams include unambiguous indirect questions, (80), and expressions that are ambiguous with free relatives, (81).

(80)   a.   You know [who had what]? is he had the beef and she had the salmon.

          b.   That's [what she did], is she went to the pharmacy.

(81)   a.   That's [what I ate], is I ate some cheese.

          b.   He wrote a book is [what he wrote].

That's x is y and many free-*be* amalgams are distinguished by having a demonstrative element in their weight clause. (Both *that* and the *wh*-clause/DP in the weight clause are anchors for specification.)

(82)   a.   **That**'s what happened, is she stormed out.

          b.   You should really think about **that** before you leave, is what your mother would say.

          c.   You see **that** again and again is that phonologists come up with theories for English which just don't work for Tiberian Hebrew.   (Massam 1999:345)

The most common demonstrative element occurring here is a proposition-anaphor, *that*, but other deictic and pronominal expressions are allowed as well:

(83)   a.   **This** is what I'm saying, is that he's presenting a false dichotomy.[12]

          b.   Can I simply say **this**, is that the parliamentary process is a difficult one. (McConvell 1988:302)

          c.   **Here**'s the problem, is that the president can't get his own party to agree to it right

---

[12]http://www.npr.org/2011/01/07/132740175/paul-offit-on-the-anti-vaccine-movement

now.[13]

    d.   **It**'s a common scavenger hunt activity for bachelorette parties around here is to go get a cock and balls.[14]

Ross-Hagebaum (2004) analyzed 207 tokens of *that's x is y* sentences in the Switchboard Corpus. In his sample, 56.5% of tokens featured a DP anchor, while 43.5% occurred with what he identifies as a relative clause. The precopular *that's x* clause is itself a so-called identificational copular sentence, expressing identity between the referent of the demonstrative and the referent of the *wh*-clause/DP. Both receive content, ultimately, from the postcopular value.

The root clause form of the weight clause in *that's x is y* highlights the fact that the weight in amalgams is propositional. Similarly, the prototypical DPs occurring in the role of anchor (see also the literature on double-*is*; Massam 1999, 2013; Coppock and Staum-Casasanto 2004; Curzan 2012) are lexically proposition-denoting DPs headed by, e.g., *problem, issue, concern, idea*, etc. DPs that can denote concealed questions (e.g., *her favorite place*, *the tallest guy in the class*) are also acceptable as anchors within a clausal weight, because concealed questions have propositional values (see further discussion in Chapter 4).

(84)   That's her favorite place, is (it's) the Grand Canyon.

The amalgam anchor can have the syntactic form of a free relative, an indirect question, or a definite description, but it is always semantically propositional. I will argue in Chapter 4 that this is the crucial distinction between canonical and specificational anchors: the former are predicates of individuals, while the latter are always predicates of propositions.

---

[13]http://www.foxnews.com/on-air/on-the-record/transcript/land-ariz-law-039cry-help039-and-039symptom039

[14]Overheard conversation about Voodoo Doughnut, in Portland, Oregon

2.3.3.2 The value

The value in an amalgam copular sentence can be of any syntactic category.[15] The following examples illustrate the categorial flexibility of the value. The examples below are all amalgam pseudoclefts, but the facts hold of the other amalgam types as well.

(85) DP

    a.   What they read is they read [*Hamlet*].

    b.   They read [*Hamlet*] is what they read.

(86) AP

    a.   What he is is he's [important to himself].

    b.   He's [important to himself] is what he is.

(87) PP

    a.   Where she went is she went [to the store].

    b.   She went [to the store] is where she went.

(88) VP

    a.   What she will do is she will [try again].

    b.   She will [try again] is what she will do.

(89) AdvP

    a.   How he runs is he runs [quickly].

    b.   He runs [quickly] is how he runs.

---

[15]Some gradience is reported in Ross (2000) with respect to the acceptability of adverb values, specifically in canonical pseudoclefts.

(90)   CP

    a.   Why he's angry is [he lost the game].

    b.   [He lost the game] is why he's angry

VP values are particularly productive in amalgams, perhaps because the "canonical" pseudo-cleft counterpart with VP as the logical subject is dispreferred (91b).

(91)   a.   I should have reported her. Hell, I should have punched her is what I should have done![16]

    b.   Punch her is what I should've done.

VP-value amalgam pseudoclefts introduced by the string *What x does/did/should do/will do is...* are highly frequent in spoken English, occurring in a range of registers.

Not only can the value be of a variety of categories, it can also serve in a variety of grammatical roles. Argument and adjunct roles are fully acceptable, although the subject role is dispreferred (95), for reasons that remain unclear.[17]

(92)   Object

    a.   What he should try is he should try [the IPA].

    b.   He should try [the IPA] is what he should try.

---

[16]David Sedaris. 2013. "Let's Explore Diabetes with Owls."

[17]It is possible that the slight degradation of subject-value amalgam pseudoclefts is related to the difficulty associated with narrowly focusing in-situ subjects, but since English generally has no trouble focusing subjects (in contrast to, e.g., French), a different explanation may be needed.

(i)   JOHN wrote the book, not Mary.

(ii)   a.   #JEAN a écrit le livre, pas Marie.
         John has written the book, not Mary.
    b.   C'est JEAN qui a écrit le livre, pas Marie.
         It's JEAN who has written the book, not Mary.

    c.   That's what he should try is (he should try) [the IPA].

(93)   Non-verbal predicate

    a.   What he is is he's really well read/all about football/a creative athlete/just one of those kids/a shortstop with a plus arm, plus speed.[18]

    b.   It's [selfish] is what it is.[19]

(94)   Adverbial adjunct

    a.   Why I left is I left [because I was angry].

    b.   She met him [at work] is where she met him.

(95)   Subject

    a.   ?What killed him is [a falling brick] killed him.

    b.   [A falling brick] killed him is what (?killed him).

    c.   That's what killed him is [a falling break] (?killed him).

Canonical pseudocleft values, by contrast, show a preference for argument roles, and no subject-object asymmetry.

(96)   a.   What he should try is [the IPA].

    b.   ?Where she met him is [at work].

    c.   What killed him is [a falling brick].

2.3.3.3 The weight

In amalgam pseudoclefts, both *wh*-initial and reverse, as well as DP-weight and double-*is* amalgams, the weight and the anchor are coextensive. As we saw above, they are CPs or DPs with

---

[18]Each example is drawn from the transcript of a different talk-show interview about a different male athlete.

[19]*Scrubs*. Television. NBC.

propositional meanings. In DP-weight amalgams, there is a strong preference for lexically proposition-denoting DPs, but concealed question DPs are also (marginally) acceptable, (97).

(97)   a.   ?Her favorite place is she likes to hang out at the mall.

         b.   ?She likes to hang out at the mall is her favorite place.

Like lexically proposition-denoting DPs (e.g., *the issue*), concealed question DPs refer to open propositions (see, e.g., Nathan 2006).

(98)   Tell me the issue.

       'Tell me the proposition p, such that p is the unique value of *the issue*.'

(99)   Tell me her favorite place.

       'Tell me the proposition p of the form *her favorite place is x*, such that x is the unique value of *her favorite place*.'

Unlike the lexically proposition-denoting DPs, however, the concealed question reading of DPs like *her favorite place* is non-basic. Because of the extra semantic computation involved in the propositional reading, these DPs fare better as amalgam weights when there is pragmatic support, for instance, a strongly presupposed open question.

(100)   A: Tell me her favorite place.

        B: Her favorite place is she likes to hang out at the mall.

Chapter 4 will discuss the properties of DP concealed question weights in more detail.

    In *that's x is y*, question-answer, and free-*be* amalgams, the weight is a full root-like proposition occupying the structural subject position of the clause.

(101)   a.   [That's the reality] is we don't have enough money.

    b.   [You know what bothers me], is we don't have enough money.

    c.   [I'm explaining why I'm asking the question] is that it seems like from the relation-

        ship that you had with her that you would have been one of those people.[20]

The form of this proposition is heavily constrained in the first two sentence types. In *that's x is y*,

as we have seen, it is an identificational copular sentence. It is also possible to find semi-copulas

and copulas associated with additional functional material in the weight clause of *that's x is y*.

(102)   a.   That might be her specialty, is getting out of situations like this.

        b.   That has always been her specialty, is getting out of situations like this.

        c.   That's not a way to convince me, is by saying that your fourth-graders listen to it.

    In the question-answer amalgam, the anchor is always embedded directly under an epistemic

predicate associated with the addressee. The question-answer amalgam is not particularly produc-

tive.

(103)   a.   You know what we did, is we left early.

        b.   Guess what we did, is we left early.

        c.   *Ponder what we did, is we left early.

The predicates *know* and *guess* directly prompt the addressee to call the question and its possible

answers to mind.

2.3.3.4   The counterweight

The sentential nature of the counterweight is perhaps the most marked feature of amalgam copular

sentences. It is a root-like finite clause that occurs as both the logical and sometimes structural

---

[20]"Serial". Narrated by Sarah Koenig. Episode 6. This American Life. NPR. October 20, 2014.

subject of the amalgam clause. This is not possible in other contexts.

(104)   a.   [She fetched a pail of water] is what she did.

          b.   [She fetched a pail of water] is the nice thing.

(105)   a.   *[She fetched a pail of water] is kind.

          b.   *[She fetched a pail of water] enabled me to drink.

Although the counterweight is structurally embedded in its precopular or postcopular position, it

exhibits a range of root properties, suggesting that it is a species of embedded root (Hooper and

Thompson 1973; Heycock 2006, etc.). For instance, it cannot be introduced by the finite comple-

mentizer *that* (106).[21]

(106)   a.   *That she fetched a pail of water is what she did.

          b.   *What she fetched is that she fetched a pail of water.

          c.   *That's what she fetched is that she fetched a pail of water.

Interrogative counterweights likewise show root, rather than embedded word order.[22]

(107)   a.   Why did she fetch a pail of water? is what is at issue.

          b.   What is at issue is why did she fetch a pail of water?

          c.   That's what's at issue is why did she fetch a pail of water?

---

[21]Some exceptions to this are *that's x is y* and free-*be* amalgams with clausal *values*.

(i)   That's the problem, is that I'm really tired.

In these constructions, a root clause, the weight, occupies the structural subject position, so the bare finite clausal
subject problem remains.

[22]The embedded question word order is of course acceptable, on the canonical specificational copular sentence
parse.

(i)   Why she fetched a pail of water is what is at issue.

Chapter 4 will show how the canonical and amalgam parses can be disambiguated, where relevant. What is of interest
at present is the fact that the root order is possible in a structural subject at all.

Topicalization (108) and locative inversion (109), other diagnostics for root status, are licit in the counterweight clause, provided the narrow focus interpretation of the value can be maintained.[23]

(108)   a.   What you should try is tomorrow, you should call them back.

        b.   To John, you should give a book is what you should do.

        c.   That's what you should do is to John, you should give a book.

        d.   #A book, you should give to John is what you should give him.

(109)   a.   What happened is out of that drain jumped a huge spider.

        b.   Into the room rode a scary clown on a unicycle is why we screamed.

        c.   That's what she described is from around the corner appeared the train.

        d.   #Down the hill rolled the baby carriage is where it rolled.

In light of the fixed information structure of specificational sentences, the counterweight is always focused, which makes it a good candidate for serving as an assertion (Hooper and Thompson 1973). We might suspect that the counterweight is the root, and the rest of the structure merely adjoined to it.

The embedded status of the counterweight proposition, despite its root form, is confirmed by the fact that it need not indicate the illocutionary force of the whole sentence. In the amalgam pseudocleft, the copular clause is the locus of the primary assertion (see Lambrecht 2001).[24] Amalgam pseudoclefts with interrogative counterweights, like (107), need not be interpreted as interrogative. (110) illustrates that such amalgam pseudoclefts can be accompanied by tag questions, which indicates the presence of assertive force in the copular clause.

---

[23] The judgments of sentences like those in (98) and (109) are admittedly tenuous; the sentences are poor in written form, and the general markedness of topicalization in English induces a register clash with the colloquial amalgam. Nevertheless, with suitable prosody, I find these sentences acceptable.

[24] Since *that's x is y* contains two complete root-like clauses, it is difficult to tease the force of the copular clause apart from the force of the weight and the counterweight. This is discussed further in Chapter 5.

(110)   a.   What is at issue is why did she fetch a pail of water, isn't it?

      b.   What is he going to do? is what he wants to know, isn't it?

The presence of assertive force in the copular clause is also supported by the fact that the interlocutor can felicitously deny or question it.

(111)   She fetched a pail of water is what she did.

      a.   No, it isn't!

      b.   Is it really, though?

The interlocutor can *also* deny or question the counterweight proposition.

(112)   She fetched a pail of water is what she did.

      a.   No, she didn't

      b.   Did she really, though?

Both the counterweight and the copular clause have root properties, but the force of the counterweight clause is subordinate.

Further evidence that the counterweight is embedded comes from recursive amalgams. In (113), weight-initial amalgam pseudoclefts are embedded within reverse amalgam pseudoclefts.

(113)   a.   And [[what they've done is they've bought... Mercedes, after Mercedes, after Mercedes], is what they've done].   (Lambrecht and Ross-Hagebaum 2006:20)

      b.   [That's my feeling, is [that's the most rewarding way to harvest deer, is to track it in the snow, and get it]].[25]

---

[25]http://www.howcast.com/videos/460894-Top-9-Tips-for-Deer-Hunting-Hunting

The counterweight clause is not the matrix, since one counterweight can be embedded within another.

When the counterweight occurs in postcopular position, it is typically expressed in full sentential form, but it can also be reduced. In Ross-Hagebaum's (2004) corpus sample of *that's x is y* tokens, the counterweight is usually a finite clause (42.2% of tokens), and less frequently a non-finite clause (25.2% of tokens), or a DP (26.5% of tokens). Even less common are PPs (4.8% of tokens). Other categories accounted for only 1.3% of Ross-Hagebaum's tokens.

The full finite clause form is basic, and the others, (114), are likely derived from it via ellipsis of the content that is identical to the weight, like a fragment answer in a discourse question-answer pair, (115).[26]

(114) a. That's what we need, is ~~we need~~ to get away for a while.

b. That's just the thing, is ~~the thing is / it's~~ a nice, relaxing bath.

c. That's where I am now, is ~~I'm~~ at the hospital.

d. That's what they are, is ~~they're~~ tall.

e. That's where they went, is ~~they went~~ to the store.

f. That's how they did it, is ~~they did it~~ quickly.

(115) A: What do we need?

B: ~~We need~~ to get away for a while.

NPI connectivity effects provide evidence that such ellipsis is available in amalgam pseudoclefts, (116) (den Dikken et al. 2000).

---

[26]Although the reduced counterweight is strongly preferred when the counterweight clause is itself copular, fully overt versions do occur.

(i)   What possibly it is is it's a class thing.   (overheard)

(116)   That's what nobody wants, is ~~nobody wants~~ any more blizzards.

The NPI in the value must be licensed by a c-commanding negation. The overt *nobody* in the weight

clause does not c-command the NPI value, so the only possible local licensor is a phonologically

empty *nobody* in the counterweight clause.[27]

2.3.4   Related phenomena

The pseudocleft and DP-weight amalgams are at the center of this thesis, since they have direct

structural counterparts in canonical form. I also draw heavily on the *that's x is y* construction,

which is productive, and offers useful comparisons, since its weight is a root clause in structural

subject position.

Potentially related phenomena involving unusual copulas in English include the nominal clause

DP-weight amalgam in (117), and the comparative correlatives in (119).

(117)   They did [[that same thing] of [it feels like you're watching a stage]].[28]

In (117), the weight *that same thing* is related to a bare sentential counterweight, *it feels like you're

watching a stage*, by a nominal copula, *of*. Den Dikken (2006) proposes that *of* spells out the

LINKER in inverse predications inside DP:

(118)   a.   a jewel of a village

         b.   an idiot of a doctor

Comparative correlatives, like copular amalgams, relate two propositional elements in one in-

tegrated syntactic structure. Culicover and Jackendoff (1999) treat the relationship as paratactic,

---

[27]Other connectivity effects, such as bound variable connectivity and opacity connectivity, are amenable to LF
analyses that do not rely on surface c-command (see Jacobson 1994; Sharvit 1999). As den Dikken et al. (2000) argue,
however, NPIs must be licensed at spell-out, so NPI connectivity effects make a direct case for the presence of an
elliptical counterweight in some reduced or superficially canonical pseudoclefts.
[28]"The Takeaway." WNYC.

which is reminiscent of the relationship between the weight and counterweight in, e.g., *that's x is y*. Both constructions allow for an asyndetic relationship (not mediated by any overt element), and at least for some speakers of English, particularly in Jamaica, both also allow a finite copula to mediate the relation (Heggie and Iwasaki 2012).

(119) a. The more unaware Megan is, **is** the more intense, obsessive and possessive Antonio gets about her.[29]

      b. The more rural the place is **is** the more traditional the celebration is.[30]

Another construction that is similar to copular amalgams in its basic configuration and use is the sentential appositive relative, (120). In this structure, two root-like clauses are juxtaposed, and the nonrestrictive relative operator *which* mediates a modification relation between them.

(120) a. There's a really great taco place on the corner, which, we should try to go there before you leave.

      b. It's supposed to be sunny tomorrow, which, we might be able to finally put away our coats.

The first proposition is asserted, and the second offers some kind of elaboration on it. Although the second clause is introduced by a relative operator, there is no gap, and the logical relation between the first and the second clause is very loose. The intonation contour of these sentences is unlike ordinary non-restrictive relatives, since there is a break before and after *which*. Here, *which* seems to function like a coordinator, rather than a relative operator.

It has a "colon function" similar to that of copular amalgams (Higgins 1979). In both sentence

---

[29]http://www.amazon.com/Mobsters-Girl-ebook/dp/B008BNF0R6
[30]http://www.virtualtourist.com/travel/North_America/Mexico/Local_Customs-Mexico-TG-C-1.html

types, two propositions are related by a functional element not ordinarily found in that environment, a relative operator and a finite copula, respectively. The two structures cannot be completely assimilated, however. First of all, the copula in amalgams is not the direct counterpart of the appositive relative operator, because it is morphologically finite, and capable of bearing illocutionary force. Moreover, appositive relatives like those in (120) are predicational, while copula amalgams are fundamentally specificational.

The prosody of the sentential apposition is also different from the weight in the reverse amalgam pseudocleft. While the sentential appositive relative adds a focused elaboration on the first sentence, with a distinct pitch accent, the post-copular weight clause in a reverse amalgam pseudocleft is backgrounded, and often deaccented.

(121)   She left EARLY, which, I hope she makes it on TIME.

(122)   She left EARLY is when she left.

Copular amalgams are also not derived from appositive relatives with *which*-deletion. If they were, then predicational copular amalgams should be possible, contrary to fact.

(123)   *She left early, ~~which~~ is unfortunate.

In fact, when the nonrestrictive relative operator is overtly added to the amalgam, the result is degraded.

(124)   ?She left early, which is what she did.

In addition, if copular amalgams were formed from appositive relativization structures, then the copula should behave like an ordinary verbal copula. The next section will demonstrate that it does not. Despite their shared colon function, sentential appositive relative constructions and copular

amalgam constructions are syntactically distinct.

2.4 Some properties of copular amalgams

This section provides a brief sketch of some syntactic and interpretive properties of copular amalgams.

2.4.1 Some syntactic properties

2.4.1.1 Restrictions on moving parts

Unlike canonical specificational sentences, amalgams are almost completely frozen: they do not license any A′-displacement operations (see also Ross 2000).

(125)   He needs a new job is the biggest problem.

     a.   *[How big of a problem]$_i$ do you think he needs a new job is t$_i$?

     b.   *[What kind of job]$_i$ do you think [he needs t$_i$] is the biggest problem?

     c.   *Is he needs a new job the biggest problem?

     d.   *[The biggest problem]$_i$, he needs a new job certainly is t$_i$.

(126)   The biggest problem is he needs a new job.

     a.   *[How big of a problem]$_i$ do you think t$_i$ is he needs a new job?

     b.   *[What kind of job]$_i$ do you think the biggest problem is [he needs t$_i$]?

     c.   *Is the biggest problem he needs a new job?

     d.   *[The biggest problem]$_i$, nobody thinks t$_i$ is he needs a new job.

(127)   That's the biggest problem is he needs a new job.

     a.   *[How big of a problem]$_i$ do you think [that's t$_i$] is he needs a new job?

     b.   *[What kind of job]$_i$ do you think that's the biggest problem is he needs t$_i$?

    c.   *Is that's the biggest problem he needs a new job?

    d.   *[That's the biggest problem]$_i$, nobody thinks t$_i$ is he needs a new job.

Chapter 4 discusses the full paradigm of movement restrictions in copular amalgams.

2.4.1.2   Absence of "baggage"

Another salient property of copular amalgams is that the copula is almost always in simple form:

is or *was*, and less frequently, *are* or *were*.

(128)   a.   He needs a new job **is** what he needs.

         b.   He needed a new job **was** what he needed.

(129)   He needs a new job and he wants to move **are** the issues.

The tense form of the copula tends to match the tense form of the lexical verb in the precopular

constituent (see Chapter 8 for an analysis).

    Chapter 4 presents experimentally obtained judgment data showing that the amalgam copula,

in contrast to the canonical copula, cannot combine with aspectual, temporal, or (deontic/root)

modal auxiliaries. I will refer to such functional material, e.g., modals, negation, and auxiliaries,

as "baggage".

(130)   a.   He needed a new job **was** / *__had been__ what he needed.

         b.   He will need a new job **is** / *__will be__ what he will need.

         c.   He needs a new job **is** / *__must obligatorily be__ what he needs, if he wants to move

            forward.

This property holds regardless of whether the counterweight clause includes matching baggage.

For example, in (130b), the future marker *will* occurs in the counterweight and weight clauses, but

still the copula must be bare.

Negation is also impossible in combination with the copula, (131), although it can occur inside of the weight or counterweight clause, (132).

(131) a. *He needed a new job **wasn't** what he needed.

       b. *That's what he needs **isn't** (he needs) a new job.

(132) a. He didn't need a new job **was / *wasn't** what he didn't need.

       b. That's not what he needs, **is / *isn't** (he needs) a new job.

2.4.2 Some interpretive properties

2.4.2.1 Specificational interpretation

Copular amalgams are quite similar to their canonical counterparts in their specificational interpretation. As we have seen, their basic components include an anchor with a variable, and a value providing the content for the variable. A simple predicational interpretation of copular amalgams is not available, (133).

(133) a. *He needs a job is an interesting idea.

       b. *That's what you meant, is unfortunate.

       c. *That's darkly colored, is you are wearing a dress.

The obligatory specificational interpretation derives from the question-answer pair meaning of the copular amalgam: the predicate has a question or open proposition meaning, which generates a set of focus alternatives; the subject provides the answer that values the variable in the open proposition.

2.4.2.2   Information structure of copular amalgams

The question-answer interpretation is also responsible for the fixed information structure of the copular amalgam. Since the counterweight provides an "answer" to the question in the weight, it is necessarily focused. The specific uses of each amalgam type in discourse are more nuanced. Below, I briefly summarize insights of Lambrecht (2001) and Lambrecht and Ross-Hagebaum (2006), who discuss the information structure of specificational sentences, including amalgams.

The basic intuition about the information structure of specificational sentences is that the anchor is a topic, and the value is a focus. Topic is not a unitary notion, however. Lambrecht (2001) identifies three types of presupposition at work in specificational sentences.

Pragmatic Presupposition, also called K(nowledge) Presupposition as defined in (134).

(134)   K-presupposition

"The set of propositions lexico-grammatically evoked in a sentence that the speaker assumes the hearer already knows or believes or is ready to take for granted at the time the sentence is uttered (the "old information")" (Lambrecht 2001:474)

To characterize the felicitous distribution of different cleft sentence types in discourse, two more narrow presupposition types are needed: C(onsciousness) Presupposition and T(opicality) Presupposition.

(135)   C-presupposition

"An entity or proposition is consciousness-presupposed (C-presupposed) if the speaker assumes that its mental representation has been activated in the interlocutors' short-term memory at the time of the utterance." (Lambrecht 2001:475)

(136)   T-presupposition

"An entity or proposition is topicality-presupposed (T-presupposed) if at utterance time the speaker assumes that the hearer considers it a center of current interest and hence a potential locus of predication. A topical denotatum is by definition a relatively predictable element in a proposition." (Lambrecht 2001:476)

2.4.2.2.1 Amalgam pseudoclefts

Consider the following canonical pseudocleft in a question-answer discourse (Lambrecht 2001:477).

(137)   A: What bothers you in Austin?

B: What bothers me is that it's so hot in the summer.

Lambrecht points out that although *it's so hot in the summer* is focused, it is not "new information" if the speaker and hearer in this context know each other to be Austin residents, and therefore familiar with its climate. The new information is thus not the focus, but the relationship between the focus and the presupposed open question: this is the assertive component of the sentence. (138) shows Lambrecht's analysis of the information structure of this pseudocleft.

(138)   Presuppositions   (Lambrecht 2001:477)

a.   K-presuppositions:

(i)   'x bothers the speaker (in Austin)'

(ii)   'it is hot in the summer (in Austin)'

b.   T-presupposition:

(i)   'KP (i) is of current interest in the discourse'

c.   Focus:

(i)   'it is hot in the summer' (=K-presupposition (ii))

d.   Assertion:

      (i)   'x=it is hot in the summer'

The relationship between the value and the variable is asserted—not simply the value or counterweight proposition. This model offers a useful insight into amalgams, in light of the fact that the assertive force of the copula is not always easily disentangled from the assertive force of the counterweight. The amalgam pseudocleft counterparts of (137) can serve the same functions.

(139)   a.   What bothers me about Austin is it's so hot in the summer.

          b.   It's so hot in the summer is what bothers me about Austin.

The relationship between the weight and the counterweight is the asserted new information.

What Lambrecht and Ross-Hagebaum (2006) hint at, but do not analyze in detail, is the fact that the reverse amalgam makes available an additional possibility. Consider (140).

(140)   He should really think more carefully about this, is his point.

In asserting (140), the speaker does not presuppose the counterweight. The speaker asserts both the counterweight, and the proposition relating it to the discourse, namely, that it is *his point*. In essence, the speaker can use (140) to explicitly assert the T-presupposition associated with the C-presupposed weight. The amalgam is used to make a metalinguistic comment.

In some instances, the weight strongly anaphoric. This is the case when it contains a concealed question DP anchor, as in (141).

(141)   She absolutely loves raspberry, is her favorite flavor.

(141) is felicitous in a context where *her flavor* has a linguistic antecedent in short-term memory, but it is much less felicitous if the open question about favorite flavors is suggested by the extralinguistic context.

It is also possible to utter a reverse amalgam without strongly presupposing the weight, in which case it may receive a focal pitch accent.

(142)   A: He's dead.

B: I know he's dead now.

A: Well, if you don't know, nobody does.

B: I should have brought in a replacement, is my f*cking POINT![31]

In (142), the weight receives a strong pitch accent. The speaker does not presuppose that the open proposition [my point is x] is of current relevance in the discourse—it asserts that it is. This type of amalgam is puzzling, from the perspective of both syntax and information structure, because it has multiple foci: the counterweight, the weight, and the assertion relating them.

2.4.2.2.2 *That's x is y*

The information structure of *that's x is y*, where the weight clause includes two variables, is also complex. Both the precopular and postcopualr clauses contain a focal element. Ross-Hagebaum (2004) proposes that *that's x is y* is an "information structure amalgam".

(143)   That's what I'm trying to do is go back to blonde.   (Ross-Hagebaum 2004:403 (5))

In uttering (143), the speaker K-presupposes the open proposition [speaker is trying to do x], and T-presupposes that the open proposition is of current interest. The pseudocleft string, *what I'm trying to do is go back to blonde*, contained in (143), conveys the assertion that the focus *go back to blonde* should be substituted for the variable in the open proposition. The identificational copular sentence string, *that's what I'm trying to do*, similarly T-presupposes that [speaker is trying to do x] is of current interest in the conversation.

---

[31] *Deadwood*. HBO. 2003.

The focus is the demonstrative *that*, and the assertion once again consists of substituting the referent of the focus *that* for the variable in the open proposition. Since *that* is referentially vague, and in some cases lacking an accessible linguistic antecedent, the amalgam construction conveys two assertions, each identifying the variable with a different focus, thereby indirectly resolving the anaphoric reference of *that*.

(144)   a.   x = that

         b.   x = go back to blond

It is important to note that my syntactic analysis of this construction does not treat the anchor+counterweight (*x is y*) string as a constituent, but nothing precludes the assertion corresponding to *x is y* from serving as an information structure unit.

2.4.2.3   The colon function of copular amalgams

Specificational sentences have a "colon function" (Higgins 1979); amalgams are particularly good at conveying what a colon does in writing. Speakers themselves tend to be aware of this. One non-linguist informant provided me with the following example and meta-linguistic commentary. She had prepared the script in (145) to read in an instructional video, but when she read it aloud, she produced (146).

(145)   Written script:

         "There are a few different ways to communicate in Office 365 and in this video I'm going to show you the most basic ways that I communicate with my coworkers, something I do very frequently all day**:** instant messaging in Lync 2013."

(146)   Spoken version:

         There are a few different ways to communicate in Office 365 and in this video I'm going

to show you the most basic ways that I communicate with my coworkers, something I do very frequently all day, **is** instant messaging in Lync 2013.

Her commentary to me was: "It kind of sounds like I was kicking off a new sentence with, 'Something I do very frequently all day...' but I think I was really saying, 'In this video I'm going to show you the most basic ways that I communicate with my coworkers, something I do very frequently all day: instant messaging in Lync 2013.' But I (inadvertently) said *is* for that colon." Lambrecht and Ross-Hagebaum (2006) likewise observe that the amalgam copula often serves a colon function (see also O'Neill 2012).

Amalgams with a root-like clause in precopular position often convey afterthoughts. It is common, for example, to find a long break, represented by ellipses in the examples below, between the precopular clause and the copula:

(147)　　a.　That's the thing ... is I need a break.

　　　　b.　I need a break ... is what I'm saying.

　　　　c.　You know what I need? ... is I need a break.

Calude (2008) examines the functions of pseudoclefts and identificational pseudoclefts like *that's x* in a corpus of 200,000 words, and finds that speakers frequently use these sentence types for the general purpose of "regulating the flow of the interaction and contributing to the general comprehension of the discourse" (p. 80). In sum, the copular amalgam clauses, because their logical subject has the form of a Force-bearing clause, are used to talk about talking. The structure I will propose for copular amalgams lends itself particularly well to this function.

2.5   The puzzle

Copular amalgams pose an analytical puzzle. Although they are quite similar to canonical pseu-doclefts in many respects, they are not easily subsumed under any standard analysis of copular sentences.

We have seen that in copular amalgam clauses, the structural subject can be a bare finite clause—a category that cannot normally serve as a subject, because it cannot enter an Agree relation with T.

(148)   a.   *He needs a break is too bad.

         b.   *It's raining upsets me.

Copular amalgams are also unusual in that their finite copula cannot combine with additional functional material. It must generally remain in bare form. If the copula were the run-of-the-mill spell-out of T, as it is in canonical specificational sentences, this would be quite unexpected. These problematic properties are real, and in need of explanation.

2.5.1   What you see is what you get

2.5.1.1   Repeated material is not structurally shared

One way to dispense with the problem of having a root clause for a subject is to say that copular amalgams are truly syntactic amalgams: two root structures overlapping on a shared constituent, like a graft (van Riemsdijk 2006).

(149)   He needs <a break> is what he needs.

         a.   He needs a break.

         b.   A break is what he needs.

An argument-sharing analysis will turn up empty, however. The distribution of argument sharing, as in Horn-amalgams, is extremely limited. For instance, although the examples in (150), below, are superficially similar to copular amalgams, they are ungrammatical.

(150) a. *That's <coffee>, is dark.

   b. *That's <Mary>, is a nice girl.

   c. *I would drink <that>, is hot and caffeinated.

In addition, one of the putative component propositions does not actually form a constituent. That is, in (149), *a break is what he needs* is not a constituent. Even in *that's x is y* sentences where only a DP value is overt, standard tests make it clear that the anchor, copula, and value do not form a constituent.

 Consider the example in (151).

(151) That's the issue, is the weather.

Suppose there were a constituent *the issue is the weather* related to the focal *that*. Predicate topicalization of this "constituent", as in (152), is absolutely out.

(152) *The issue, is the weather, that certainly is.

 Now suppose that this putative constituent is not a finite proposition, per se, but a non-restrictive relative-clause-like apposition containing the anchor and the value as one extended DP (e.g., Massam 2013; similar to the structure argued against in section 2.3.4). If the anchor, *the issue*, is pronominalized, as in (153), the value remains outside the domain of pronominalization, unlike in relative clauses, which form constituents with their heads.

(153) a. That's it, is the weather.

     b.   *I saw him, who is a nice guy.

2.5.1.2 Repeated material is not copied

A superficial look at repeated-material amalgams may also suggest the hypothesis that they are de-

rived from their canonical clause counterparts via copying, which, e.g., Bošković (1997a) proposes

for canonical specificational sentences and their simple clause counterparts. Aside from raising an-

alytical issues (what kind of copy operation would this be?), this hypothesis is easily falsified for

amalgams. The "repeated" material need not be strictly identical across the two clauses, (154).

(154)   a.   She**'d like** to leave early is what she really **wants** to do.

          b.   What she really **wants** is she**'d like** to leave early.

          c.   I really **enjoy** the oldies, is what I really **listen** to most of the time. (Lambrecht and

               Ross-Hagebaum 2006:9)

    Finally, not all of the structures considered here under the umbrella of the copular amalgam

construction include a surface amalgamation. In DP-weight amalgams, and others where the anchor

is propositional, the whole counterweight serves as the value, and no material is repeated.

(155)   a.   She'd like to leave early is the issue.

          b.   He wanted the vegetarian option was his preference.

These structures exhibit the same unusual properties as the other "repeated-material" amalgams:

bare sentential structural subjects, frozenness, and limited functional structure associated with the

copula. A multi-dominance or copying analysis of shared-material amalgams therefore offers little

insight into the unusual properties they share with DP-weight amalgams.

2.5.1.3 The weight and counterweight are integrated

One could also hypothesize that the counterweight is not syntactically integrated with the rest of the structure, something akin to a "radical orphan" approach (Haegeman 2009). Perhaps they are related to each other at the level of discourse instead.

If this is the right way of looking at copular amalgams, then at least for the types with a root-like clause in precopular position, the structure would be something like (156).

(156)   That$_i$'s [what I worry about]$_i$. e$_i$ is our finances$_i$.

The shared constituent is overtly expressed in the first clause, and binds a topical null subject in the second.

Such an approach captures the "afterthought" use of amalgams—the speaker asserts the first proposition, and follows it up with a second separate sentence. It also removes the puzzle of accounting for the occurrence of a bare finite clause in subject position. Nevertheless, there are reasons to reject an approach like the one represented in (156).

First of all, it predicts that negation should be possible on the amalgam copula, contrary to fact. While the structure in (157) is predicted to be good, the string in (157) is not.

(157)   *That$_i$'s not [what I worry about]$_i$. e$_i$ isn't our finances$_i$.

In fact, the grammatical *that's x is y* amalgam version of (157), (158a), cannot feature negation on the amalgam copula. The string in (158a) without the negation would be contradictory in its hypothetical unintegrated counterpart (158b).

(158)   a.   That's not what I worry about, is our finances.

          b.   *That$_i$'s not [what I worry about]$_i$. e$_i$ is our finances$_i$.

Another strike against a pro-drop analysis in a putative second sentence is that although spoken English does allow subject and expletive drop (159) (and even referential pronominal null subjects, particularly in so-called "diary" contexts; Haegeman 2013), it does not allow null subjects with the non-contracted form of the copula.

(159)   a.   Seems like a nice idea.

          b.   's raining.

(160)   a.   *Is a nice idea.

          b.   *Is raining.

The copula in amalgams is almost never contracted, which indicates that pro-drop is not the right approach to the structure of the second clause.

Indirect evidence against an unintegrated bisentential analysis of amalgams comes from cross-linguistic data. The type of topic-drop illustrated by the second sentence in (159) is quite readily available in other non-pro-drop languages like German (e.g., Rizzi 1994; Trutkowski 2010):

(161)   a.   Kennst du   den Hans?
                 know   you the Hans
                 'Do you know Hans?'

          b.   Na klar,   ____ ist mein Nachbar.
                 of course, ____ is my   neighbor
                 'Of course, he's my neighbor.'   (Trutkowski 2010:206 (1))

The relatively contentless *it* referring to an antecedent proposition in a discourse like (162) should be a prime candidate for such topic-drop.

(162)   a.   Er soll   ein neues Auto kaufen.
                 he should a   new   car   buy
                 'He should buy a new car.'

b.   ____ ist was ich meine.
      ____ is what I   think
      'That's what I think.'

Nevertheless, the counterpart of the English subject-initial amalgam is unacceptable in German, according to several speakers I surveyed.[32]

(163)   a. *Er soll   ein neues Auto kaufen ist was ich meine.
             he should a   new   car   buy   is what I   think
            'He should buy a new car is what I think.'

        b. *Er braucht ein neues Auto ist was er braucht.
             he needs   a   new   car is what he needs
            'He needs a new car is what he needs.'

        c. *Er hat kein Geld   ist das Problem.
             he has no   money is the problem
            'He has no money is the problem.'

The only way to render these grammatically is to use the anaphoric subject *das* in the second clause, so that there are two separate sentences present.

(164)   Er soll   ein neues Auto kaufen. Das ist was ich meine.
        he should a   new   car   buy.   that is what I   think
        'He should buy a new car. That's what I think.'

Given the sharp contrast between the acceptability of topic-drop and the acceptability of the amalgam pseudocleft in German, I conclude by analogy that topic-drop is not the source for the amalgam pseudocleft in English.

2.5.2   The bare sentential subject

As we saw in section 2.3.3.4, the subject clause in amalgams lacks the complementizer that enables finite clauses to serve as subjects in other contexts (compare (165) and (166)).

---

[32]I suspect that the ungrammaticality of this sentence type stems from the restrictions on true CP arguments in other Germanic languages (e.g., Koster 1978; Lohndal 2013). CP arguments are embedded within a DP structure, so they do not provide the root-like content needed for the interpretation of the amalgam pseudocleft.

(165)   a.   She reads a lot is what she does.

        b.   *That she reads a lot is what she does.

        c.   *For her to read a lot is what she does.

Bare finite clauses cannot serve as structural subjects of ordinary clauses, whether they are copular or not (166), so their occurrence in amalgam pseudoclefts (and other specificational copular amalgam sentences) must be explained.

(166)   a.   *She reads a lot is surprising.

        b.   That she reads a lot is surprising.

        c.   For her to read a lot would be surprising.

Also problematic is the role of the bare finite counterweight as logical subject. From a compositional perspective, a root-like finite clause cannot combine with an $<e,t>$-type property of individuals (denoted by canonical small clause predicates, vP, and TP), because it is an expression of type t.

There are a number of analyses on the market for what prevents root-like CPs from occurring as structural subjects (see Chapter 7 for detailed discussion). Either the root clause *lacks* a feature that enables it to Agree with T (Pesetsky and Torrego 2001; Landau 2007); it *has* a feature that makes it clash with T (Stowell 1981a); or it has a null complementizer that cannot be licensed in the structural subject position (Kayne 1981; Pesetsky 1982; Bošković and Lasnik 2003; Landau 2007, etc.). None of these approaches can be easily adapted to account for the distribution of sentential subjects in amalgams.[33]

---

[33]Note that assuming the precopular clause is a non-subject A′-position does not solve the problem, since dislocated sentences, like in-situ sentential subjects, must be introduced by a functional element.

(i)   a.   *They had invited him, everyone regretted.
       b.   It shocked everyone *(that) they had invited him.

2.5.2.1 Against *that*-drop

Suppose there were a way to reconcile the root-like subject with an analysis that normally prevents null-headed CPs from serving as subjects. This is a non-starter: the counterweight in an amalgam is distinct from the null-headed CPs that occur in, for instance, complement position. Root-like CPs in such environments alternate with headed CPs.

(167)   a.   I heard you left.

          b.   I heard that you left.

The counterweight of the amalgam pseudocleft, however, does not.[34] The paraphrase of a repeated-material amalgam pseudocleft with a headed CP is nonsensical; it forces an entity-like interpretation of the counterweight, forcing the whole counterweight to be interpreted as the value.

(168)   a.   [He needs a break] is what he needs.

          b.   *[**That** he needs a break] is what he needs.

                $\longrightarrow$ #'He needs the proposition that he needs a break.'

(169)   a.   What he ate is [he ate a bagel].

          b.   *What he ate is [**that** he ate a bagel].

                $\longrightarrow$ #'What he ate is the the proposition that he ate a bagel.'

The same absurdity arises if bare sentential weights, e.g., in *that's x is y*, include the complementizer.

---

[34]The exceptions to this are of course the DP-weight amalgams, and any pseudoclefts where the value is propositional. These counterweights alternate with headed CPs, but their structure is still quite different; see Chapter 4.

(i)   She left, is my problem.
        a.   [That she left] is my problem.

(170) a. [That's what he ate] is he age a bagel.

     b. *[**That** that's what he ate] is he ate a bagel.

        $\longrightarrow$ #'The proposition that that's what he ate has the value/property of being the proposition that he ate a bagel.'

Clearly, CP subjects in amalgams do not function like ordinary CP subjects.

2.5.2.2 Against direct quotation

There is an environment where a root-like clause can be found in preverbal position: direct quotation.

(171) "We should really leave soon," said John.

Suppose that at least some counterweight-initial amalgams are generated from their weight-initial counterparts via a quotative inversion process, similar to (171).

(172) a. My issue is: "we need to leave."

     b. We need to leave, is my issue.

This seems like a promising avenue, since quotative inversion and DP-weight amalgams have a similar profile. I will briefly argue that this is in fact not the right approach to amalgams. (In Chapter 8, I revisit the issue and propose that the similarities between the two structures are nevertheless not coincidental.)

If all counterweights were direct quotations, then any direct speech report should be quotable in counterweight position. This is not the case, however:

(173) "What a wonderful world!" exclaimed John.

(174) ??"What a wonderful world!" is my opinion.

Moreover, in quotation-fronting, both inverted and non-inverted orders of the remnant are possible, unlike in amalgams. Compare (175) and (176).

(175)   "I am so tired," John said.

(176)   *"We need to leave," my issue is.

(177)   *"He needs a break," what I mean is.

Amalgam pseudoclefts do not pattern syntactically with quotative constructions.

  Further evidence against treating amalgams as quotatives comes from the distribution of indexicals. In true direct speech reports, as in (178), indexical pronouns and adverbs are interpreted relative to the quoted utterance's context, rather than the matrix context, namely, that of the speech report (Sharvit 2008).

(178)   a.   "What will $I_{k/*I}$ do tomorrow?" she$_k$ wondered.

         b.   *"What will she$_k$ do tomorrow?" she$_k$ wondered.

         c.   "My$_{j/*I}$ car has been stolen!" exclaimed John$_j$.

         d.   *"His$_j$ car has been stolen!" exclaimed John$_j$.

In amalgams, indexical pronouns, adverbs, and tenses in the counterweight are interpreted relative to the matrix utterance context, as in (179), even where the whole counterweight serves as the value.

(179)   a.   What will $I_{*k/I}$ do tomorrow is what she$_k$ wondered.

         b.   What will she$_k$ do tomorrow is what she$_k$ wondered.

         c.   My$_{*j/I}$ car has been stolen is what John$_j$ exclaimed.

         d.   His$_j$ car ??has/had been stolen is what John$_j$ exclaimed.

The indexical present tense shows the same asymmetry. In quotations, and even in free indirect discourse, where first person need not be shifted, it is interpreted as present with respect to the local evaluation context; in amalgams, it must be interpreted with respect to the matrix context, and is in fact impossible under a past form of the copula.[35]

(180)   "I need a break," I thought.

(181)   *I am so tired was my problem.

Finally, in quotations, the speaker does to commit herself to the reported speech act. I can quote someone else's question without requesting information, and I can quote someone else's assertion without committing myself to its truth.

(182)   a.   He asked, "Who is the president?" But of course, everyone knows it's Barack Obama.

         b.   "Two plus two is five," they told me, but I knew it was a lie.

In an amalgam, by contrast, the speaker commits herself to the assertion expressed by the counterweight.

(183)   a.   #She needs to get a new job is what she needs. But I don't think she needs a new job.

         b.   #They should've left while they were ahead is the problem. But I think they should have stayed.

If the speaker reports anyone's speech act in the amalgam, it is his own.

---

[35]The facts are more complex; a detailed analysis is given in Chapter 8.

2.5.3 What is the copula doing?

The other central question in accounting for amalgams is: what is the copula's role? It bears finite

form ([tense] and [ɸ]), but it fails to combine with other functional baggage. It takes a bare finite

clause as its structural subject, just in its specificational instantiation. It relates two proposition-

denoting elements, much like a coordinator, although it is inflected.

It can encode the assertive Force of the whole proposition. The subordinate status of the coun-

terweight's force contrasts with other constructions with embedded roots (Krifka 2014), like inter-

rogative slifts, where the speaker is taken to be asking the putatively embedded root question, and

the two cannot have independent force (Haddican et al. 2014:100 (88)).

(184)   How old is she? did she say?

      a.   45.

      b.   *She did, 45.

      c.   45, but she didn't say it.

      d.   *Yes/No.

In the amalgam, the speaker can report the question expressed by the counterweight without per-

forming an interrogative act. Consider the possible responses in (185a–c).

(185)   What time should they leave for the train station? is what they haven't figured out yet.

      a.   Yeah, they'd better check the schedule soon.

      b.   No, they said they checked the schedule already.

      c.   They should leave at noon.

The two questions in (184) are posed by the speaker, while in (185), only the copular clause is

necessarily associated with Force. I return to this issue when I analyze the structural position of the copula in Chapter 5.

The other puzzling attribute of the copula is that although it is inflected for tense, it does not associate with semantic tense. That is, it does not encode temporal displacement of the proposition, unlike an ordinary verb.

(186)   a. *What I need **will be** I need a break.

   b. ***Yesterday**, what I needed was I needed a break. (on the reading where *yesterday* modifies the copular clause)

The absence of an independent temporal interpretation associated with the copula is crucial to the structure I propose, and is the main topic of Chapter 8.

2.6 Summary

The copular amalgam sentence shares the same basic configuration and function with its canonical specificational copular sentence counterpart, but its unusual syntactic composition makes it difficult to accommodate in standard theories.

Chapter 3

Methods

3.1 Grammaticality and acceptability

My aim in the next part of the dissertation is to present an exhaustive profile of the copular amalgam sentence, by comparing it to its better-studied counterpart, the canonical specificational copular sentence. First, a note on my approach to the data is warranted.

3.1.1 Copular amalgams are grammatical

Copular amalgam sentences have often been relegated to the periphery of the grammar. For instance, the reverse amalgam pseudocleft, exemplified in (1), is reported as ungrammatical in Akmajian (1979), Higgins (1979), Declerck (1988), and den Dikken et al. 2000.

(1)   He went to the store is where he went.

Higgins (1979:86) finds even *wh*-initial amalgam pseudoclefts "irremediably anacoluthic," setting a precedent in syntactic studies of specificational copular sentences.

In contrast to these reports in the literature, copular amalgams occur frequently in natural speech, informal written registers, and fictional dialogue in novels, television, and films. The fact that they occur frequently in fictional dialogue suggests that writers are aware of their effectiveness at representing naturalistic speech. Corpus research on *that's x is y* amalgams, *wh*-initial amalgam

pseudoclefts, and the double-*is* construction finds occurrences in the United States, Canada, the United Kingdom, Australia, and New Zealand (e.g., Massam 1999; Andersen 2002; McConvell 2004; Brenier and Michaelis 2005; Lambrecht and Ross-Hagebaum 2006; Curzan 2012; O'Neill to appear). In other words, copular amalgams are found across English—they are not an isolated phenomenon.

Double-*is* stands out as the most conventionalized copular amalgam sentence type; it is even gaining ground as a written form (Andersen 2002). It occurs across registers and national and socio-demographic speech communities. For example, one of a handful of Language Log posts observing double-*is* behavior documents President Obama's frequent use of double-*is* over the course of three 2012 Presidential debates (Zimmer 2012), e.g., (2).

(2) "And if you are going to save any money through what Governor Romney's proposing, uh, what has to happen **is, is** that the money has to come from somewhere."

A range of copular amalgam sentence types is in fact totally acceptable to English speakers. Moreover, their judgments are consistent with respect to the effect of different syntactic manipulations on the copular amalgam. Even speakers who judge baseline examples as degraded can reliably detect grammaticality contrasts between different copular amalgam forms. The stability of these relative judgments suggests that copular amalgams are indeed generated by the grammar.

Acceptability judgments provide a reliable source of data for investigating copular amalgams, but because a host of factors mediate between the performance of a judgment task and the mental grammar, judgments should be elicited under controlled conditions, whenever possible. In particular, the colloquial, or even stigmatized status of the copular amalgam negatively affects acceptability judgments. Since most speakers are unaccustomed to providing acceptability judgments of colloquial language, acceptability ratings are expected to skew somewhat lower than ratings of an

equivalently grammatical construction that belongs to a more formal register. Many speakers that I have consulted reject the amalgam pseudocleft as unacceptable, and then produce several tokens spontaneously in the same conversation.[1] Another common speaker comment is: "Well, I would never write that, but I guess I might say it." Prescriptive stigma is just one of the factors obscuring the relationship between acceptability and grammaticality.

3.1.2 Canonical specificational copular sentences are variable

A further complication in the present research is that *canonical* specificational copular sentence forms exhibit rampant inter- and intra-speaker variation. Since the canonical sentence type provides a structural baseline for analyzing the properties of the amalgam, this is an unwelcome confound.

Much of the early literature on canonical pseudoclefts reports stringent syntactic restrictions (e.g., Akmajian 1979; Higgins 1979; Williams 1983; Heggie 1988). This literature reports that pseudoclefts cannot be freely embedded, that their *wh*-clause weights cannot undergo raising to subject, that their copula must be in simple form, and that it cannot undergo T-to-C movement, for example. The standard methodology in syntactic research, especially in those decades, was to elicit binary judgments from just a few speakers. In some cases, only the judgments of the author are presented. Some authors report gradience, but only three levels of acceptability are typically distinguished, and discussion of relative acceptability is sparse.

Many of these early descriptions of the formal behavior of specificational pseudoclefts are not replicated under controlled experimental conditions. I will show in this work through a series of controlled acceptability experiments that canonical pseudoclefts are much more flexible than what is reported in the literature.

---

[1]Henry (2005) observes the same behavior among Belfast English speakers, and suggests methods for compensating for prescriptive stigma when the features under study are recognized as part of the local identity. Since copular amalgams signal no particular socio-demographic affiliation, it is unclear how to adapt these methods.

Gradience and variation are facts of life when it comes to acceptability judgment data. One of the linguistic factors conditioning the gradience and variation in judgments of pseudoclefts is the unavoidable structural ambiguity of the canonical pseudocleft string. A pseudocleft string like (3), for example, can be parsed as a canonical pseudocleft (den Dikken at al.'s 2000 "Type B" pseudocleft) or as a partially elliptical amalgam pseudocleft ("Type A" pseudocleft).

(3)   Where he went is to the store.

    a.   Where he went is to the store.   canonical

    b.   Where he went is ~~he went~~ to the store.   elliptical amalgam

The effect of this ambiguity on the acceptability of certain pseudocleft strings was not taken into consideration in the early literature. For instance, the amalgam pseudocleft is not compatible with negation (Chapter 4), so if a speaker parses a canonical pseudocleft string as an amalgam, and encounters negation, she may judge the string as degraded. As far as I am aware, only den Dikken et al. (2000), Ross (2000), and Schlenker (2003) explicitly discuss the confounding relationship between the canonical and amalgam type source for pseudoclefts.

3.2   Sources of data for this dissertation

An empirical database fraught with stigma and variation presents a methodological challenge. What kinds of judgment data, if any, should be used? How should gradient and variable acceptability judgments be treated as evidence for grammaticality?

The present work confronts these challenges by integrating data from a variety of sources. It draws on a mixture of naturalistic and constructed examples, as well as informally and formally obtained judgments.

Attested examples from the Corpus of Contemporary American English (Davies 2008–), web

searches, and fictional dialogue are used to illustrate the surface formal properties of sentence types, and to demonstrate their role in discourse. In some cases, I report naturally occurring examples that I have overheard. These examples are rendered verbatim, where possible; otherwise, they are reconstructed to the best of my recollection.

When available, corpora give a useful picture of the relative frequency of different forms. The absence of certain forms is suggestive, but it does not entail that the form in question in ungrammatical (see Schütze 1996, 2009; Cowart 1997; Henry 2005, Sprouse and Almeida 2010; Schütze and Sprouse 2014 for discussion). In the case of the copular amalgam, a corpus does not yield an exhaustive portrait. For example, it does not offer evidence for what is ungrammatical, and it is difficult to ensure that putative copular amalgam tokens are not instances of anacoluthon or disfluency. There is also a serious logistical barrier to using corpora for copular amalgam research: most copular amalgam sentence types have no particular searchable string. In the absence of widely available and comprehensive *parsed* corpora of spoken English, amalgam sentence types other than double-*is* and *that's x is y* do not lend themselves to large-scale searches.

Acceptability judgments of idealized data form the bulk of the data presented here. Since I am a productive user of amalgam pseudoclefts, DP-weight amalgams, and *that's x is y*, my intuitions about baseline examples are robust and categorical.

Whenever possible, example paradigms that are not categorically (un)acceptable are tested experimentally. Experimental testing for acceptability compensates for the cognitive biases of the researcher, and controls for factors like repeated exposure to a structure, lexicalization, and sample size. The experiments consist of crowd-sourced acceptability judgment surveys. Surveys are designed according to the increasingly accepted best practices in syntactic research (e.g., Sprouse and Almeida 2010; Sprouse 2011b; Gibson et al. 2011; Schütze and Sprouse 2014; Erlewine and

Kotek to appear; see section 3.3).

In some cases, a large-scale survey was not feasible. For example, the reverse amalgam pseudocleft presents a garden path, which is costly to repair in reading. When the same pseudocleft is embedded, the garden path effect is worse: the counterweight is parsed as the complement of the verb, and the string introduced by the copula cannot be integrated with the parsed structure. The square brackets in (4) show the relevant processing unit.

(4)   [I heard that you needed a new job] was what you needed.

The string is particularly difficult to integrate in reading, given the length of the sentence. The processing cost associated with such sentences makes it difficult to obtain meaningful acceptability judgments. In cases like these, I opted for small-scale surveys, conducted orally with a small number of speakers.

Other judgment questions were too complex to be investigated in a web-based survey. For example, it is cognitively taxing to interpret and judge structures involving multiple levels of temporal embedding, e.g., (5) (based on an example in Abusch 1988; Chapter 8).

(5)   Last week, John decided that in ten days he would say to his mother that what he needed was he needed his own place.

An informant needs detailed instructions, feedback from the experimenter, and excellent concentration. The target data relate not only to syntactic acceptability, but also to semantic interpretation. The anonymous setting of the web-based survey cannot easily get at such data.

Instead, when a quantitative survey was not appropriate, I used traditional informal methods, consulting three to five speakers. To improve the accuracy of the judgments, I focused on patterns of relative acceptability, comparing target sentences to each other and to a baseline. When possible,

I consulted the same speakers over a period of time, to improve the precision of the findings. An advantage of this method is that speakers often provide unsolicited insights about the meaning and function of the sentences (Matthewson 2004).

3.3 Formal acceptability experiment design

This section briefly outlines the experimental design methods I employed in the surveys reported in this dissertation.[2]

3.3.1 The nuts and bolts of AMT

The Internet is an invaluable source of linguistic data. Population samples for traditional acceptability judgment surveys are quite limited, often consisting only of university students or linguists. Tools like Amazon Mechanical Turk (AMT) have gained popularity among social science researchers in recent years, because they grant fast and low-cost access to a varied and vast sample.

In AMT, the researcher "Requester" posts a task "Human Intelligence Task", and solicits a particular number of participants "Workers". Workers are identified to the Requester only by their Worker ID number, but they have the option of sending direct messages to the Requester if they have questions or concerns. AMT allows the Requester to specify "Qualifications", criteria that the Workers must meet in order to accept a task. For example, Workers in my surveys were all adults with USA-based IP addresses. Qualifications can also screen Workers by their skill level. AMT records how many tasks each Worker has completed, and what percentage of these tasks was accepted by the Requester. Workers with particularly strong records are designated as "Masters". For a higher commission fee to AMT (20%), the Requester can make the task available only to Master Workers. I have found that results are just as good when I rely only on Workers who have

---

successfully completed more than 5000 tasks with at least a 95% task approval rate.[3]

The quality of the linguistic judgment data obtained via AMT is just as good, if not better, than judgment data obtained using other recruitment platforms, e.g., university psych pools composed of students enrolled in introductory courses (see Sprouse 2011b for a thorough validation of AMT as a source of judgment data). Out of a sample of 351 participants across several different surveys, only two submitted responses that showed a lack of understanding or attention to the task, which is a remarkably low rejection rate.[4]

AMT's (2011) Best Practices Guide for Requesters makes several recommendations in order to protect Requesters and Workers. Before accepting a task, Workers read a description, including screening criteria, an estimation of the task's duration, criteria for approval of the results, and the amount they will be paid. (In the survey itself, they give informed consent, according to institutional standards for the responsible and ethical conduct of human research.) The Requester decides how much to pay workers per task; the standard is to base the payment on an estimated hourly rate. Although the payment for comparable tasks is often much lower than the federal minimum hourly wage, my target was always $8.00/hour. The Requester-Worker relationship is reputation-based, so it is important to approve or reject results promptly and according to explicit criteria. For instance, if a technical glitch prevented a Worker's results from being submitted successfully, they were still paid for their participation, and their task approved, in order to protect their reputation (as Masters or >%95 approved). I made payments and responded to all Worker concerns (typically pertaining to server errors) within 24 hours.

In the present research, all surveys were hosted externally to the AMT environment, but AMT also offers internal survey tools (see, e.g., Gibson et al. 2011; Sprouse 2011b; Erlewine and Kotek

---

[3]I thank Jon Sprouse for suggesting this screening strategy.

[4]Several other sets of responses were excluded for other reasons, which are explained on a case-by-case basis in Chapter 4.

to appear).

3.3.2   Ibex Farm

The surveys in this research were hosted by Ibex Farm, a web-based platform designed by Alex Drummond. Ibex Farm gives the researcher a great deal of control over survey design and the style of item presentation. The interface for survey participants is spartan, minimizing distractions and mistakes.

Each sentence item to be judged is presented by itself in the middle of the screen, with a numerical Likert scale below it. Ibex Farm does not allow participants to backtrack, so controls for ordering effects were crucial to the experimental design.

3.3.3   Item design

Each of the surveys reported in this dissertation tested a maximum of three different factors. The basic method in linguistic analysis is to identify minimal contrasts. If two items differ in just one syntactic feature, and they have different acceptability, then that syntactic feature is the culprit. Applying this principle to quantitative survey research requires either a very small number of factors or a very large number of items.

In carefully constructed traditional judgment tasks, contrastive analysis typically motivates the use of paradigms. A consultant is presented with a large set of minimally distinct examples, and asked to judge them. While this method can yield reliable results, particularly if the judgments are robust, it is problematic when the contrasts are subtle. The effect of repeated exposure (to similar structures and to the same lexicalizations of those structures) can desensitize the informant to the relevant contrasts. It then becomes difficult to replicate the judgments over a larger number of sentence tokens. For this reason, linguists can be unreliable informants; most of us at some

point have exclaimed, "I'm losing my judgments!" The use of counterbalancing techniques and between-subjects comparisons in formal experiments can control for this repeated exposure.

Cowart (1997) recommends that an acceptability survey testing simple interactions include no fewer than 4 tokens from each experimental condition (member of the test paradigm). Crucially, each of these tokens should have the exact same structure, but use different lexical items to instantiate it. This lexical variation controls for any unanticipated quirks of individual items that may introduce error into the results, and it weakens the effect of repeated exposure. In order to compare members of the experimental paradigm to each other, however, the lexicalization must be held constant. The conflict between the goals of presenting minimal contrasts and avoiding repetition with the same speakers motivates the use of a Latin square design, with between-subjects analysis.

In a Latin square design, blocks of items are composed by arranging lists of item paradigms into a square. Each list uses a different lexicalization, but includes the same conditions. Consider a hypothetical 2x2 design (four experimental conditions). Four paradigms are constructed to instantiate those conditions with different lexical items. The four lexicalization paradigms are arranged in a square (e.g., conditions 1–4; lexicalizations a–d). Items are then distributed across four blocks in a diagonal pattern (e.g., Block 1: 1a, 2b, 3c, 4d; Block 2: 1b, 2c, 3d, 4a). The resulting blocks represent one token from each paradigm, and one token from each lexicalization. Since each survey version needs to include four tokens of each condition, sixteen lexicalizations are constructed. The Latin square assignment is applied to each set of lexicalizations: 1–4, 5–8, 9–12, and 13–16, yielding sixteen blocks. The resulting lexically counterbalanced blocks are then distributed among versions by the same procedure, with four blocks in each version. The final result is that each version includes sixteen experimental items, controlling for presentation order, including a single member of each lexicalization and four distinct tokens from each experimental condition. This

counterbalancing procedure was used in all of the surveys presented in Chapter 4.

Although individual participants do not judge paradigms of examples, between-subjects comparison serves as a proxy, allowing paradigmatic example sets to be compared. They also control for anomalous lexicalizations, items, and participants, facilitating outlier analysis.

Another crucial control is the use of filler items. Fillers serve multiple purposes in survey research. They help maintain participants' attention by making the survey less repetitive, they catch errors, they encourage full use of the rating scale, and they provide benchmarks. The fillers used in the surveys presented here were evenly divided into acceptable and unacceptable items, which allowed each subject's results to be normalized. The surveys used the same number of fillers as experimental items. The Latin square procedure summarized above included blocks of fillers, which were randomized and interpolated into the experimental item blocks. A 2x2 design therefore included 32 items in total.

3.3.4 The measurement scale

An advantage of formal acceptability experiments over the traditional binary judgment task is that they can capture more nuanced contrasts in acceptability (Schütze 1996; Cowart 1997; Keller 2000; Sprouse and Almeida 2010; Schütze and Sprouse 2014). The cognitive system is not directly accessible to introspection—even if the principles of grammar are not gradient, the output of an acceptability judgment is. A binary task is often too crude to distill the effect of the grammar from other factors affecting the performance of the judgment.

There is some disagreement as to what sort of measurement scale and what sort of task is best suited to detecting these contrasts. For instance, Bard et al. (1996) and Keller (2000) use Magnitude Estimation to simulate a continuous scale. In Magnitude Estimation, a method often used in psychometric research, the participant judges items relative to a benchmark, and controls

the size of the scale, making proportional judgments as the survey progresses. For example, if they rate a benchmark item with an arbitrary value "100", and a subsequent item at "200", they can freely rate a third item at "10,000" if that proportion seems appropriate to them. The idea is that an unbounded scale is better at measuring perception than a discrete Likert scale or forced choice task; participants can decide in response to the data which contrasts are relevant, and how large they perceive them to be.

Sprouse (2011a) criticizes the use of Magnitude Estimation in linguistic acceptability experiments. The purported advantages of Magnitude Estimation are not demonstrated in validation studies. Sprouse (2011a) shows that participants tend to fall back on a more traditional discrete scale, even when they are trained to use the freer scale. In view of the limited advantages, if any, of using a more logistically complex method, I opted to use a Likert scale to measure acceptability, with endpoints labeled "unnatural" and "natural". The use of controlled fillers as the basis for normalization eliminated any quirks in individual participants' use of the scale.

3.3.5 Instructions

The same instructions were used for all of the acceptability surveys in this research. Instructions were presented as illustrated in Figure 3.1. Although it is tempting to be more forceful in coaching participants to avoid making judgments based on prescriptive norms, Cowart (1997) observes no significant effect of such instructions on judgments, and additional instructions may in fact cause confusion.

In each survey, participants were trained on three practice items, with feedback giving open-ended suggestions about how to use the range of the scale. They then rated six unmarked "warm-up" items that were excluded from the analysis. Items were presented one-at-a-time with no con-

Instructions

You will see a series of approximately [X number of] sentences in English. Your job is to read each sentence aloud and rate how natural you find it on a scale of 1–9.

Some of the sentences sound completely natural, some sound completely unnatural, and some are in between.

A rating of 1 means: "This sentence sounds completely unnatural in spoken English—this is not something an English speaker would say."

A rating of 9 means: "This sentence sounds completely natural in spoken English—this is something an English speaker would say."

Feel free to use the full range of the scale. There are no right or wrong answers, since we want to know what sounds natural to you in a casual context, not what is considered to be correct or proper English.

No specific context is given for each sentence. Feel free to imagine each sentence in any context that makes sense to you.

Figure 3.1: Instructions for acceptability judgment surveys

text. Participants read each sentence in full, and clicked a box in the Likert scale to rate its acceptability and advance to the next item. When they finished the task, they answered three "debrief" questions about their own usage of amalgam sentence types. Once they entered a completion code into AMT, and the results were approved, they were paid for their time. A sample item is depicted in Figure 3.2.

3.4 Analysis of results

First, raw results for each participant were converted to z-scores based on the fillers. Normalization controls for individual skew and variability in participants' use of the rating scale. For example, if a participant is biased toward high ratings, using the maximum rating (9) for acceptable sentences, and the medial rating (5) for sentences that are lowest in acceptability, a rating of (7) means some-

What she likes seems to be she likes the bacon.

| 👎 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 👍 |

Click a box to answer.
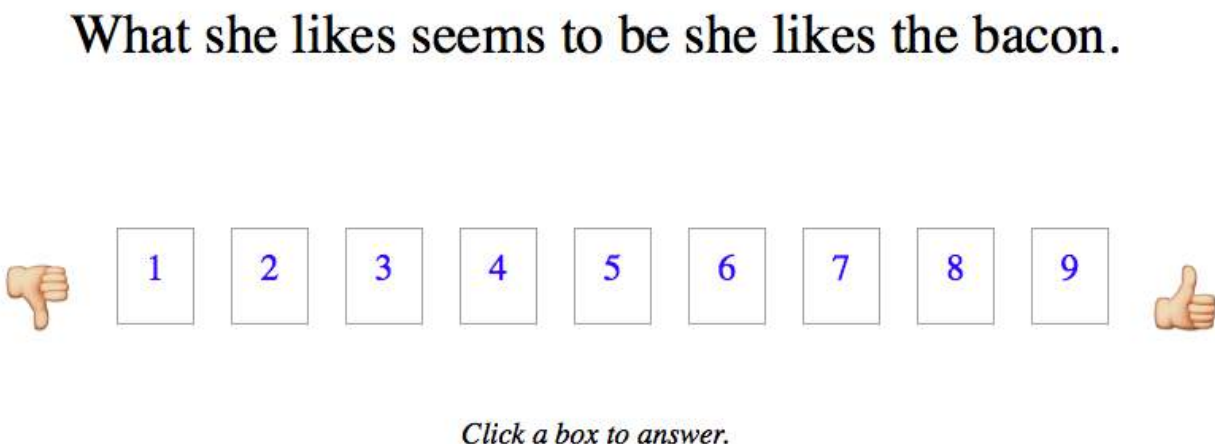
Figure 3.2: Sample item presented in Ibex Farm

thing different than it does for a participant who uses the full rating scale. I make three types of inferences based on these normalized judgments.

The first approximates the categorical binary judgment of grammaticality. The normalized midpoint (z=0) marks a non-arbitrary (but crude) threshold of acceptability for a given participant: it is the linear midpoint between totally acceptable and totally unacceptable. The mean normalized judgment for any given sentence type can be identified as at least somewhat acceptable or at least somewhat unacceptable relative to this threshold. I will report, for example, that the effect of aspectual auxiliaries on the amalgam copula reduces acceptability into the negative range, while aspectual auxiliaries on the canonical copula reduce acceptability, but not into the negative range. My interpretation of this result is that aspectual auxiliaries are *grammatical* in canonical pseudo-clefts, although they are less *acceptable* than the bare form of the copula.

The second type of inference arises from simple pairwise acceptability comparisons (t-tests). For example, there is a significant negative effect of adding modal and aspectual auxiliaries to the *wh*-initial amalgam pseudocleft (t=-4.3).

Finally, linear mixed effects models are used to test for interactions between sentence type

(e.g., canonical vs. amalgam) and other factors. The models are generated by statistical packages in R (R Core Team 2013). The surveys presented here primarily ask whether a given syntactic feature, e.g., temporal auxiliaries on the copula, has the same effect on canonical specificational sentences as it does on amalgams. The number of effects considered in each model is quite small. Interactions are identified by comparing a "null" model with a "test" model. The "null" model includes fixed effects for the linguistic factors in the experimental conditions (at most three factors) and the random effects of subject, item, and lexicalization. The "test" model is the same, but adds a fixed effect for the interaction of the two linguistic factors. The significance of the interaction term is detected using an ANOVA of the null model and the test model. For example, a model with the interaction between the future auxiliary factor (sentences with and without *will* associated with the copula) and the sentence type factor (canonical vs. amalgam sentences) accounts for significantly more variance in the judgment data than a model without this interaction term.

3.5 Geographical distribution of copular amalgams

A total of 351 AMT Workers participated in the surveys reported in this dissertation. Interestingly, there are no particularly salient demographic trends with respect to the acceptability of the baseline amalgam pseudocleft. Speakers from diverse regions of the U.S.A. accepted them, and similarly, speakers from all over rejected them. One subtle pattern is the lower ratings of the reverse amalgam pseudocleft among speakers in the Midland region. There is no salient effect of speaker age or gender either.[5]

The maps below illustrate, for a subset of the responses, the geographical distribution of the judgments, based on participants' self-reports of their place of origin. For each participant, the

---

[5]I leave further quantitative analysis of demographic factors for future research. One factor that will be interesting to explore is the relationship between normalized judgment means and speaker's self-reports in the debrief questions.

mean acceptability of baseline (e.g., bare copula) amalgam and reverse amalgam pseudoclefts was calculated. These mean values were geocoded, using CartoDB, for all participants who entered a recognized city of origin.[6]Mean ratings above the normalized baseline are plotted as "accepted", and ratings below the baseline are plotted as "not accepted". Figure 3.3 shows the acceptability of the *wh*-initial amalgam pseudocleft, and Figure 3.4 shows the acceptability of the counterweight-initial (reverse) amalgam pseudocleft.
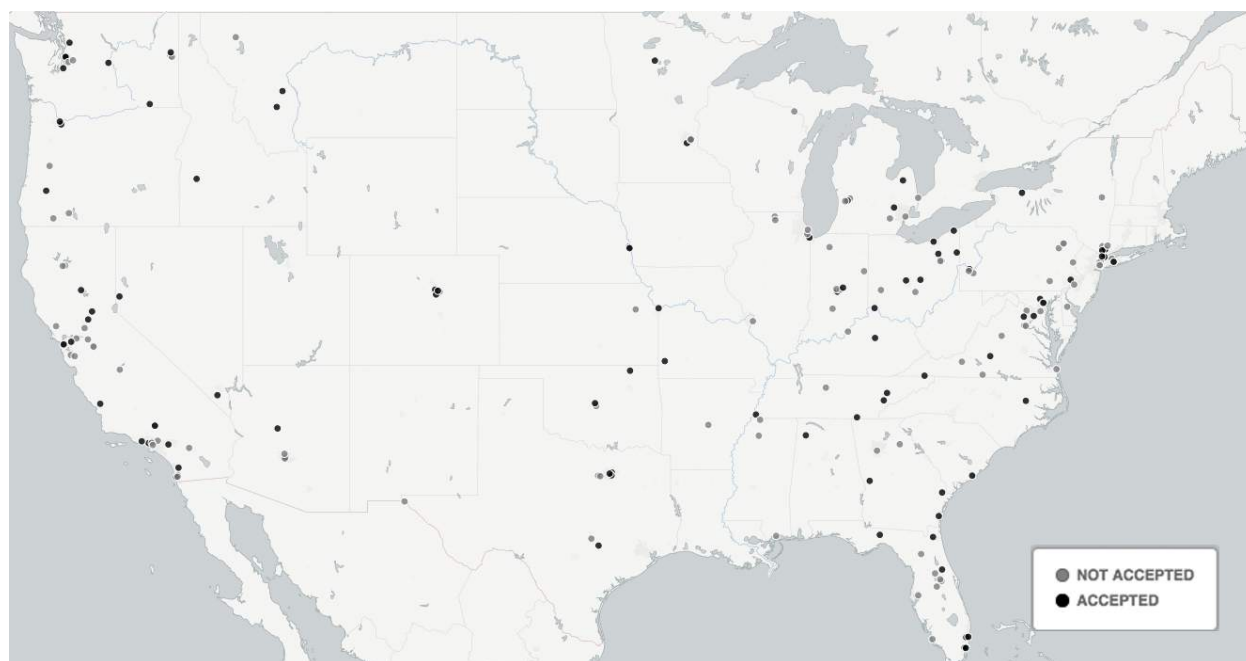
Figure 3.3: Acceptability of *wh*-initial pseudoclefts

3.6 Summary

By bringing together data from a variety of sources, and testing subtle acceptability contrasts experimentally, this dissertation presents the first thorough empirical profile of the copular amalgam construction in English. It further illustrates how experimental methods can be brought to bear on marginalized linguistic features in syntactic research.

---

[6]Some participants indicated vague regions, e.g., "Western USA", which were not plotted.
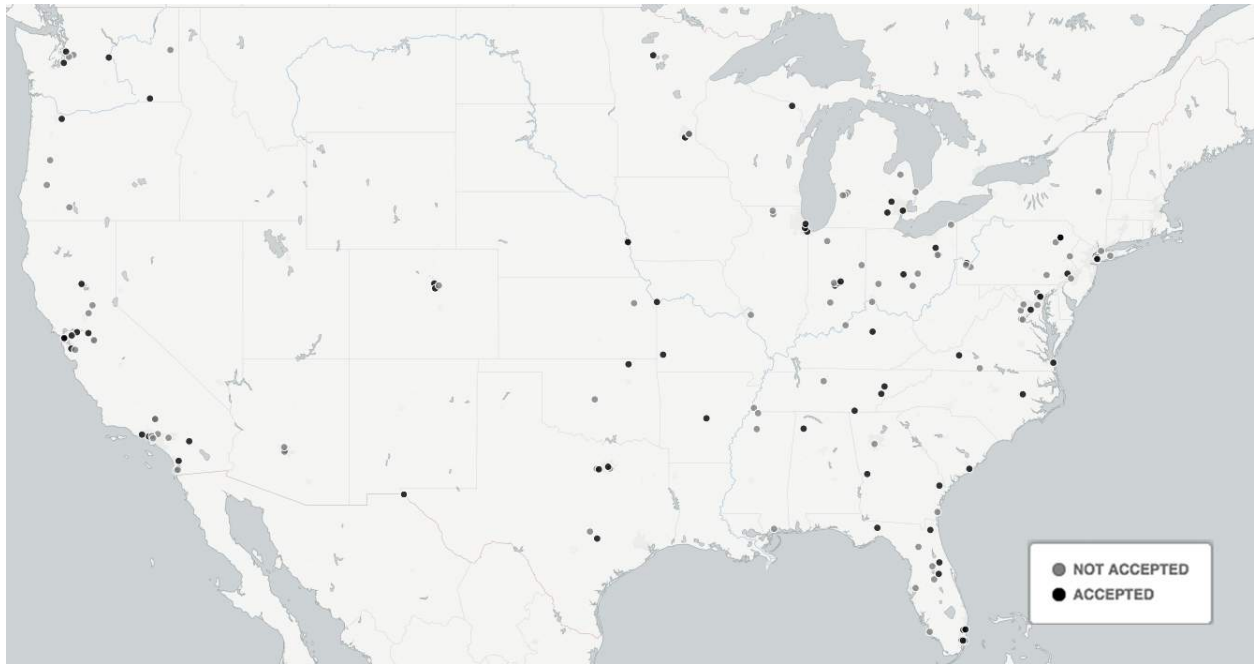
Figure 3.4: Acceptability of reverse amalgam pseudoclefts

Chapter 4 turns to reporting the findings of these empirical comparisons of the canonical and amalgam sentence types. Then, Chapter 5 develops a syntactic analysis that accounts for the differences between them.

PART 2: The Analysis

That depends on what your definition of 'is' is.

Chapter 4

Canonical versus amalgam pseudoclefts

4.1 Introduction

Specificational copular amalgam sentences have not yet received a comprehensive analysis in generative syntax. This second part of the dissertation develops such an analysis. The contribution of the present chapter is primarily empirical: its purpose is to contrast copular amalgam sentences with canonical specificational copular sentences.[1] Sections 4.2–4.5 examine syntactic properties, including movement, embedding, and the form of the copula. Finally, section 4.6 compares the semantic properties of canonical and amalgam pseudoclefts.

I will assume a predicate-inversion model of canonical specificational copular sentences, as discussed in Chapter 2. The predicate-inversion model of canonical specificational sentences provide a practical starting point for the analysis of amalgams, since the two sentence types are formally and functionally similar. The model of canonical sentences offers a set of null hypotheses for how the syntax of amalgams might work. The present description focuses on the syntactic features that allow these hypotheses to be tested. Since the amalgam pseudocleft/DP-weight amalgam has a direct structural counterpart in the canonical pseudocleft, while the other amalgam copular sentence types do not, this chapter focuses on the pseudocleft/DP-weight amalgam type.

---

[1]Ross (2000) gives an empirical profile of several specificational sentence types, including some of the same properties, but he does not explicitly distinguish the canonical type from the amalgam type.

An additional goal of this chapter is to replicate and clarify the syntactic behavior of canonical pseudoclefts. Other than Ross (2000) and Heycock (2012), few analyses of canonical pseudoclefts take careful stock of speaker variation and gradient acceptability, which has led to some inconsistencies in the literature. I speculate that some of the variation, particularly in the judgments of weight-initial pseudoclefts, is due to the inherent structural ambiguity of canonical pseudocleft strings. They can be parsed as either canonical copular clauses or partially elliptical amalgam clauses. While it is usually not possible to disambiguate in favor of the canonical pseudocleft, when the two surface forms are compared directly, clear contrasts emerge.

4.2 Moving parts

4.2.1 Reversibility

Reversibility is the hallmark of specificational copular sentences: either the logical subject or the predicate may occupy the precopular position. The examples below demonstrate the reversibility of both double-NP and pseudocleft specificational sentences.[2]

(1) a. The most important thing is the budget.   (anchor < value)

    b. The budget is the most important thing.   (value < anchor)

    c. What John wants is a vacation.   (anchor < value)

    d. A vacation is what John wants.   (value < anchor)

The same property holds of DP-weight and pseudocleft amalgams:

(2) a. The most important thing is we need to revise the budget.   (anchor < value)

    b. We need to revise the budget is the most important thing.   (value < anchor)

    c. What John wants is he wants a vacation.   (anchor < value)

---

[2] As always, the term "specificational" describes the meaning of the sentence, not its word order properties.

    d.   John wants a vacation is what he wants.   (value < anchor)

Now consider the other copular amalgam sentence types: *that's x is y*, the question-answer amalgam, the free-*be* amalgam, and double-*is*. These types are crucially *not* reversible: they occur only in the weight-initial order, as in (3).

(3)   a.   [That's what John wants]$_{weight}$ is [(he wants) a vacation].

        b. *[(John wants) a vacation] is [that's what John wants]$_{weight}$.

Suppose that the amalgam pseudocleft string within the *that's x is y* sentence (*x is y*) were a constituent excluding *that's*, as implied by the pragmatic analysis of the construction in Ross-Hagebaum (2004) and Calude (2008). We might then expect the pseudocleft part of the string to be reversible, leaving *that's* at the left edge of the sentence. This reversal remains impossible, as (4) illustrates, reinforcing the conclusion reached in Chapter 2 that the *that's x* string forms a constituent separate from the counterweight string.

(4)   *That's (John wants) a vacation is what he wants.

The fixed weight-initial order of the other amalgam sentence types is illustrated in the following examples.

(5)   Question-answer amalgam

        a.   [You know what she found]$_{weight}$, is [(she found) his drug stash].

        b. *[(She found) his drug stash] is [you know what she found]$_{weight}$.

(6)   Free-*be*

        a.   [I'm explaining why I'm asking the question]$_{weight}$ is that it seems like from the rela-

tionship that you had with her that you would have been one of those people.[3]

b. *That it seems like from the relationship that you had with her that you would have

been one of those people is [I'm explaining why I'm asking the question]$_{weight}$.

(7)   Double-*is*

a.   [The problem]$_{weight}$ is, is (that) he left early.

b.   (That) he left early is <*is> [the problem]$_{weight}$ <*is>.

The striking ill-formedness of the counterweight-initial versions of these amalgams suggests a

syntactic difference between these and amalgam pseudoclefts. It also sheds light on the status of

amalgams in languages like German, Dutch, and American Sign Language. In these languages,

and for some speakers of English, the weight-initial amalgam pseudocleft is acceptable, but the

reverse amalgam pseudocleft is categorically unacceptable (den Dikken et al. 2000; Caponigro and

Davidson 2011).[4] What these non-reversible amalgam sentence types have in common, excluding

double-*is*, is that their weight clause is a root, and it must occur in precopular position.

Predicate-inversion approaches to specificational copular sentences account for reversibility by

allowing either the subject or the predicate to raise out of a small clause into a higher position.

In Chapter 2, I adopted an A-movement model of predicate inversion (e.g., den Dikken 2006).

An inverted predicate must be a topic, but it need not occupy a left-peripheral Topic position in

canonical clauses (contra Heggie 1988). In the non-reversible amalgam sentence types, as I will

argue in the next chapter, the Topic-Comment order is base-generated. Because of this, inversion

is neither motivated nor possible. The other properties considered in this section follow from the

---

[3]"Serial". Narrated by Sarah Koenig. Episode 6. This American Life. NPR. October 20, 2014.

[4]Caponigro and Davidson (2011) in fact argue that the ASL equivalent of the amalgam pseudocleft is not a pseudocleft, but a "question-answer clause", because the full range of indirect question clauses, including polar questions, can occupy the initial position. The relevant point is that ASL has an amalgam sentence type, but lacks the reverse version with the counterweight in initial position (Kate Davidson, p.c.).

inversion-plus-A-movement analysis of canonical copular sentences.

4.2.2 A′-movement

A general resistance to A′-movement sets specificational pseudoclefts apart from predicational copular sentences (Akmajian 1979; Higgins 1979; Williams 1983; Heggie 1988; Moro 1997; Heycock and Kroch 1999; den Dikken 2006). Canonical and amalgam copular sentence types are subject to extraction restrictions. Focus movement out of both canonical and amalgam copular sentences fails. Topicalization in canonical specificational sentences is licit, provided it does not disrupt the fixed information structure of the construction. Subject-auxiliary inversion is also licit in canonical specificational sentences. In amalgams, by contrast, all of these movements are impossible. Judgments reported in this section are categorical, so no experiments were undertaken.

4.2.2.1 Focus movement

The focus of a pseudocleft cannot undergo long A′-extraction:

(8)   a.   What John is is tall.

      b.   *[How tall]$_i$ is what John is e$_{is}$ t$_i$?

      c.   *[How tall]$_i$ what John is is t$_i$!   (Higgins 1979:279 (2))

(9)   a.   What I don't like about John is his tie.

      b.   *[What/which tie]$_i$ is what you don't like about John e$_{is}$ t$_i$?   (Higgins 1979:279 (3))

      c.   *[His tie]$_i$ what I don't like about John is t$_i$.

Even in double-DP specificational sentences, the postcopular focus cannot extract.

(10)   a.   *[Whose child]$_i$ do you think the problem is t$_i$?

      b.   That school has nice teachers and vindictive teachers. *[Which kind of teacher]$_i$ do

you think the problem is t_i? (Heycock and Kroch 1999:378 (50))

Reverse pseudoclefts show some of the same sensitivity to focus-extraction, but the data must be handled with care. Since canonical pseudoclefts are ambiguous between predicational and specificational interpretations, it is difficult to isolate the specificational type once the sentence has been manipulated. Consider, for example, the sentences in (11).

(11)   a.   John is the teacher.

       b.   Coffee is what she likes.

On the predicational reading, extraction of the subject is perfectly legitimate:

(12)   a.   Who$_i$ do you think t_i is the teacher?

       b.   What$_i$ do you think t_i is what she likes?

In the specificational version, however, the subject is a focus, and is therefore unable to serve as the focus of a different clause. Compare (13), where the focus cannot be extracted, to (14). (14) uses a value exhibiting bound anaphor connectivity with the weight clause, to ensure the specificational reading.

(13)   a.   She only likes [coffee]$_F$.

       b.   *[What$_i$]$_F$ do you think that she only likes t_i?[5]

(14)   a.   Important to himself$_j$ is what John$_j$ is.

       b.   *[Important to himself$_j$]$_i$ nobody said t_i is what John$_j$ is, not humble.

---

[5]It is telling that (13b) is acceptable on the reading where the verb *likes* associates with focus; e.g., *What do you think that she only LIKES, but does not actually LOVE?*

The same set of restrictions on focus movement holds for amalgam pseudoclefts. (The amalgam structure is obscured by the simple *wh*-form of the counterweight clause, so the echo-question form in (15c) illustrates the ungrammaticality of questioning the focus more clearly.)

(15)   a.   What I don't like about John is I don't like his tie.

        b.   *What$_i$ is what you don't like about John e$_{is}$ t$_i$?

        c.   *You don't like [which tie]$_i$ is what you don't like about John e$_{is}$ t$_i$?

        d.   *[I don't like his TIE]$_i$ what I don't like about John is t$_i$.

The reverse order, likewise, resists focus extraction:

(16)   a.   I don't like John's tie is what I don't like.

        b.   *[She doesn't like John's tie]$_i$ nobody said t$_i$ is what she doesn't like, not [his poor choice of cologne]/[she doesn't like his cologne].

In the non-reversible amalgam types, the same property holds: the focus clause cannot be moved. The focus-movement examples below could generously be described as word salad.

(17)   a.   That's what I worry about is (I worry about) paying rent.

        b.   *What$_i$ is that's what you worry about e$_{is}$ t$_i$?

        c.   *[You worry about which expense]$_i$ is that's what you're worried about e$_{is}$ t$_i$?

        d.   *[She worries about paying rent]$_i$ that's what she worries about is t$_i$.

(18)   a.   You know what she worries about, is (she worries about) paying rent.

        b.   *What$_i$ is you know what she worries about e$_{is}$ t$_i$?

        c.   *[She worries about paying rent]$_i$ you know what she worries about is t$_i$.

(19)   a.   The problem is, is he forgot his lunch.

    b. *What$_i$ do you think the problem is, is t$_i$?

Subextraction out of the subject of a specificational copular sentence is ill-formed.

(20)   a. *Who$_i$ do you think [the beauty of t$_i$] was the cause of the war?

          b. *Who$_t$ do you think the cause of the war was [the beauty of t$_i$]?

(21)   a. *[Which fruit]$_i$ do you think [she ate t$_i$] was what she ate?

          b. *[Which fruit]$_i$ do you think what she ate was [she ate t$_i$]?

(22)   *Where$_i$ do you think that's the author of the play is [a man from t$_i$]?

(23)   *[How big of a piece of cake]$_i$ do you suppose guess what she ate was [she ate t$_i$]?

(24)   *What$_i$ do you think the problem is, is [he forgot t$_i$]?

The ungrammaticality of examples like these is not surprising, given that they involve extraction out of a subject island and the implausible act of questioning an element of a focus that the speaker is supplying precisely to answer a question. Even in the more plausible scenario of non-*wh*-extraction, these structures are totally out.

(25)   a. *Helen$_i$ [the beauty of t$_i$] was the cause of the war.

          b. *Helen$_i$ the cause of the war was [the beauty of t$_i$].

(26)   *Germany$_i$ that's the author of the play is [a man from t$_i$].

(27)   *[His green card]$_i$ guess what he finally got is [he got t$_i$].

(28)   *[His lunch]$_i$ the problem is, is [he forgot t$_i$], (not his briefcase).

It is also well known that the weight of a pseudocleft cannot undergo focus movement (Higgins 1979; Declerck 1988; Heggie 1988; Moro 1997; Heycock and Kroch 1999; Ross 2000; Rothstein 2001; Mikkelsen 2005; den Dikken 2006; and others), as illustrated in (29).

(29)   a.   Imogen thinks that the best candidate is Brian.

       b.   *[How good a candidate]$_i$ does Imogen think t$_i$ is Brian? (den Dikken 2006:125 (103c))

Chapter 2 described how the predicate-inversion approach to specificational copular sentences accounts for the frozenness of the weight clause when it occurs in the inverted position. Since the weight expression is obligatorily interpreted as a topic, it cannot be dislocated to a position where it receives a focus interpretation—this is incoherent (Mikkelsen 2005; den Dikken 2006).

The non-inverted order bars focus-movement of the predicate for a similar reason. Although the subject is not frozen in an A′-position, specification requires focus on the value; hence, the weight cannot be focused. (30) includes a bound anaphor in the value to force the specificational interpretation; (31b) has an acceptable predicational counterpart where *Mary* is a topic. The subscribed *F* indicates that in the relevant specificational reading, *Mary* is focused.

(30)   a.   [Important to himself]$_F$ is what John is.

       b.   *[What John is]$_F$ important to himself is.

(31)   a.   Mary$_F$ is the tallest girl in the class.

       b.   *[The tallest girl in the class]$_F$ Mary$_F$ is.

Contrast these with focus fronting of the predicate in a predicational sentence, which is grammatical.

(32)   a.   Mary is [a tall girl]$_F$.

b. [A tall girl]$_F$ Mary is.

This flexibility is not found in specificational sentences, because they have fixed information structure, while predicational sentences do not.

The weight of an amalgam is likewise barred from undergoing focus movement. Amalgams are always specificational, so the counterweight clause is always the focus. The reason for this will be discussed further in section 4.6: they relate an expression with a question denotation to an expression with an answer denotation. Focus fronting of the weight clause, regardless of the base order, is ungrammatical in amalgams. I illustrate this in (33)–(35) for DP-weight amalgams, and subsequently for some non-reversible amalgams.

(33) a. The biggest problem is we don't have any more money.

b. *[How big of a problem]$_i$ do you think t$_i$ is we don't have any more money?

c. *[The biggest problem]$_i$ nobody said t$_i$ is we don't have any more money.

(34) a. We don't have any more money is the biggest problem.

b. *[How big of a problem]$_i$ do you think we don't have any more money is t$_i$.

c. *[The biggest problem]$_i$ nobody said we don't have any more money is t$_i$.

(35) a. We don't have any more money is the biggest problem.

b. *[The biggest problem]$_i$ we don't have any more money is t$_i$.

(36) a. That's my main worry is we don't have any more money.

b. *What$_i$ do you think t$_i$ is we don't have any more money?

(37) a. The biggest problem is, is he forgot his lunch.

b. *[How big]$_i$ do you think [the t$_i$ problem] is, is he forgot his lunch?

Neither canonical nor amalgam specificational sentences can be the input to focus movement.

4.2.2.2 Topic movement

Now consider the effect of topicalization in specificational sentences. Topic movement is generally more free than focus movement. For instance, the cartographic model of the clause allows Topic heads to iterate, while Focus is unique (Rizzi 1997, 2004, 2010). There are restrictions on extraction *across* a Topic, but there is no uniqueness requirement. As long as topicalization does not set up an information structure clash or violate locality constraints, it is acceptable (see discussion in den Dikken 2006).

Recall from Chapter 2 that the acceptability of long-topicalization of the predicate provides an argument in favor of treating the landing site of predicate inversion as an A-position.

(38)   a.   [The best candidate]$_i$, Imogen thinks t$_i$ is Brian.   (den Dikken 2006:125 (103d))

        b.   [The best candidate]$_i$, which everyone thinks t$_i$ is Brian, will have a lot to look forward to.

While the fronted predicate of a canonical specificational sentences is certainly interpreted as a topic, it does not occupy a Topic position. If it did, it ought to observe "criterial freezing". In a cartographic framework, once the expression satisfies a Topic-Criterion (checking a discourse-feature in the left periphery), the A$'$-chain can be extended no further (Rizzi 2006, 2010). Topicalization of a topicalized expression fails (39), in contrast to (38), which involves topicalization out of an A-position.

(39)   *Into the room$_i$, nobody claimed (that) t$_i$ there walked a circus clown.

The amalgam counterparts to (38) are similarly ungrammatical. Negative quantifiers are used in the matrix subject in order to prevent the matrix clause from being interpreted as a parenthetical.

(40)   *[What it needs]$_i$, nobody thinks t$_i$ is it needs paint.

(41)   *[That's what it needs]$_i$, nobody said t$_i$ is it needs paint.

(42)   *[You know what we should do]$_i$, nobody thinks t$_i$ is we should leave early.

(43)   *[What she thought e$_{is}$]$_i$, nobody said t$_i$ e$_{is}$ is she should quit.

Topicalization of the predicate in a relative clause structure shows the same asymmetry between canonical and amalgam specificational copular sentences.[6]

(44)   [What she likes]$_i$, which everyone agrees t$_i$ is coffee, could make her sick.

(45)   a. *[What she likes]$_i$, which everyone agrees t$_i$ is she likes coffee, might shock you.

        b. *[That's what she likes]$_i$, which everyone agrees t$_i$ is she likes coffee, might shock you.

Unlike canonical pseudoclefts, amalgams do not allow their weight constituent to be topicalized. The weight clause of an amalgam shows the behavior Heggie erroneously predicts for canonical weights; therefore, they seem to occupy an A$'$-position.

4.2.2.3 Subject-auxiliary inversion

An additional difference between the canonical and amalgam sentence types is the availability of subject-auxiliary inversion. In the seminal works on canonical pseudoclefts, it was claimed that the weight-initial order cannot be the input to inversion. Judgments in the examples below are from their original sources.

(46)   a. *Is what John is proud?

---

[6](44b) is of course also ruled out by the fact that the root clause *that's what she likes* cannot serve as the subject of the ordinary predicate *might shock you*.

    b.   *Was what Mary did (to) wash herself?

    c.   *Was what he wanted to marry an actress?   (Higgins 1979:283 (19)–(22))

Similarly, Faraci (1971:69 (62)–(63)) contrasts inversion in predicational sentences with inversion in specificational sentences.

(47)   a.   Is where John is going a nice place?

       b.   *Is where John is going San Francisco?

Several researchers have since pointed out that these judgments do not reflect most speakers' intuitions (Williams 1983; Hedberg 1993; Heycock and Kroch 1999; Ross 2000, 2004; den Dikken 2006; etc.). I suspect that any degradation in the examples in (46) can be explained on a case-by-case basis. For example, VP-value canonical pseudoclefts are often dispreferred in favor of their amalgam counterparts. If the preferred parse of the version of (46b) without *to* is an amalgam pseudocleft, for example, then it is not surprising that some speakers should judge inversion as degraded.

    Hedberg (1993) gives the following as perfectly acceptable examples of subject-auxiliary inversion in pseudoclefts; I confirmed these judgments in informal surveys.

(48)   a.   Is what you're writing on clefts or pseudoclefts?

       b.   Isn't where he's going San Francisco?

Den Dikken (2006:156 (160)) gives additional examples.

(49)   a.   Was the cause of the riot a picture of the wall?

       b.   Is the biggest problem the factory closings?

The availability of this inversion confirms the analysis of the fronted predicate as occupying an A-position (contra Heggie 1988), in contrast to fronted topics, where auxiliary inversion is not possible:

(50)   *Did in that cave, a monster live?

In specificational amalgams, by contrast, subject-auxiliary inversion is categorically ungrammatical.[7]

(51)   a. *Is what she ate she ate some pizza?

        b. *Is she ate pizza what she ate?

(52)   a. *Is the main issue you lost the game?

        b. *Is you lose the game the main issue?

(53)   *Is that's what she means she's quitting?

(54)   *Is guess who's coming over John (is coming over)?

The structure of the copular amalgam does not allow this inversion, suggesting that either the pre-copular constituent does not occupy Spec,TP, or the copula's ordinary spell-out position is not T. I will argue that *both* factors are behind the unavailability of subject-auxiliary inversion in amal-

---

[7]Interestingly, subject-auxiliary inversion *within* the weight clause is possible in *that's x is y*:

(i)   Is that what she means? is she's quitting?

I represent this example with two question marks to signal that both sentence boundaries are pronounced with rising intonation, and both have interrogative force. Similarly, double-*is* clauses marginally allow inversion:

(ii)   Is the issue is that he quit?

The fact that the weight clause can include interrogative force and inverted syntax further demonstrates that it is a constituent independent of the counterweight. I defer further discussion of the distribution of illocutionary force in amalgams and double-*is* until Chapter 5. The story for double-*is* is complicated by the fact that ordinary double-*is* sentences may be structurally ambiguous. On their simplest parse, such inversion is not available.

gams.

4.2.3 A-movement

The facts associated with A-movement out of specificational copular sentences are more subtle. This subsection discusses the distribution of raising to subject in different sentence types.[8]

4.2.3.1 Variation in the literature

Judgments of A-movement out of canonical pseudoclefts are sometimes reported as categorical, but the facts are actually nuanced. Higgins (1979) reports that raising is unacceptable in the *wh-*initial order, though it is acceptable in the value-initial order. Judgment diacritics below represent those given by Higgins (1979:285 (28)–(31)) for the specificational reading.

(55)   a. *[What John is]$_i$ tends t$_i$ to be conceited t$_i$.

        b. *[What John was]$_i$ was unlikely t$_i$ to be tall t$_i$.

        c. *[What Mary did]$_i$ began t$_i$ to be to wash herself t$_i$.

        d. *[What I don't doubt]$_i$ seems t$_i$ to be that anyone left t$_i$.

The lexical heterogeneity of the raising predicates in this example raises suspicion for Higgins's (1979) empirical claim. These examples, particularly (55c) in my own judgment, do show some degree of degradation, but the problem is not with A-movement, but with the meaning of the raising predicates. When considered on a case-by-case basis, there are other explanations for their marked status.

In (55a) and (55c), for example, the aspectual meanings of *tend* and *begin* clash with the meaning of specification. The episodic meaning of *tend* is incompatible with the intensional, time-

---

[8]Passive is of course unavailable in copular sentences, since they are not transitive. Even Rothstein (2001) and Mikkelsen (2005), who assume that the copula is associated with a *v* head, are clear that it is not a potential accusative-case assigner.

dependent interpretation of the weight clause *what John is*. It is not referential, so it does not have different states that can serve as the bearers of the property ascribed by *tend*. Compare (55a) to the non-specificational sentence with an intensional subject in (56), which illustrates the same clash (the non-restrictive relative is included to make the intensional reading prominent):

(56)   *The best editor, which I am looking for, tends to return manuscripts promptly.[9]

In (55c), the inchoative interpretation of *begin* is similarly incompatible with the intensional weight expression; it requires a referential subject. Even in a double-DP specificational sentence, inchoative is very odd:

(57)   *[The president]$_i$ began t$_i$ to be Barack Obama t$_i$.

Non-copular sentences with intensional subjects behave the same way.

(58)   *A skilled contractor, which is hard to find, began to work early in the morning.

(55d) differs from the other cases, since there is nothing wrong with an evidential predicate mediating specification. The problem with (55d) is pragmatic: it is odd for a person to ascribe to herself a lack of doubt toward some proposition, while at the same time hedging it with *seem*. This example improves dramatically when it is changed to third person.

(59)   [What he doesn't doubt]$_i$ seems t$_i$ to be that anyone left t$_i$.

The degraded examples of raising to subject observed by Higgins do not constitute solid evidence that raising to subject is syntactically unavailable in pseudoclefts.

Williams (1983) gives examples, marked ungrammatical, which differ from Higgins's (1979) above, since they pose no semantic problem:

---

[9]This sentence is acceptable only on the generic reading.

(60)　　a.　[Important to himself$_j$]$_i$ seems t$_i$ to be t$_i$ what John$_j$ is.

　　　　b.　*[What John$_j$ is]$_i$ seems t$_i$ to be important to himself$_j$ t$_i$.

Den Dikken (2005b:50 (184)) gives additional examples, citing Hankamer (1974):

(61)　　a.　*What they were doing seems likely to be claimed to be washing each other.

　　　　b.　*What she never wants to hear again appears to be any more WH Cleft sentences.

Hedberg (2008), on the other hand, points out the following examples, claiming that raising is acceptable (citations are hers).

(62)　　a.　[What he's after]$_i$ seems t$_i$ to be her money t$_i$.

　　　　b.　[What he's asking] appears t$_i$ to be whether there will be any beer t$_i$.　(Gundel 1977)

　　　　c.　[What John is]$_i$ seems t$_i$ to be proud of himself t$_i$.　　　　　　(Culicover 1977)

　　　　d.　[What John wants]$_i$ seems t$_i$ to be never to be left alone t$_i$.　　　(Halvorsen 1978)

Once again, when additional factors are taken into account, raising does not seem to be the sole culprit. To my ear, the examples in (61) have not earned their categorical negative judgments. (61a) presents an instance of long A-movement that is difficult, likely for processing reasons, even in a predicational sentence.

(63)　　?What they were doing seems likely to be claimed to be a bad idea.

Since the nonreferential, intensional meaning of the anchor in the specificational version must be recovered after several levels of embedding, perhaps (61a) is slightly worse than (63). The former also includes a long-distance reciprocal anaphor, which may contribute to the processing load. An additional disfavoring factor is that (61a) has a VP-value, which tends to favor an amalgam parse. (61b) includes an NPI, which similarly favors the amalgam parse, and does not strike me as

categorically out.

As den Dikken (2005b:50) points out, "The empirical status of specificational pseudoclefts of the type in [(61)–(63)] is a contentious issue in the literature." This is an empirical issue, straightforwardly addressed in a controlled experiment.

4.2.3.2 Experiment 1 design

Because of the heterogeneity of judgments demonstrated above, and the relatively unexplored status of raising in amalgams, I conducted a formal acceptability experiment to test the effect of raising on the two pseudocleft types. Since the degraded status of canonical pseudoclefts with raising in examples drawn from earlier literature (e.g., Higgins 1979; Akmajian 1979; Williams 1983; Heggie 1988) can be explained with reference to other factors, like the lexical aspect of the raising predicate and structural ambiguity with an amalgam-type source, canonical pseudoclefts are predicted to be relatively acceptable in a controlled judgment task. These earlier works did not control for lexical factors or inter-speaker variation. Amalgams, however, are predicted to be relatively unacceptable. A combination of within-subjects and between-subjects analyses of controlled judgment data will show that these predictions are borne out: canonical pseudoclefts occur in raising-to-subject constructions, while amalgam pseudoclefts in this environment are degraded.

The raising experiment employed the methodology described in Chapter 3. I will summarize the application of the method to the raising experiment. Throughout the rest of this chapter, as a series of similar experiments are presented, the discussion of the design will be more abbreviated.

The survey used a 2x4 factor design, crossing a raising factor with a sentence-type factor. The raising factor included two levels: bare (simple copular structure) with raising (embedding of the copular structure under the raising predicate *seem*). Sentence type included four levels: canonical pseudocleft, reverse canonical pseudocleft, amalgam pseudocleft, and reverse amalgam

pseudocleft. Example stimuli are presented in Table 4.1.

	Bare	Raising
Canonical	What we got is the salmon.	What we got seems to be the salmon.
Rev. Canonical	The salmon is what we got.	The salmon seems to be what we got.
Amalgam	What we got is we got the salmon.	What we got seems to be we got the salmon.
Rev. Amalgam	We got the salmon is what we got.	We got the salmon seems to be what we got.

Table 4.1: Factor design for the raising survey

To implement the counterbalancing method laid out in Chapter 3, 32 different lexicalizations were used (8 conditions by 8 lexicalizations, for a total of 4 squares). The conditions and lexicalizations were distributed across 32 blocks by Latin square and randomized, with an equal number of fillers interpolated into each block. The blocks were then assigned to 8 versions by the same method. The blocking and randomization were used to control for ordering effects, and for random variance associated with specific lexical items.

Each version thus included 32 experimental items, with 4 tokens per condition. Each block also included 32 fillers, for a total of 64 sentences per version. Each item was judged by 4 different participants; each participant judged each condition 4 times; and no participant saw more than one item from a given lexicalization.

Participants were recruited via Amazon Mechanical Turk (AMT), and accessed the survey outside of the AMT environment. The survey was administered using the web-based survey platform Ibex Farm. They were instructed to rate the "naturalness" of the sentences they read. The same instructions were used for all of the acceptability surveys reported in the present chapter (see Chapter 3; Figure 3.1 for details). A total of 32 participants completed the survey.

4.2.3.3 Experiment 1 results

Results were converted into z-scores based on the distribution of filler item judgments. Since the fillers are evenly divided between totally acceptable and totally unacceptable, normalizing against the fillers provides a reliable indication of how a given participant uses the measurement scale. Normalization also establishes a crude "grammaticality" threshold, z=0. This threshold indicates the mid-point between what a participant judges to be "completely natural" and "completely unnatural". It provides a useful reference point for the discussion of relative acceptability and the relationship of acceptability to grammaticality, which is necessarily an approximation. An item or condition with a positive score is "closer" to a fully acceptable sentence than to a fully unacceptable sentence. I analyze such items as grammatical.

Outlier analysis by subject showed that 6 participants either rated both weight-initial and counterweight-initial bare-copula amalgams below a floor of z=-0.5 (0.5 standard deviations below the filler mean), or they gave erratic judgments of filler items. These 6 sets of results were excluded, leaving 26 completed surveys in the analysis.

The results of the survey are illustrated in Figure 4.1. These results exhibit the variation reported in the literature. While the reverse canonical pseudocleft shows almost no effect of raising, the *wh*-initial canonical pseudocleft shows decreased acceptability. The majority of the acceptability ratings are in the *positive* range, however, bearing out my initial prediction. In short, for most participants, *wh*-initial canonical pseudoclefts with raising are grammatical.

In the amalgam condition, by contrast, the raising factor pulls the judgment value below the normalized baseline. One caveat is that this batch of speakers happens to have a low tolerance for *wh*-initial amalgam pseudoclefts (note that the ratings for bare *wh*-initial amalgams skew low), so even a small difference between the bare and raising conditions would have this effect. Neverthe-
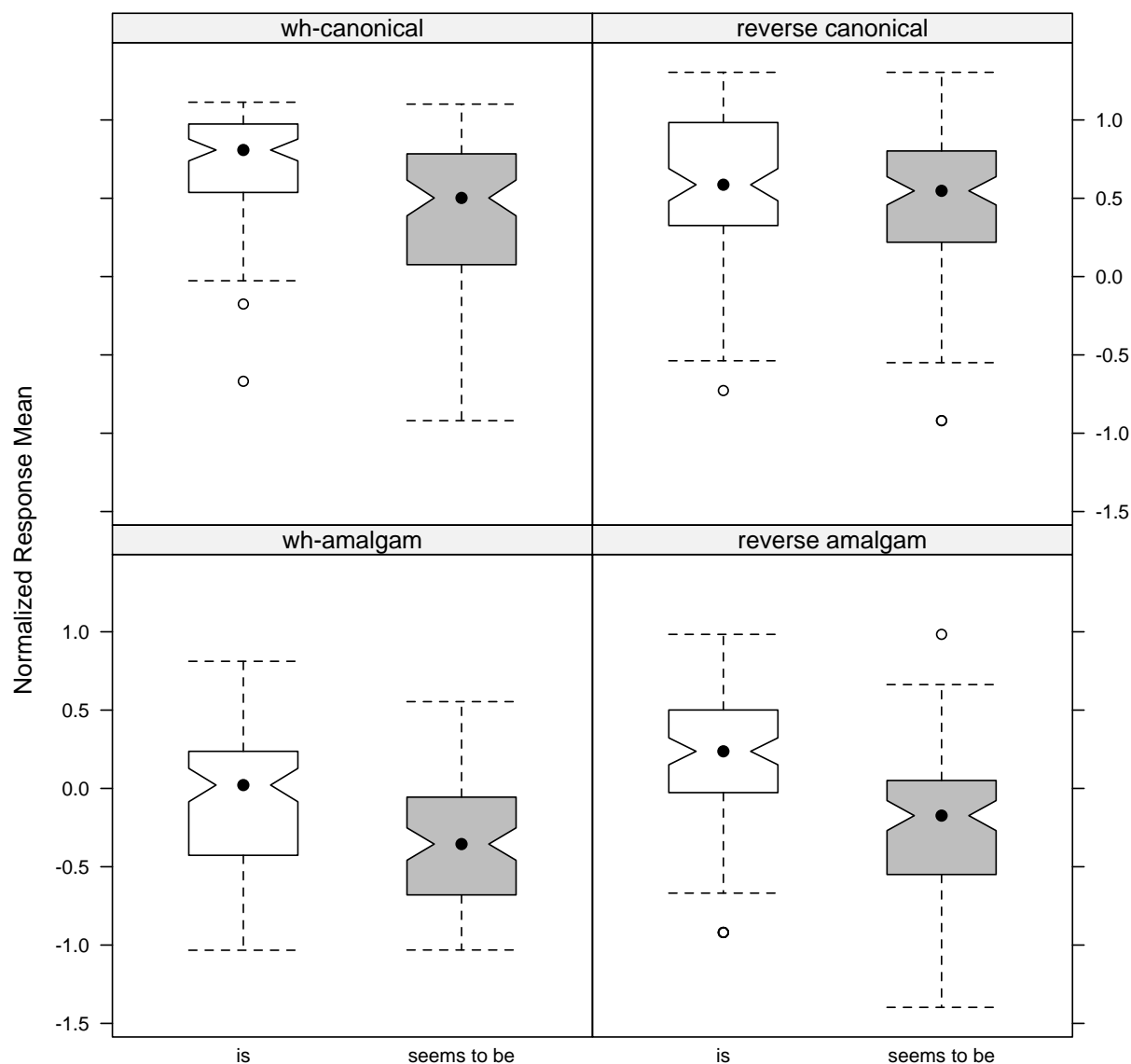
Figure 4.1: Effect of raising on specificational pseudoclefts

less, a comparison of the relative-to-baseline acceptability of raising across the sentence type factor

bears out the initial prediction. While raising receives a positive acceptability rating for canonicals,

it receives a negative acceptability rating for amalgams.

In the reverse pseudocleft condition, the asymmetry between the canonical and amalgam type is

striking. The reverse amalgam pseudocleft is judged as significantly worse when the counterweight

undergoes raising to subject than when it does not, while the canonical pseudocleft shows almost no difference. Both amalgam conditions show negative acceptability relative to the baseline when their precopular constituent undergoes raising.

A linear mixed effects model ("test") with random intercepts for subject and item, and fixed effects for sentence type, raising, and the interaction between sentence type and raising, was significant when compared with a "null" model omitting the interaction term ($p<0.05$) (Table 4.2).[10,11]

There is no interaction when only the weight-initial versions are considered; the interaction is significant only in the reverse pseudocleft condition. Table 4.3 shows the analysis for the reverse pseudocleft condition alone.

Model	Df	AIC	BIC	logLik	deviance	Chisq	Chi	Df	Pr(>Chisq)
Null	8	786.39	823.66	−385.19	770.39				
Test	11	784.14	835.39	−381.07	762.14	8.2514		3	0.04109*

Table 4.2: Fixed effects linear model of interaction between raising and sentence type

Model	Df	AIC	BIC	logLik	deviance	Chisq	Chi	Df	Pr(>Chisq)
Null	6	433.39	457.31	−210.70	421.39				
Test	7	428.87	456.78	−207.44	414.87	6.521		1	0.01066*

Table 4.3: Fixed effects linear model of interaction between raising and sentence type for reverse pseudoclefts

The tendency in the quantitative response data is confirmed by informal survey findings. Most informants rejected amalgams with raising predicates, but some accepted them, or gave them

---

[10]Function call for the analysis in R:

```
library(lme4)
data$sent.factor<-relevel(data$type.factor, ref="revcanonical")
test.lmer<-lmer(response.norm ~ raising*type + (1|subject)+(1|item), data)
null.lmer<-lmer(response.norm ~ raising+type + (1|subject)+(1|item), data)
anova(test.lmer, null.lmer)
```

The same types of models are used for the other experiments as well, so the specific code is not reproduced for each analysis.

[11]Significance codes in tables: 0 '∗∗∗' 0.001 '∗∗' 0.01 '∗' 0.05 '.' 0.1 ' ' 1

marginal status. Additional informal surveys show a reduction in acceptability for the other amalgam sentence types as well. Some speakers (who accept baseline amalgams whole-heartedly) reject them categorically, and some tolerate them. I will use the traditional question mark diacritic to indicate their reduced acceptability.

(64)   a.   ?That's the problem, seems to be we ran out of money.

         b.   ?Guess what he did, seems to be he quit his job.

         c.   ?He should give that a try, seems to be he should try cooking.

In summary, amalgams with raising are unacceptable when compared with an acceptability baseline, in contrast to canonical pseudoclefts with raising. For some speakers, amalgams with raising are marginal or close to acceptable. This result will ultimately constitute a (principled) exception to the tendency for the copula in amalgams to remain bare. I will return to this issue in Chapter 5.

4.3 Missing parts

Canonical and amalgam copular sentences differ with respect to whether they allow the copula to be omitted or elided.

4.3.1 Optionality

In varieties of English without zero-copula, the copula must be pronounced in canonical specificational sentences, regardless of order.[12]

(65)   a.   John *(is) the best candidate.

---

[12]An interesting question I leave for future research is whether zero-copula Englishes have copular amalgam sentences, and if so, whether zero-copula in amalgams is subject to the same constraints as in canonical specificational contexts.

  b. The best candidate *(is) John.

(66) a. A bagel *(is) what John ate.

  b. What John ate *(is) a bagel.

Likewise, DP-weight and pseudocleft amalgams, particularly in the counterweight-initial order, require an overt copula.

(67) a. We ran out of money *(is) the problem.

  b. The problem *(is) we ran out of money.

(68) a. John ate a bagel *(is) what he ate.

  b. What John ate *(is) he ate a bagel.

Some examples of a weight-initial pseudocleft without the copula can be observed in casual speech. Take, for example, the following sentence from television dialogue.

(69) All I'm trying to do right now, I'm trying to avoid eye contact with the CIA director.[13]

Sentences like (69) are rare, and sound like hanging-Topic constructions. They require characteristic topic intonation, and a strong boundary between the initial clause and the main clause.

In the non-reversible amalgam types, not only is the copula optional, but its absence has no discernible effect on interpretation.

(70) a. That's what John ate (is) (John ate) a bagel.

  b. Guess what John ate (is) (John ate) a bagel.

  c. John might want to order that next time (is) (John might want to order) a bagel.

---

[13]*The West Wing*. Television. 2000.

Formally, the copula-less versions appear to consist of two syntactically unintegrated sentences, but they stand in the same specificational relationship regardless of the copula's presence or absence: their semantic, pragmatic, and prosodic features are fixed. When the copula is absent, the sentences are in a paratactic relationship—a type of coordination (see Chapter 5 for further discussion).

4.3.2 Gapping

Another restriction on specificational sentences observed in Higgins (1979) concerns gapping. For example, while gapping of the copula is perfectly acceptable in predicational sentences and double-DP specifications (71), it is less acceptable in *wh*-initial pseudoclefts (72) (see discussion in den Dikken 2005b:348).

(71) a. John is the teacher, and Mary _____ the principal.

  b. The teacher is John, and the principal _____ Mary.

(72) a. Important to himself is what John is, and important to herself _____ what Mary is.

  b. *What John is is important to himself, and what Mary is _____ important to herself.

Hedberg (2008) gives the following example analogous to (72b), judging it to be acceptable.

(73) A: So what does Jane want, and what does Bill want?

  B: What Jane wants is never to go out, and what Bill wants, never to stay home.

The availability of gapping in *wh*-initial pseudoclefts seems to be sensitive to information structure—Hedberg's example is situated within a context where there are two open questions. Gapping in a specificational context is improved with this pragmatic support. Moreover, informal surveys do not confirm the categorical contrast between (72a) and (72b). Speakers tend to accept gapping in examples like (73), but they reject it in pseudoclefts with anaphor connectivity, like (72).

Gapping is not testable in the amalgam sentence types shown in the previous discussion to have an optional copula, but it can be tested for amalgam pseudoclefts. (The relevant parse of (74b) is of course not the one with the hanging topic parse.)

(74)   a. *John ate a bagel is what he ate, and Mary ate a salad _____ what she ate.

          b. *What John ate is John ate a bagel, and what Mary ate _____ Mary ate a salad.

The judgments of these examples are categorical, for speakers I surveyed informally. Because of the length of the test sentences, the gapping feature does not lend itself to a long-form controlled experiment. Given the impracticality of running a controlled experiment, and the categorical informal judgments, I take the result in (74) to be sufficient evidence that copula gapping in amalgams is illicit.

The difference between the two sentence types stems from the fact that the copula in amalgams is not in the domain of the ellipsis operation responsible for gapping.[14] As I will argue in Chapter 5, the amalgam copula does not occupy a V or T-domain position.

4.4 Embedding

4.4.1 Finite contexts

Canonical pseudoclefts are easily embeddable in finite contexts, particularly under bridge verbs.

(75)   a.   She said that coffee was what he liked.

          b.   She said that what he liked was coffee.

Non-bridge verb environments also allow both types of canonical pseudocleft:

(76)   a.   She regretted that coffee was what she drank.

---

[14]Even if gapping is not ellipsis (Johnson 2009), it involves V-domain structure not present in amalgams.

       b.   She regretted that what she drank was coffee.

Amalgams show a general preference for root contexts, but they can be embedded under verbs that encode a speech or mental event (see Bianchi 2003; Giorgi 2010).

(77)   a.   She said/supposed/believed that he liked coffee was what he liked.

       b.   She said/supposed/believed that what he liked was he liked coffee.

Double-*is* sentences, likewise, can only be embedded under bridge verbs. (The examples labeled COCA are from the Corpus of Contemporary American English; Davies 2008–.)

(78)   a.   We **decided**, by the way, that the only way they're going to let the news media count the ballots in the state of Florida was is if they threw you and Bob in the same room together and let you do it and that would be acceptable.   (COCA)

       b.   I **guess** the question is, is is it right to use our open seas as a testing lab?   (COCA)

(79)   *She **regrets** that her mistake is, is that she forgot her lunch.

4.4.1.1   Experiment 2

To verify the judgments of pseudocleft examples, I conducted a small-scale controlled survey. The survey compared root pseudoclefts with pseudoclefts embedded under bridge verbs, examining the effect of embedding on both canonical and amalgam pseudocleft. I predicted no interaction: the two sentence types are expected to behave the same way with respect to embedding. The embedding factor had three levels: bare (no embedding), first person, and third person. Since amalgams are frequently used to make metalinguistic comments, I hypothesized that embedding under first person verbs would be preferred to embedding under third person. This prediction was ultimately not borne out.

The garden path encountered at the first sentence boundary in the counterweight clause makes embedded counterweight-initial amalgams extremely difficult to process in reading (80). Because of the processing difficulty, the sentence type factor only included the weight-initial versions.

(80)   a.   [She supposed [that [he liked coffee]]] *was what he liked.

        b.   [She supposed [that [he liked coffee] was what he liked]].

The factor design for the survey is illustrated in Table 4.4.

	Bare	First person	Third person
Canonical	What Bill watches is documentaries.	I believe that...	Gail believes that...
Amalgam	What Bill watches is he watches documentaries.	I believe that...	Gail believes that...

Table 4.4: Factor design for embedding experiment

All the participants preferred the root versions. Embedding in both canonical and amalgam pseudocleft types caused a similar reduction in acceptability. Judgments of embedded amalgam pseudoclefts collected at the baseline (z=0), but variation is robust. There was no effect of sentence type, or of the person of the embedding verb's subject, confirming the results of informal consultations. Because the sample was so small (n=8), and no interaction was expected, no statistical analysis was performed on these results.

4.4.1.2   Embedding under non-bridge verbs

The canonical and amalgam sentences types contrast with respect to their acceptability under non-bridge verbs. For instance, informal consultations with amalgam users yielded an acceptability contrast for embedding under manner-of-saying verbs.

(81)   a.   She whispered that coffee was what he liked.

        b.   She whispered that what he liked was coffee.
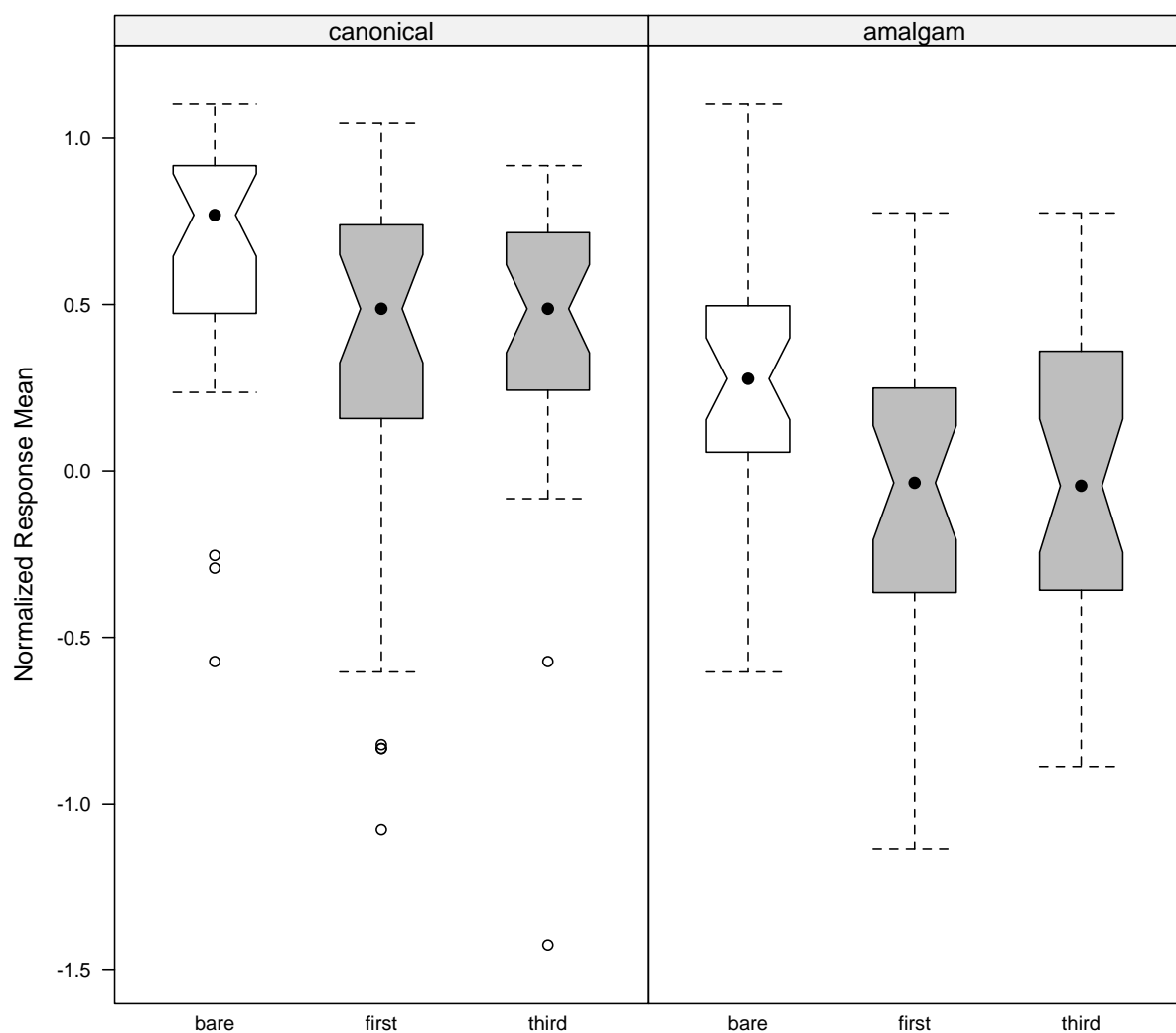
Figure 4.2: Effect of embedding in specificational pseudoclefts

(82)   a. ??She whispered that he liked coffee was what he liked.

         b. ??She whispered that what he liked was he liked coffee.

This effect is expected if the amalgam type involves left-peripheral clausal structure not present in

the complement of a non-bridge verb—the analysis I will develop in Chapter 5.

      To the extent that the sentences in (82) are acceptable, they strike me as instances of free

indirect discourse, where some of the indexical content of the reported utterance event is preserved;

compare (83), where the indexical *tomorrow* can be interpreted relative to the subject's *whisper-*

time, instead of the matrix speaker's time.

(83)   She whispered that she would leave tomorrow.

The free indirect discourse interpretation of (82) is that the reported speech act was an amalgam pseudocleft. A counterpart of (82) with a bridge verb does not require verbatim fidelity to the form of the reported speech act: a speaker can report someone else's utterance using an embedded amalgam, even if the original utterance was not an amalgam.

(84)   Original: 'He liked coffee.'

         Report: She said he liked coffee was what he liked.

(85)   Original (whispered): 'He liked coffee.'

         Report: #She whispered that he liked coffee was what he liked.

(86)   Original (whispered): 'He liked coffee was what he liked.'

         Report: ?She whispered that he liked coffee was what he liked.

The central contribution of the indexical content of the amalgam clause to its interpretation will be analyzed in Chapter 8.

4.4.2   Non-finite contexts

Embedding in non-finite contexts is another notorious locus of variation in canonical pseudoclefts. Under ECM/raising-to-object verbs, the value-initial pseudocleft is acceptable with or without an overt non-finite copula, while the weight-initial order requires the copula (Higgins 1979; Williams 1983; Heggie 1988; Moro 1997; den Dikken 2006, etc.).

(87)   a.   I consider John (**to be**) the best candidate.

    b.   I consider the best candidate *(**to be**) John.

Once again, the empirical status of the anchor-initial pseudocleft is unclear. Heggie (1988), for instance, reports that *wh*-initial pseudoclefts cannot be embedded, which lends support to her proposal that the fronted predicate occupies an A′-position that is incompatible with the selectional requirements of an ECM predicate. I illustrate the reported contrast with the now-familiar example:

(88)   a.   I consider important to himself (**to be**) what John is.

        b. *I consider what John is (**to be**) important to himself.

Many speakers do not find (88b) bad (see also Hankamer 1974; Ross 2000). There is an alternative explanation for the reduced acceptability of some such examples. In an informal survey, a naïve speaker spontaneously narrated her experience of encountering a garden path. To paraphrase her comments, she parsed *I consider what John is to be* as a constituent, expecting that the remainder of the sentence would provide a value for *what John is to be*. The structure she had in mind would generate the following sentence, which is obviously quite difficult to process:

(89)   I consider [[what John is to be] to be a gentleman].

When I altered my prosody to indicate a constituent boundary at *what John is*, her judgment improved.[15]

---

[15]While the reduced acceptability of canonical pseudoclefts in this environment is not fully explained, embedding under *consider*, *expect*, and *believe* requires a particular interpretation, which may clash with the intensionality of the fronted predicate in specificational sentences. The meaning of such sentences is that the subject believes *of a referential object x* that some property holds. This is a predicational sentence type, so even in the finite paraphrase, the specificational versions of these sentences are odd:

(i)   a.   *I believe of what John is that it is important to himself.
        b.   *I believe of what John ate that it is an apple.

For reasons of space, I do not discuss this issue in detail here, but I suspect that the unacceptability of embedding certain specificational sentences under these ECM verbs is semantic, not syntactic.

I presented the same speaker with amalgams in the ECM environment (90), and she quite spontaneously said, "Yuck!"

(90) a. *I consider John is important to himself **to be** what John is.

    b. *I consider what John is **to be** he is important to himself.

Compare also (91).

(91) a. *I expect what John ate **to be** John ate an apple.

    b. *I expect John ate an apple **to be** what he ate.

These examples are simply ill-formed.

I consulted other speakers who likewise gave categorical judgments of amalgams in ECM contexts, matching my own very strong judgments, so I did not test these in a controlled experiment.

The non-reversible amalgam sentence types demonstrate the same pattern.

(92) *I consider that's the thing **to be** we ran out of money.

(93) *I consider guess what we should do **to be** we should go to the library.

(94) *I consider John should give that a try **to be** he should try the new pizza place.

(95) *I consider my mistake {∅ / **to be**}, {∅ / **to be**} that I forgot my lunch.

Interestingly, it is possible to embed a *that's x* clause in an ECM context, in which case the amalgam copula remains bare, while the *that's x* clause is non-finite.

(96) a. I consider [that (**to be**) the problem] **is** we ran out of money.

    b. *I consider [that'**s** the problem] **to be** we ran out of money.

The fact that the *that's x* shows the non-finite structure associated with ECM, and not the amalgam copula, offers further evidence that the *that's x* string is a constituent distinct from the counterweight clause. In fact, the structure in (96a) is perhaps best characterized as a free-*be* amalgam, since its weight and counterweight constituents are as in (97).

(97)   [I consider that (**to be**) the problem] **is** / ****to be** [we ran out of money].

In so-called *with*-absolutes, which embed non-finite clauses including a participial form of the copula, amalgams are similarly unacceptable, in contrast to canonical pseudoclefts.[16]

(98)   a.   With John **being** the teacher, we should really try to be on time. (We know how much he values punctuality.)

       b.   With the teacher **being** John, we should really try to be on time.

(99)   a.   With Minnesota **being** where he lives, we should bring parkas.

       b.   With where he lives **being** Minnesota, we should bring parkas.

(100)   a.   *With he lives in Minnesota **being** where he lives, we should bring parkas.

        b.   *With where he lives **being** he lives in Minnesota, we should bring parkas.

---

[16]I observed one example of an amalgam embedded under a *with*-absolute that does seem to succeed in the stylized dialogue from the 2014 FX television series *Fargo*, adapted from the 1996 film by the Cohen Brothers. To illustrate the regional and stylistic features used by the writers, the context is included.

(i)   Uh, my neighbor, he does the watch. Said he saw this fella sitting in his car outside the apartment. May even have followed me here, I'm thinking. So the neighbor—Jewish fella—he gives the guy a warning, you know, a kinda "move along." Said the fella was real menacing in return. Had a police scanner, he said. Talked about how maybe he was gonna come back and kill the neighbor and his family.
       —Aw, jeez. Did you call it in?
       Well, see, I don't think they'd believe me is the thing. I mean, not **after, you know, I arrested the wrong fella, being what they think**. So there's that.

It is difficult to tell whether the bolded string is intended to constitute a completely integrated structure, since the style is somewhat fragmented and includes vigorous pro-drop and instances of reported speech, but perhaps for some speakers, amalgams can be embedded under *with*-absolutes. Note also the amalgam in the sentence preceding the bolded one. The construction is extremely prolific in the series' script.

This section has shown that while canonical pseudoclefts participate in a restricted set of A-and A′-movement operations, amalgams are completely frozen.

4.5   Baggage

The copula in pseudoclefts shows a general tendency to remain bare (Akmajian 1979; Higgins 1979; Declerck 1988). It typically takes the form *is* or *was*, marked only for [tense] and [3sg] agreement. Akmajian and Higgins observe that when a pseudocleft contains temporal, modal, or aspectual marking, for example, it tends to associate with the lexical verb in the weight clause, rather than with the copula, (101).[17] These authors also observe limited examples of functional material on the copula, which tends to be "harmonious" with marking on the lexical verb, (101b). They propose a generalization: the copula's form may either harmonize with that of the lexical verb, or it can take simple present. Their claim is that (101c) should be ungrammatical. (I leave off acceptability diacritics for the following examples, since their empirical status is at issue.)

(101)   a.   What he will want is a raise.

      b.   What he will want will be a raise.

      c.   What he wants will be a raise.

Other functional elements, like negation, certain adverbials, and plural are supposed to be incompatible with the copula.

(102)   a.   What he wanted wasn't a raise.

      b.   What he wanted probably is a raise.

      c.   What he wants are two new employees.

---

[17]I will use the term "baggage" throughout this dissertation as a shorthand for functional material in the verbal domain of the clause (temporal, modal, and aspectual auxiliaries and negation).

This section evaluates these empirical claims, comparing the canonical pseudocleft to the amalgam in formal experiments. In previous research on copular amalgams, the general observation is that the amalgam copula strongly resists functional material—it must occur in simple form (den Dikken et al. 2000; Ross 2000; Ross-Hagebaum 2004). For example, Ross-Hagebaum's (2004) corpus study of *that's x is y* finds only bare tokens of the copula, nearly all of which are in simple present form *is*, with the rest taking the form *was*. The experiments in the present section indicate a robust contrast between the canonical and amalgam environments: the canonical copula can combine with a range of baggage, while the amalgam copula must take simple form. The next chapter argues on the basis of these findings that the lower domains of clausal structure (V and T) are absent in amalgams.

4.5.1 Negation

Let us begin high in the T-domain, with negation. It has been claimed that the canonical pseudocleft does not allow sentential negation; instead, it only allows constituent or "contradiction" negation of the value (Higgins 1979; Declerck 1988; Mikkelsen 2005; see also discussion in den Dikken 2005b). That is, the specification relation itself cannot be negated.

(103)   What he likes is: not coffee, but tea.

In fact, syntactic negation on the copula, and not merely constituent negation, is well formed, as evidenced by the availability of tag questions.[18] (104b) shows that the same tag question is incompatible with constituent negation.

(104)   a.   What he likes isn't coffee, is it?

       b.   *What he likes isn't coffee, but tea, is it?

---

[18]I thank Bill Haddican for suggesting this diagnostic.

The negative judgments reported in the literature may stem from a pragmatic problem. Negating a specification relation requires a specific context. It is particularly appropriate when the focus is explicitly contrasted with the value that holds of the variable in the actual world.

(105)   a.   What he likes isn't coffee... it's tea.

          b.   The teacher isn't John... it's Mary.

This contrast interpretation may be felicitously expressed with constituent negation. Despite this pragmatic constraint, specificational pseudoclefts do in fact allow sentential negation. This claim is supported by the results of an acceptability survey.

4.5.1.1 Experiment 3 design

The survey tested effect of sentential negation in six copular sentence types. The goal of the experiment was to determine whether the copula shows the same level of tolerance for negation in canonical pseudoclefts and amalgams. Given the preliminary pattern discussed above, the prediction is that canonical specifications pseudoclefts admit sentential negation, while amalgams do not.

Six sentence types are compared: canonical pseudoclefts (both orders), amalgam pseudoclefts (both orders), and two versions of *that's x is y*. The first version is the familiar one, and the second includes sentential negation in the weight clause (*that's not x*), in order to control for the effect of semantics. It is plausible that negation on the amalgam copula could be rejected because it clashes with the affirmative *that's x* clause; the additional control rules out this possibility. *That's x is y* serves as a proxy for the non-reversible amalgam types in this survey. It is a practical test case, because of its fixed form and frequency in written registers. Sample items illustrate the factor design in Table 4.5.

	Bare	Negation
Canonical	What you will enjoy is the dressing.	What you will enjoy isn't the dressing.
Rev. Canonical	The dressing is what you will enjoy.	The dressing isn't what you will enjoy.
Amalgam	What you will enjoy is you will enjoy the dressing.	What you will enjoy isn't you will enjoy the dressing.
Rev. Amalgam	You will enjoy the dressing is what you will enjoy.	You will enjoy the dressing isn't what you will enjoy.
That's x is y	That's what you will enjoy is the dressing.	That's what you will enjoy isn't the dressing.
That's not x is y	That's not what you will enjoy is the dressing.	That's not what you will enjoy isn't the dressing.

Table 4.5: Factor design for negation experiment

Items from 48 different lexicalizations were distributed among 12 versions by Latin square. Each version included 48 experimental items, with 4 items representing each condition in Table 4.5, along with 48 filler items. All lexicalizations had salient specificational readings.

In some cases, it was possible to interpret the pseudoclefts as predicational copular sentences (i.e., with a referential weight clause and predicative value), but this reading is always pragmatically dispreferred. A surefire way of disambiguating sentences in favor of the specificational reading is to use sentences with connectivity effects, but this strategy is problematic. These sentences make the amalgam parse of the canonical string more salient, since connectivity effects are licensed straightforwardly by the presence of a full counterweight clause. Since the purpose of these experiments is to contrast amalgam and canonical pseudoclefts, the items rely primarily on pragmatic disambiguation cues, rather than syntactic ones.

50 speakers from a variety of regions in the United States participated in the survey. Because of a high rate of rejection of baseline amalgam sentences in this survey, 20 sets of results were excluded (see discussion below), leaving 30 sets of results.

4.5.1.2 Experiment 3 results

Results were normalized against the filler mean. Normalized values were compared to the acceptability threshold of z=0, and to each other. The analysis used a linear mixed effects model with item, lexicalization, and subject as random effects, and sentence type, negation, and the interaction of sentence type and negation as fixed effects. Raw results are depicted in Figure 4.3.
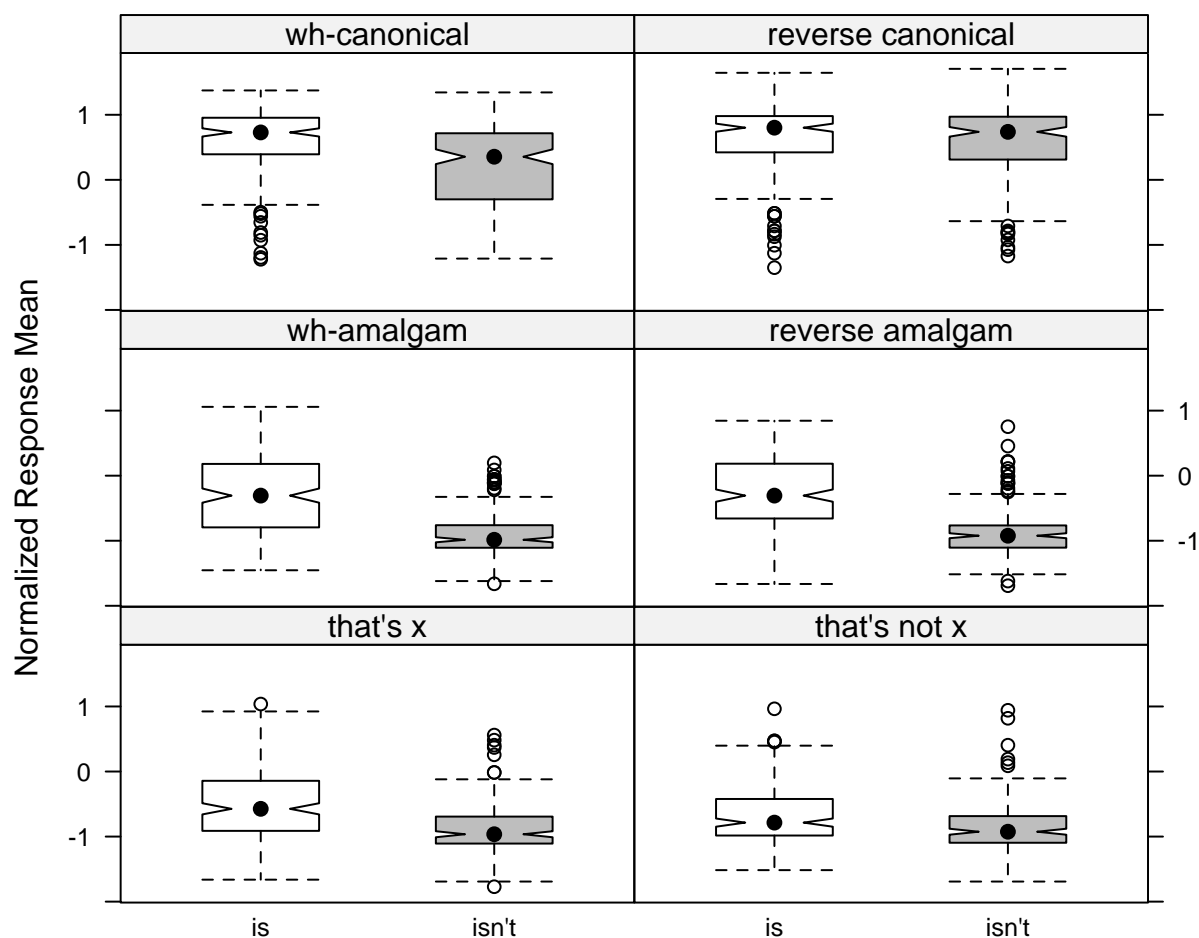
Figure 4.3: Negation in specificational sentences

Weight-initial and counterweight-initial pseudoclefts did not differ significantly from each other with respect to the effect of negation, while in the canonical pseudocleft sentence conditions, order and negation interact. More importantly, an ANOVA found that the interaction between sen-

tence type and negation was significant (p<0.001), when compared to the null model (Table 4.6).

	Df	AIC	BIC	logLik	deviance	Chisq	Chi	Df	Pr(>Chisq)
Null	10	3196.8	3254.6	−1588.4	3176.8				
Test	15	3112.2	3198.9	−1541.1	3082.2	94.569		5	<2.2e^{-16}∗∗∗

Table 4.6: Fixed effects linear model of interaction between negation and sentence type (2x6)

A similar analysis of the two *that's x is y* conditions (*that's x* and *that's not x*; see Figure 4.4) found a significant main effect of negation on the copula on both. There was also a significant interaction between negation and the form of the *that's x* clause; however, this interaction is not triggered by a main effect of negation, since the negated versions in both conditions are rated near the floor of acceptability. The interaction only reflects the difference between the two weight forms in the *bare* copula condition. This finding highlights the fact that additional negation inside one of the major constituents of the amalgam does not facilitate negation on the copula; i.e., *that's x is y* does not exhibit harmony or some other matching process for negation.

Now consider the effect of negation in canonical pseudoclefts. The amalgam conditions (see Figure 4.3) all show a floor effect: many participants simply did not accept the amalgam sentence types. This is not surprising, because they are prescriptively stigmatized, and not often represented in writing (see discussion in Chapter 3).

Preliminary analysis of the raw data finds a significant interaction between sentence type and negation. This finding should be validated for speakers who find the baseline amalgams acceptable. Setting a threshold of z=−0.25 (below the normalized baseline), I excluded results from subjects who rated *both* amalgam pseudocleft sentence types in the bare copula condition below that threshold. This identifies a subset within the participant population of people who find bare amalgam pseudoclefts at least somewhat acceptable. Figure 4.5 shows the by-subjects threshold. On the basis of this analysis, 20 sets of results were excluded, leaving 30.
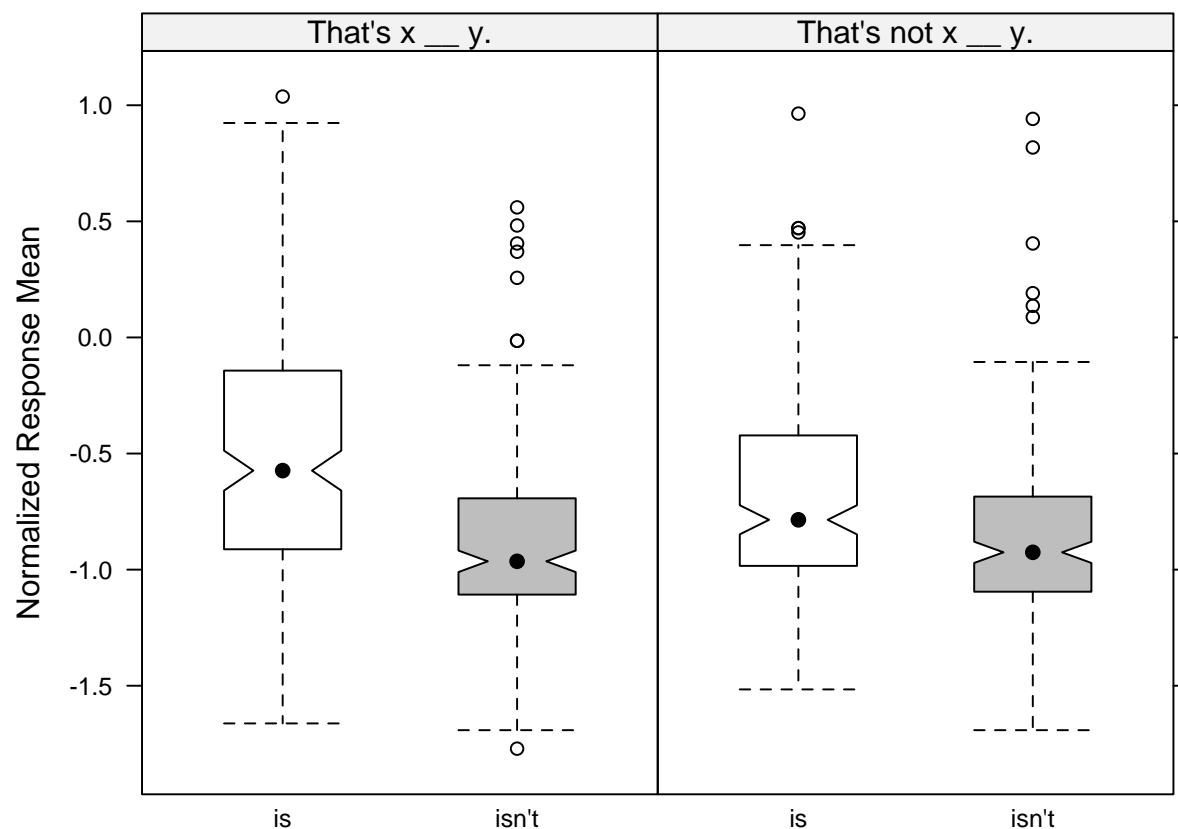
Figure 4.4: Floor effect of negation in *that's x is y*

The overall pattern in the results is the same (compare Figure 4.3 and Figure 4.6).[19] Refitting

the mixed effects models to the smaller dataset yields very similar results to the models of the larger

dataset. Table 4.7 shows that the interaction between sentence type and negation is significant

(compare to the strikingly similar results in Table 4.6).

Model	Df	AIC	BIC	logLik	deviance	Chisq	Chi	Df	Pr(>Chisq)
Null	10	1922.7	1975.4	-951.35	1902.7				
Test	15	1811.4	1890.5	-890.68	1781.4	121.34		5	$<2.2\mathrm{e}^{-16}$∗∗∗

Table 4.7: Fixed effects linear model of interaction between negation and sentence type (n=30)

---

[19]The fact that the pattern is preserved suggests two hypotheses: English grammars all include copular amalgams, but some speakers do not find them acceptable; or English speakers can perform reliable judgments of amalgams even if their own dialect lacks them. The idea that some grammatical features are not realized in performance, whether because of prescriptive stigma or other reasons, is developed in, e.g., (Barbiers 2005) and (Blanchette 2013).
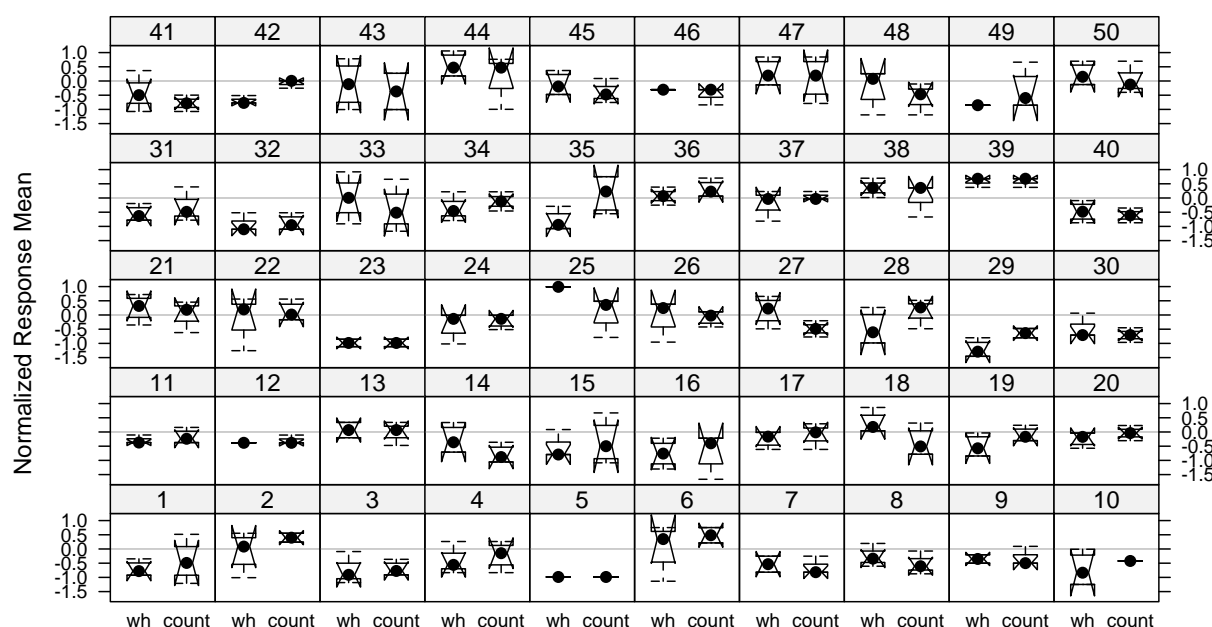
Figure 4.5: By-subjects analysis of baseline; *wh*-initial vs. counterweight-initial

To ensure that the interaction is not solely due to the low acceptability of *that's x is y* among these speakers, I repeated the analysis on a subset of the data. This subset included only the four pseudocleft conditions—once again, the interaction between sentence type and negation is significant (p=2.953e^{-16}). In addition, collapsing the four-level sentence type factor into two levels: canonical and amalgam, shows a similarly strong interaction (Table 4.8). This additional comparison highlights the effect of the canonical vs. amalgam sentence type factor.

Model	Df	AIC	BIC	logLik	deviance	Chisq	Chi	Df	Pr(>Chisq)
Null	7	2294.0	2331.6	-1140.0	2280.0				
Test	8	2249.2	2292.2	-1116.6	2233.2	46.762		1	8.015e^{-12}***

Table 4.8: Fixed effects linear model of interaction between negation and sentence type (2x2)

In Figure 4.6, it is clear that the reverse (value-initial) pseudocleft shows almost no effect of negation (as expected), while negation does reduce the acceptability of the weight-initial pseudocleft. Although the negated weight-initial pseudocleft has lower acceptability than the bare (affir-
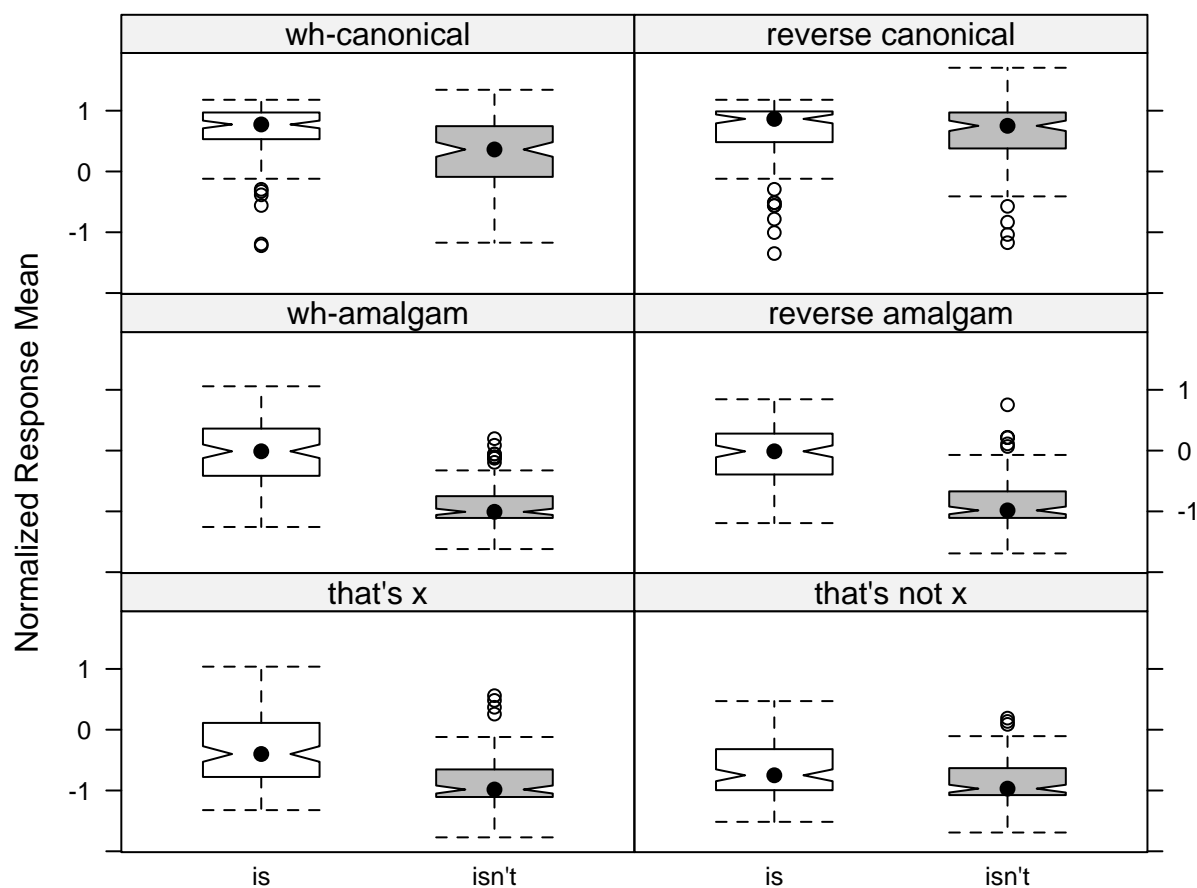
Figure 4.6: Negation in specificational copular sentences (n=30)

mative) version, it remains well above the normalized mean across speakers. That is, canonical

pseudoclefts with negation are not judged as absolutely unacceptable; they are merely judged as

less acceptable than pseudoclefts without negation, when no context is given.

What remains is to compare the canonical *wh*-initial pseudocleft to its amalgam counterpart:

does negation have a stronger effect in the amalgam condition? An analysis of the subset of data

including only the *wh*-initial pseudoclefts finds that the answer is affirmative (p<0.01; t-value

for the canonical-negation interaction term=3.086). The significant result in the other models was

therefore not due only to the reverse pseudocleft or *that's x is y* condition: the asymmetry between

canonical and amalgam copular sentences holds across the board.

The unavailability of negation in the amalgam sentences reinforces the argument from Chapter 2 that amalgams are not unintegrated biclausal structures, for example, with pro-drop in a second clause. If the amalgam copula were the main verb of a second clause with pro-drop, then negation should be required when the precopular clause is negated. Otherwise, a contradiction would ensue:

(106)   a.   *That$_i$'s not my problem; it$_i$/my problem$_i$ is [that it's ugly]$_i$.

         b.   That$_i$'s not my problem; it$_i$/my problem$_i$ isn't [that it's ugly]$_i$.

The counterpart of (106a) is the only grammatical option; the integrated *that's x is y* amalgam counterpart of (106b) is totally ungrammatical. The amalgam copula is not the matrix verbal element of an unintegrated sentence subsequent to the precopular constituent.

4.5.2   Plural

The distribution of plural agreement in specificational copular sentences, pseudoclefts in particular, is another empirically hazy area. Den Dikken (2005b) reviews the discussion in the literature, and I will briefly comment on his findings here. Specificational sentences (in weight-initial order) favor the singular form of the copula.

(107)   a.   What you have bought **is** fake jewels.

         b.   What you have bought **are** fake jewels.   (Declerck 1988:79)

Declerck's (1988) observation is that the plural agreement requires the predicational reading. The paraphrase of the plural sentence is: *The things that you have bought have the property of being fake jewels.* Nevertheless, plural agreement in sentences that are most naturally construed as specificational strikes me as extremely natural, and I hear such examples frequently. Even when the

weight is formally singular, the plural copula is possible.

(108)   The thing that bothers me most **are** the calories.

In a sentence like (108), the weight in the structural subject position is singular, so there must be a mechanism in English for agreeing with a postcopular subject of predicate inversion (see Heycock 2012 for further discussion).

As for amalgams, den Dikken (2005b:29) reports that Chris Wilder in a personal communication finds reduced weight-initial amalgam pseudoclefts ungrammatical with the plural copula, e.g., (109).

(109)   *What nobody brought **were** any crackers.

In den Dikken et al.'s (2000) analysis of pseudoclefts with NPI connectivity, (109) is analyzed as a partially elliptical amalgam pseudocleft. NPIs must be licensed at spell-out, so there must be a local licensor for *any*—this licensor is the elliptical clause-mate *nobody*. (LF approaches to connectivity cannot account for NPI connectivity.)

Once again, my consultants give a different result.

(110)   a.   What nobody brought **were** any crackers.

         b.   That's what nobody brought **were** (any) crackers.

         c.   You know what nobody brought, **were** (any) crackers.

Such examples are judged in informal contexts as quite natural—some speakers, including me, prefer the singular, but accept the plural. I prefer the singular even when the local value is plural.

(111)   That's all we are, **is** information brokers.[20]

---

[20]http://www.jobshadow.com/interview-with-a-private-investigatorfirm-owner/

The analysis of this sentence type must therefore accommodate both the singular and the plural forms. In the fully overt counterpart of (111) (and its counterweight-initial version), the copula tends to remain singular, since the formally and notionally plural focus finds itself embedded in the counterweight clause.

(112)   a.   What nobody brought **was**/??**were** nobody brought any crackers.

        b.   Nobody brought any crackers **was**/??**were** what nobody brought.

For speakers who accept the plural agreement version, the plural seems to be coming either from the underspecified *what* or the embedded value, which may be accessible for agreement attraction due to its status as a narrow focus (den Dikken 2014a). When the weight is a plural DP, by contrast, it can easily license plural on the amalgam copula, regardless of order.

(113)   a.   **The main problems are** he missed the shot and he argued with the referee.

        b.   He missed the shot and he argued with the referee **are the main problems**.

Examples like (113) raise a question: is plural in specificational sentences triggered by the value/counterweight or by the weight? The question is an important one, because the presence of [ɸ]-features is central to structural subject-hood, and one of the goals of this thesis (and the main topic of Chapter 7) is to account for the occurrence of bare sentential subjects in amalgams, in contrast to canonicals. To address this question, let us focus on specificational sentences with sentential values, like (113). This sentence type allows for a direct comparison of canonical and amalgam specificational copular sentences, since they both have sentential logical subjects.

Sentential subjects in canonical contexts are associated with singular/default agreement:

(114)   That Mary should be so disappointed surprise*(s) me.

In a DP-weight specificational sentence, therefore, canonical or amalgam, the copula takes singular form if the weight is singular.

(115)   a.   That he missed the shot is the problem.

           b.   The problem is that he missed the shot.

(116)   a.   He missed the shot is the problem.

           b.   The problem is he missed the shot.

If a conjunction of propositions serves as a subject (in a non-specificational context), however, both singular and plural agreement are possible (see also McCloskey 1991).

(117)   That it's raining and that I live far from the train **is** making me mad.

(118)   a.   That he will leave and that he will stay both **terrify** me.

           b.   That he will leave and that he will stay **are** equally possible.

In non-specificational contexts, like (118), there is no identity relation between the subject and predicate (and the predicate is formally non-nominal), so it is clear that the plurality is coming from the subject.

In specificational sentences, the case is not so clear. If the conjunction of propositions is notionally plural, then the anchor is notionally plural—it is then odd for the weight to be formally singular. (119) means that the value of *my worry* is a plurality of propositions. It is much more natural in this case to use the plural weight *my worries* (120).

(119)   ??That he will leave and that he will stay **are** my worry.

(120)   That he will leave and that he will stay **are** my worries.

Since wherever plural on the copula is licensed, the weight becomes formally plural, it is difficult to determine whether a plural copula is valued by the subject or the predicate.

One way to lessen the effect of this confound is to use a weight that is formally underspecified for number: one introduced by a proposition-denoting *wh*-word. The same weight form is compatible with both singular and plural contexts. We might expect, therefore, that it will simply match the notional number of the value/counterweight. If a value/subject consisting of a conjunction of *that*-CPs can be formally plural, then the copula will be compatible with plural. If the conjunction of CPs cannot be plural, then the copula will tend to remain singular.

A difference in the [ɸ]-feature content of the value/counterweight clause in the canonical and amalgam environments raises central questions about the nature of the "structural subject" relation in amalgams, and the role of the overt complementizer in the structural subject relation. If the nominality of the *that*-CP that allows it to serve as an ordinary subject is tied to the presence of [ɸ]-features, then an asymmetry between the canonical and amalgam specificational sentence types is predicted. Coordinated CPs in the canonical sentence can trigger plural, while coordinated bare sentential subjects cannot.

4.5.2.1 Experiment 4 Design

The data needed to test this prediction are fraught with confounds, and the judgments are rather delicate, so I conducted a controlled survey. The survey examined the effect of the plural copula on the acceptability of sentential-value specificational sentences with a conjunction of propositions in subject position. To mitigate the effect of the DP-weight's number feature, formally plural DP weights were also compared to underspecified *wh*-weights. Since it is impossible to completely remove "interference" from the number feature of the weight (because the underspecified *wh*-weights may behave like plurals as a result of the specification relation), this survey is exploratory.

The survey had a 2x2x2 design, comparing sentence type, weight number, and copula form (Table 4.9). Sentence type, in this survey, includes only value/counterweight-initial sentences, and distinguishes between canonical ([*that*-CP and that-CP]) and amalgam ([S and S]). Weight form includes formally underspecified free relatives and formally plural DPs. Eight versions were designed (32 lexicalizations, 32 items per version, 4 from each condition). 32 participants recruited from AMT completed the survey.

Type	Weight	Bare	Plural
Canonical	*Wh-*	That he took her purse and he knocked her down is what I saw.	That he took her purse and he knocked her down are what I saw.
	DP_{pl}	That he took her purse and that he knocked her down is the worst parts.	That he took her purse and that he knocked her down are the worst parts.
Amalgam	*Wh-*	He took her purse and he knocked her down is what I saw.	He took her purse and he knocked her down are what I saw.
	DP_{pl}	He took her purse and he knocked her down is the worst parts.	He took her purse and he knocked her down are the worst parts.

Table 4.9: Factor design for plural experiment

4.5.2.2 Experiment 4 results

Raw results of the survey are depicted in Figure 4.7. Outlier analysis excluded two participants and one lexicalization. The plurality of the weight expression was a highly signifiant predictor of the acceptability of the plural form of the copula, which is expected (Table 4.10).

Model	Df	AIC	BIC	logLik	deviance	Chisq	Chi	Df	Pr($>$Chisq)
Null	6	1425.8	1454.8	-706.91	1413.8				
Test	7	1376.6	1410.4	-681.28	1362.6	51.245		1	$8.152e^{-13}$***

Table 4.10: Main effect of plural weight on acceptability of plural copula

Less clear is whether the effect of the form of the sentential value can be distinguished from that of the weight. An ANOVA of a model of the three-way interaction between amalgam vs. canonical sentence type, underspecified vs. plural weight, and bare vs. plural copula, compared with a
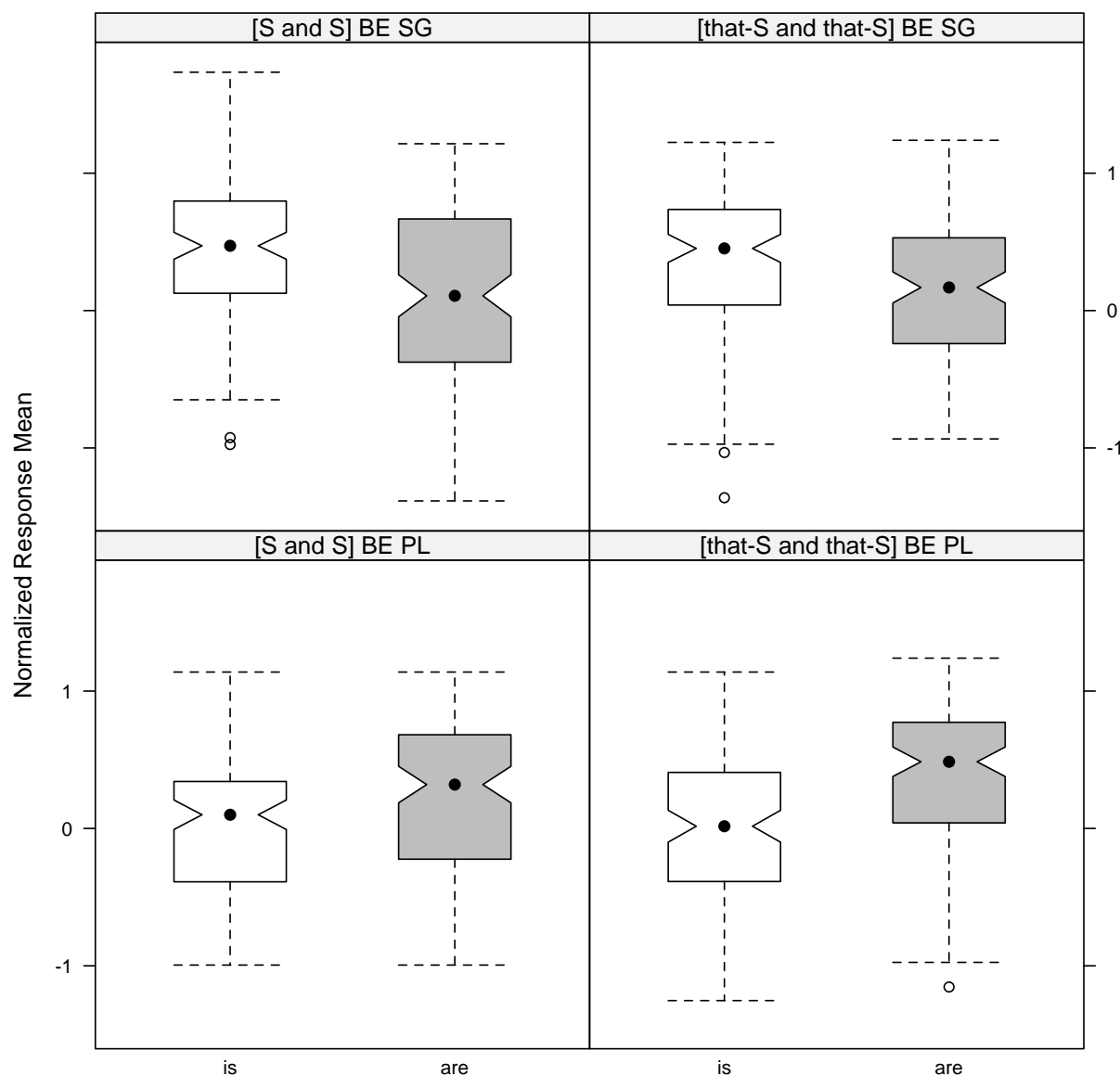
Figure 4.7: Effect of plural on the copula in DP-weight amalgams

model with only a two-way interaction between weight form and copula form (since weight form is known to be a strong main effect), does not reach significance (p=0.077). Amalgam vs. canonical sentence type (coordinated S or coordinated CP) made some difference, however. When the weight form factor is not considered, i.e., when all amalgams are compared to all canonicals, regardless of weight form, an interaction between sentence type and copula form approaches significance,

when compared to a model without the interaction term (p=0.05054). The amalgam copula shows a stronger preference to remain singular in plural environments than its canonical counterpart. Figure 4.8 shows the two-way comparison.
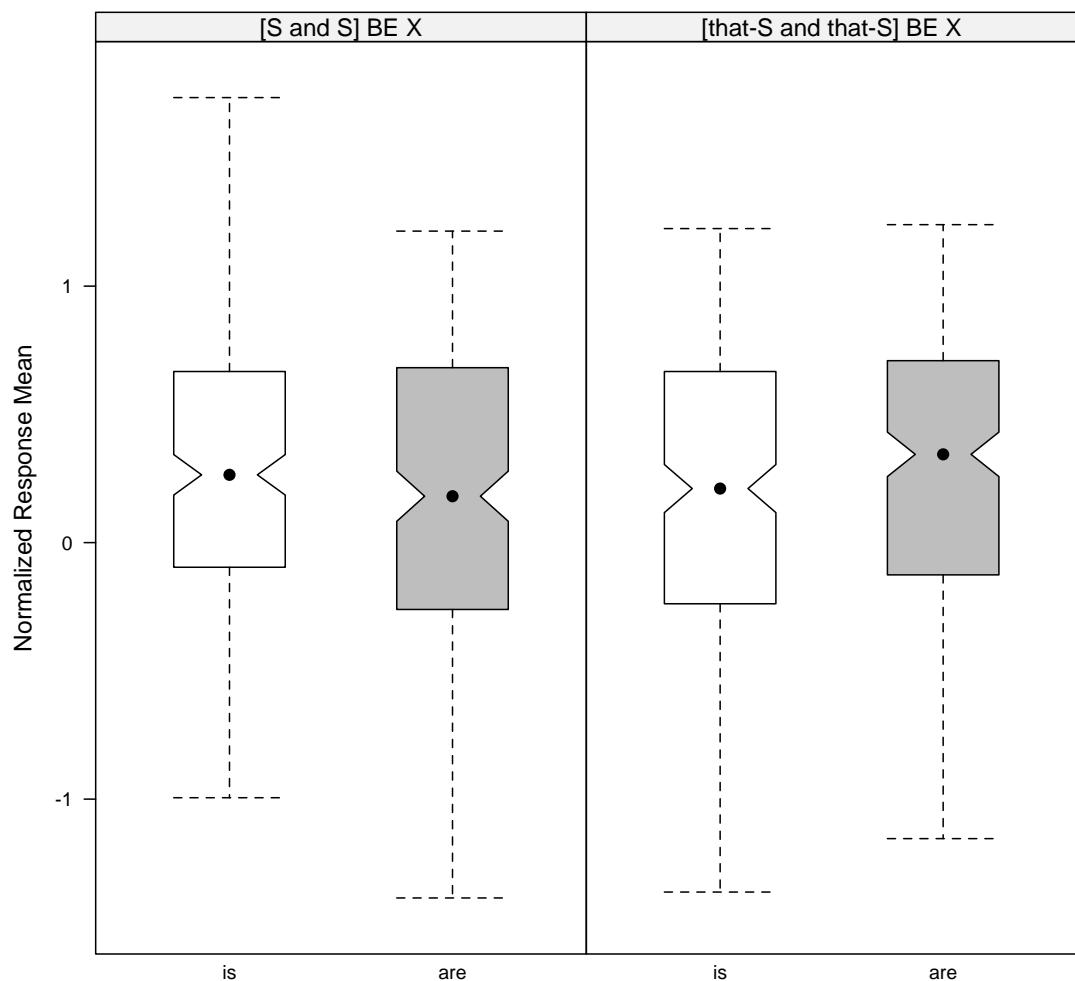
Figure 4.8: Effect of plural on the copula in amalgam and canonical specificational sentences

If we look within the condition where the weight is unspecified for number (thus, not a source for plural [ɸ]), and compare the canonical condition to the amalgam condition, there is no significant difference, which does not support the initial prediction. I suspect that this result is due to an additional factor that this experiment did not control for. There is a strong preference for interpreting coordinated CPs as notionally singular, particularly when they describe sub-events of

a single event or context (like the examples in Table 4.9). This sub-event interpretation arises quite naturally in specificational sentences, where the two clauses jointly specify the value of a variable.

Consider now an additional example where this confound is removed. Prosody and pragmatics can facilitate the construal of the coordinated CPs as two separate objects, rather than a notionally singular object. Imagine that the following dialogues take place in a context where a basketball player seems inordinately upset after missing a shot.

(121)   A: Is the reason he's upset just that he missed the shot?

B: No, that he is in trouble with the COACH **AND** that he has a KNEE injury **is/are** why he's upset.

(122)   A: Is the reason he's upset just that he missed the shot?

B: No, he is in trouble with the COACH **AND** he has a KNEE injury **is/*are** why he's upset.

In (122) and (121), the asymmetry between the canonical and amalgam number agreement patterns is more robust. I verified these judgements in informal surveys. Ideally, a larger-scale study would use pragmatic and prosodic cues to support the notionally plural reading, to replicate this finding under controlled circumstances.

Since the informal judgment data are stable across speakers, I conclude that the original prediction is in fact correct: canonical CP values have [ϕ]-features, in contrast to amalgam counterweights. In Chapter 7, this contrast is attributed to a formal difference between the two sentential subject types: *that*-headed CPs are formally nominal, while root CPs are not.

4.5.3 Simple tense

Tense in specificational copular sentences presents some analytical puzzles, which have never been solved conclusively. I will propose a new analysis in Chapter 8, which is devoted to examining the meaning of tense and anchoring to the utterance context in canonical and amalgam pseudoclefts. This section offers a summary of the simple tense form combinations that are available in pseudoclefts.

Canonical and amalgam pseudoclefts show the same range of simple tense combinations, summarized in Table 4.11.

(123)   a.   What he wants is another drink.

    b.   What he wanted is another drink.

    c.   What he wanted was another drink.

    d.   *What he wants was another drink.

(124)   a.   Another drink is what he wants.

    b.   Another drink is what he wanted.

    c.   Another drink was what he wanted.

    d.   *Another drink was what he wants.

(125)   a.   What he wants is he wants another drink.

    b.   What he wanted is he wanted another drink.

    c.   What he wanted was he wanted another drink.

    d.   *What he wants was he wants another drink.

(126)   a.   He wants another drink is what he wants.

    b.   He wanted another drink is what he wanted.

    c.   He wanted another drink was what he wanted.

    d. *He wants another drink was what he wanted.

	[pres[pres]]	[pres[past]]	[past[past]]	[past[pres]]
wh-canonical	✓	✓	✓	✗
rev. canonical	✓	✓	✓	✗
wh-amalgam	✓	✓	✓	✗
rev. amalgam	✓	✓	✓	✗

Table 4.11: Simple tense patterns in pseudoclefts

The literature on tense patterns in specificational contexts emphasizes the tendency for the copula's tense to be "congruent" with that of the weight term, or to remain in simple form. Higgins (1979:294) claims that "the tense of the copula in a specificational pseudo-cleft sentence is not an independent variable". Declerck (1988) proposes that the tense meaning (and form) of the copula is "neutralized" in the environment of the tensed lexical verb. These requirements are analyzed in both morphological and semantic terms. In Higgins (1979), Akmajian (1979), and Declerck (1988), these different components of tense are not effectively teased apart, as I will argue later in this thesis. Sharvit (2003) and Romero (2004) give formal semantic analyses of the matching pattern (which can also accommodate the optional simple present on the copula), but I will show in Chapter 8 that their analyses under-generate and do not account for the difference between canonical and amalgam pseudoclefts.

4.5.4 Future *will*

Consider now the distribution of the temporal auxiliary *will* in pseudoclefts. Akmajian (1979) observes when the specificational copula is marked by *will*, the predicational interpretation becomes salient:

(127)   What you are holding in your hand **will be** a small brown butterfly.

'The thing you are now holding in your hand will in the future be a small brown butterfly.'

Sharvit (2003) reports that in fact the modal *will* is incompatible with specificational copula for semantic reasons: it forces a contradictory meaning onto the weight's Topic Time. This empirical claim is not accurate, however. Declerck (1988) and Ross (2000) observe that when the future also occurs on the weight clause, the specificational interpretation is readily available. (128) gives a specificational counterpart of (127).

(128)   What you *will* hold in your hand **will be** a small brown butterfly.

Hedberg (2008) gives other examples, cited from written corpora.

(129)   a.   But what really may be at issue when this comes to term **will be** deployment of SDI.

       b.   What Soviets will gain from reading Dr. Zhivago **will be** a somewhat better under-standing of their recent history.

In these examples, the specification relation itself is clearly future-shifted.

An important question is whether the copula can be marked for future when the weight clause is not; that is, is future on the copula only possible when it is "congruent" with the lexical verb? Another question is: do canonical and amalgam pseudoclefts behave the same way with respect to *will*?

4.5.4.1   Experiment 5 design

Given the variability in reported judgments, and the straightforward structure of the data in question, I tested the effect of future marking experimentally. Once again, an asymmetry between the canonical and amalgam sentence types is predicted. Only the canonical specificational pseudocleft is expected to allow the future auxiliary in combination with the copula.

The experiment used a 4x3 design, crossing sentence type with future marking. The sentence type factor had four levels: canonical pseudocleft, reverse canonical pseudocleft, amalgam pseudocleft, and reverse amalgam pseudocleft. The future-marking factor had three levels: bare (simple copula), future on the copula alone, and future on both the copula and the lexical verb. The latter two levels were differentiated in order to test whether the "harmonious" configuration, with future on both the copula and the lexical verb, was indeed the only grammatical option.

The experiment's 12 conditions were represented in 48 lexicalizations, distributed among 12 versions by Latin square. The versions were constructed using the same blocking and randomization procedure that was used for the other experiments reported in this chapter. 48 participants completed the survey. The design is summarized in Table 4.12.

	Bare	[will be [pres]]	[will be [will V]]
wh-canonical	What they'll need is a good rest.	What they need will be a good rest.	What they'll need will be a good rest.
rev. canonical	A good rest is what they'll need.	A good rest will be what they need.	A good rest will be what they'll need.
wh-amalgam	What they'll need is they'll need a good rest.	What they need will be they need a good rest.	What they'll need will be they'll need a good rest.
rev-amalgam	They'll need a good rest is what they'll need.	They need a good rest will be what they need.	They'll need a good rest will be what they'll need.

Table 4.12: Factor design for future experiment

4.5.4.2 Experiment 5 results

Normalized results were inspected for outliers. The participant pool for the future experiment included 17 participants who rated baseline amalgams below a floor of $z=-0.5$, so their results were excluded, following the same reasoning as the outlier analysis in the negation experiment, leaving 31 sets of results. The results are summarized in Figure 4.9.

The summary shows that future marking on the copula reduces acceptability in all conditions,
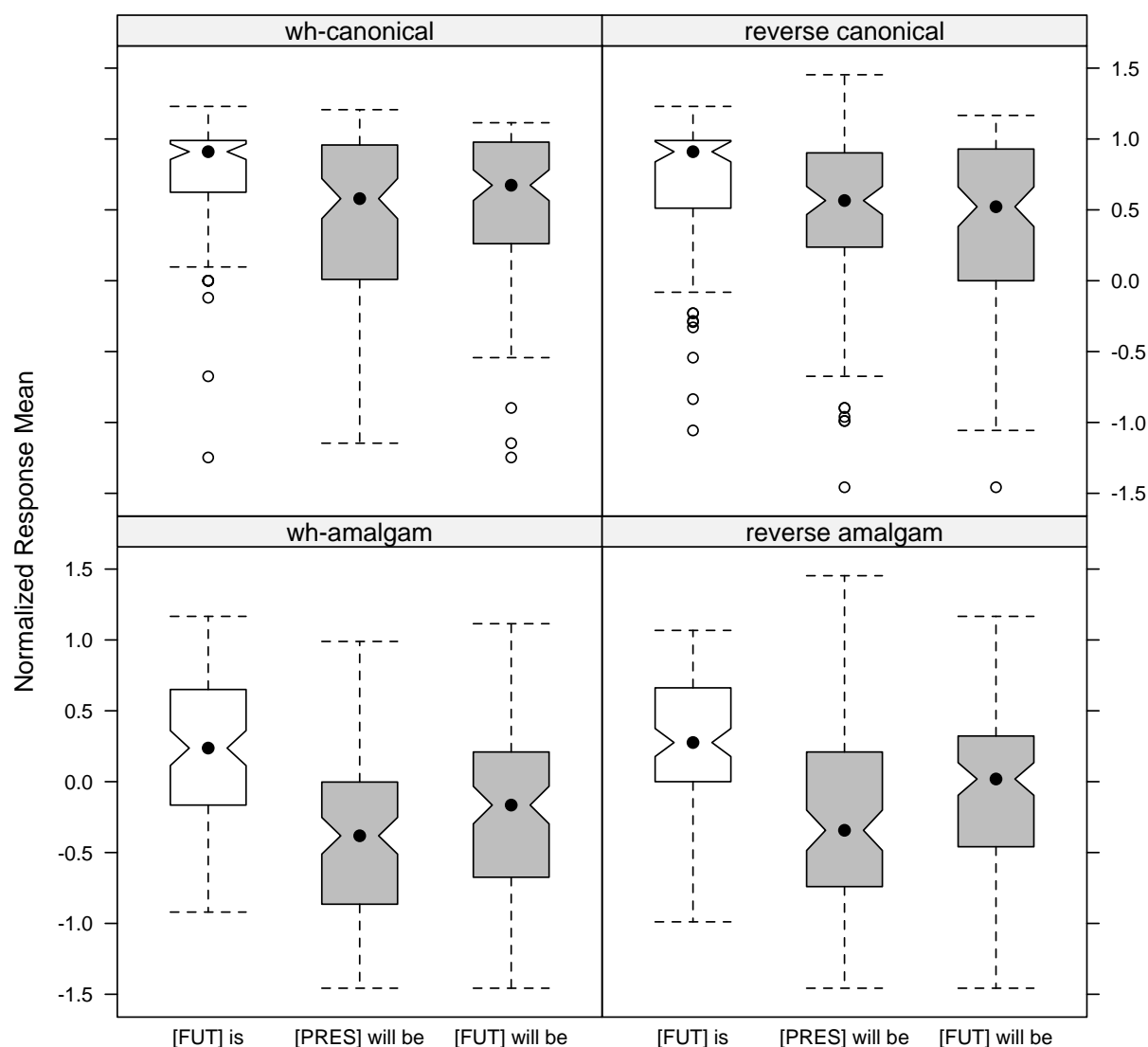
Figure 4.9: Effect of future copula on specificational pseudoclefts (n=31)

but the effect appears stronger in the amalgam conditions.[21] Contrary to the claims in the litera-

ture, future marking on the copula alone is judged as acceptable in canonical pseudoclefts: over

---

[21] An interesting and rather surprising result in these data is the high number of low-score outliers in the bare reverse canonical pseudocleft condition, e.g., *A good rest is what they'll need*. While I do not have a conclusive account of this result, I speculate that some speakers were primed by the frequency of the amalgam structure in the survey. A reverse canonical pseudocleft string can be derived from an amalgam structure via backward ellipsis—perhaps some speakers rated these baseline examples lower than other participants, because they had to recover the elided material. Ad hoc inspection of the results offers no conclusive evidence for this hypothesis, but there was one participant who rated the bare reverse canonical pseudoclefts lower than the reverse amalgam pseudoclefts, which (s)he accepted; this accounts for some of the outlier items.

three-quarters of the responses are distributed in the positive acceptability range. In the amalgam condition, however, well over half of the responses are in the negative range.

A linear mixed effects model including the interaction between sentence type and copula form is significant with respect to the null model (Table 4.13). An additional test of the two-way interaction between bare and non-bare forms of the copula (independent of the form of the lexical verb) also yields a significant result when compared with the null model ($p<0.01$). Lastly, two tests of the interaction between copula vs. amalgam sentence type (without distinguishing for order) and future (one with all three levels distinguished: bare copula, independent future copula, future copula + future V; the other with only two levels distinguished: bare copula vs. future copula) are both significant ($p<0.001$). The simpler two-way interaction is visualized in Figure 4.10.

Model	Df	AIC	BIC	logLik	deviance	Chisq	Chi	Df	Pr($>$Chisq)
Null	9	1869.3	1916.1	-925.67	1851.3				
Test	15	1862.9	1940.8	-916.43	1832.9	18.467		6	0.005165**

Table 4.13: Effect of future marking on specificational pseudoclefts

The high error rate in this data set warrants further discussion. Notice the wide distribution of responses to the future form copula condition in Figures 4.9 and 4.10. I speculate that the skew is due to a confounding factor. There is a salient epistemic use of the modal *will* (e.g., (130)), which may make future *marking* more acceptable in environments where future *semantics* is not.

(130)   A: Who's that at the door?

         B: Oh, that **will be** John.

This speculation is corroborated by informal survey findings that disambiguate the meaning using context and temporal adverbials. In these contexts, informants judged the canonical sentences as acceptable, and the amalgams as unacceptable. The informal judgments of the following contrasts are extremely robust, matching my own intuition.
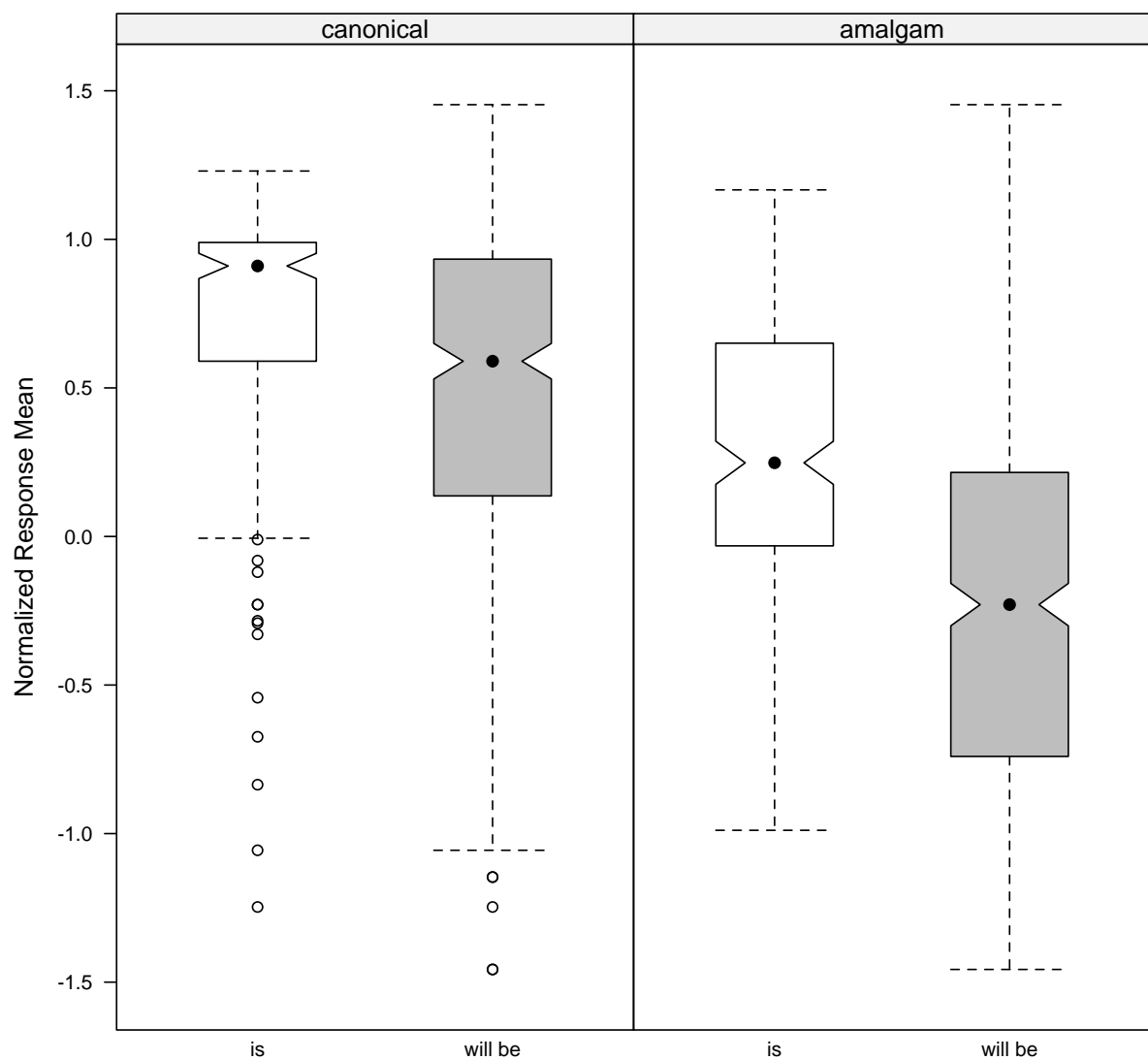
Figure 4.10: *Is* vs. *will be* on the canonical vs. amalgam copula

(131)   a.   John has been drinking too much lately. If he keeps this up, soon what he needs **will be** rehab.

        b.   *John has been drinking too much lately, If he keeps this up, soon what he needs **will be** he needs rehab.

(132)   a.   John has been drinking too much lately. He thinks he needs to cut back a little, but

I think that if he keeps this up, then rehab **will be** what he needs.

b. *John has been drinking too much lately. He thinks he needs to cut back a little, but

I think that if he keeps this up, then he needs rehab **will be** what he needs.

Epistemic modals contrast with deontic modals not only in meaning, but also in their syntactic position (e.g., Condoravdi 2002; Cinque 2004; Hacquard 2009, 2010). The model that I will develop in Chapter 5 predicts the (limited) availability of epistemic *will* in amalgams.

To summarize, participants in formal and informal surveys find future marking on the amalgam copula to be unacceptable, in contrast to the canonical copula. Contradicting previous literature, my results show that even when the lexical verb is in present tense form, the copula in a specificational context may be marked for future. On the semantics side, informal surveys find that under a *bona fide* future interpretation, the contrast between canonical and amalgam sentences with a future copula is categorical.

4.5.5 Modals

Now let us turn to the distribution of other modals in specificational sentences. The "congruence" requirement on tenses is thought to hold of modals as well; that is, modal auxiliaries on the specificational copula require the presence of matching modal auxiliaries on the lexical verb. (Diacritics in (133) are based on the general claims in Higgins 1979; Akmajian 1979; Declerck 1988; see further discussion in den Dikken 2005b.)

(133)   a.   What he **might be** reading **might be** *War and Peace*.

b.   What he **might be** reading **is** *War and Peace*.

c.   *What he **is** reading **might be** *War and Peace*.

Casual observation shows that this generalization is too strong (see also Ross 2000). Speakers frequently produce structures like (134).[22]

(134)   a.   What he needs **might be** a vacation.

         b.   What she wanted **must have been** the vegan option.

         c.   What they do **should be** to call back later.

Both deontic and epistemic uses of English modals are compatible with the canonical pseudocleft. The two readings are distinguished by context, since English modal auxiliaries are ambiguous. Consider the deontic examples in (135).

(135)   a.   If he wants to keep fans interested, what he writes about next **must be** himself.

         b.   To win the race, what she chose **had to be** to train every day.

Since the obligation conveyed by the deontic modal is most naturally ascribed to the referential subject of the *wh*-clause, the epistemic reading of the modal on the copula is the default. Sentences like those in (136) are perfectly acceptable, interpreted with speaker-oriented epistemic modality.

(136)   a.   What John is **must be** important to himself.

         b.   What she chose **had to be** to train every day.

         c.   What they needed **might have been** a raise.

---

[22]The example in (134c) is less acceptable without *to*. My intuition that the absence of *to* favors the amalgam parse, where the deontic modal is not acceptable. Perhaps because amalgam pseudoclefts introduced by a *do*-clause are so frequent, the amalgam parse is the default for pseudoclefts with VP values. If the preference for omitting *to* whenever possible is robust across speakers, it would suggest that the amalgam structure is uniformly grammatical, and that the higher frequency of amalgams with *do* mitigates whatever extralinguistic factors lead many speakers to reject them in judgment tasks. I was not able to conduct an experiment on the interaction between the presence of *to* and canonical vs. amalgam underlier (e.g., using disambiguating baggage on the copula) for this dissertation, but it is a potentially revealing question I leave for future research, which gets at the heart of the competence vs. performance problems that copular amalgams confront.

I did not undertake a formal experiment to test the acceptability status of modals, because the context needed to distinguish deontic and epistemic modals would render the amalgam items quite long, and difficult to judge in writing. Informal surveys show an asymmetry between the deontic and epistemic interpretations. Speakers categorically rejected amalgams in the deontic contexts, as in (137) and (138).

(137)   a.   *If he wants to keep fans interested, what he writes about next **must be** he writes about himself.

      b.   *To win the race, what she chose **had to be** she chose to train every day.

(138)   a.   *If he wants to keep fans interested, he writes about himself **must be** what he writes about next.

      b.   *To win the race, she chose to train every day **had to be** what she chose.

Amalgams were judged as at least marginal when the modals were interpreted epistemically. I leave off judgment diacritics for the following examples, but in traditional notation, some speakers would give them '?' status.

(139)   a.   What John is **must be** he's important to himself.

      b.   What she chose **had to be** he chose to train every day.

      c.   What they needed **might have been** they needed a raise.

(140)   a.   John is important to himself **must be** what he is.

      b.   She chose to train every day **had to be** what she chose.

      c.   They needed a raise **might have been** what they needed.

Similarly, epistemic but not deontic adverbs are accepted in amalgams.

(141) a.   It's a working class language, **probably**, is what it amounts to.[23]

     b.   What it amounts to **probably** is it's a working class language.

(142) a.   *What they did was **obligatorily** they used the service entrance.

     b.   *They used the service entrance was **legitimately** what they did.[24]

Higgins (1979), Bošković (1997a), and others claim that adverbs cannot be placed to the right of the copula in *wh*-initial pseudoclefts; however, den Dikken (2005b) points out examples in Huddleston (1971) and Declerck (1988) that are judged as acceptable. My informants preferred the adverbs to the right of the copula, but accepted both orders.

(143) a.   What he is is **probably** he's angry with himself.

     b.   He's angry with himself is **probably** what he is.

In summary, pseudoclefts in general are tolerant of epistemic modal modification, while only the canonical type can combine with deontic modals. The acceptability of epistemic modals constitutes an exception to the simple form requirement shown so far in this chapter to hold of the amalgam copula, one which is predicted by the analysis in Chapter 5.

4.5.6 Aspectual auxiliaries

The familiar pattern occurs also with aspectual auxiliaries: canonical pseudoclefts permit them, while amalgam pseudoclefts do not. I assume a neo-Reichenbachian treatment of aspect, where the aspectual predicate orders the Event Time identified by the VP with respect to a Topic Time—a time under discussion (Reichenbach 1947; Klein 1994; Demirdache and Uribe-Etxebarria 2007, 2014, etc.). In English, the present and past perfect forms are odd without an adverbial modifier or

---

[23]http://www.slate.com/blogs/lexicon_valley/2014/10/02

[24]I have noticed an epistemic use of *legitimately* among young speakers. Perhaps for these speakers, it can occur in amalgams.

a context to distinguish the Topic Time from the Event Time of the verb:

(144)   a.   #It has been sunny.

        b.   It has been sunny lately.

Similarly, out-of-the-blue perfect aspect on the copula of a pseudocleft is odd, unless there is an adverb or a perfect in the weight clause.

(145)   a.   #What he likes **has been** coffee.

        b.   What he likes **has always been** coffee.

        c.   What he **has always liked** has been coffee.

(146)   a.   #Coffee **has been** what he likes.

        b.   Coffee **has always been** what he likes.

        c.   Coffee **has been** what he has always liked.

In amalgams, however, I find the non-simple form of the copula degraded.

(147)   a.   *What he likes **has always been** he likes coffee.

        b.   *He has always liked coffee **has been** what he likes.

This pattern of judgments is verified experimentally.

4.5.6.1   Experiment 6 design

The experiment testing aspect in pseudoclefts used a 4x2 design (see Table 4.14). The sentence type factor had four levels: canonical pseudocleft, reverse canonical pseudocleft, amalgam pseudocleft, and reverse amalgam pseudocleft. The auxiliary factor had two levels: bare copula and auxiliary (perfect aspect with another perfect on the lexical verb). 8 counterbalanced versions were constructed using the Latin square procedure outlined for the other experiments. 32 participants

completed the survey.

	Bare	Auxiliary
Canonical	What you had studied was chemistry.	What you had studied had been chemistry.
Rev. Canonical	Chemistry was what you had studied.	Chemistry had been what you had studied.
Amalgam	What you had studied was you had studied chemistry.	What you had studied had been you had studied chemistry.
Rev. Amalgam	You had studied chemistry was what you had studied.	You had studied chemistry had been what you had studied.

Table 4.14: Factor design for the aspect survey

4.5.6.2 Experiment 6 results

Normalized results were inspected for outliers. 13 participants in this survey rated bare amalgams below a threshold of $z=-0.5$. Their results were excluded, leaving 19 in the sample. While aspectual auxiliaries reduce the acceptability of all pseudoclefts, the effect is more deleterious in amalgams. The judgments of aspectual auxiliaries in canonical pseudoclefts remain in the positive range (compared to the $z=0$ threshold), while the judgments of aspectual auxiliaries in amalgams are primarily in the negative range, skewing low. These results are summarized in Figure 4.11.

A mixed effects model summarized in Table 4.15 finds a significant interaction between sentence type (canonical vs. amalgam) and aspect, as compared to the null model. Figure 4.12 shows the two-way comparison of sentence types.

Model	Df	AIC	BIC	logLik	deviance	Chisq	Chi	Df	Pr(>Chisq)
Null	6	700.08	726.24	-344.04	688.08				
Test	7	697.08	727.59	-341.54	683.08	5.0043		1	0.02528*

Table 4.15: Fixed effects linear model of interaction between aspect and sentence type (2x2)
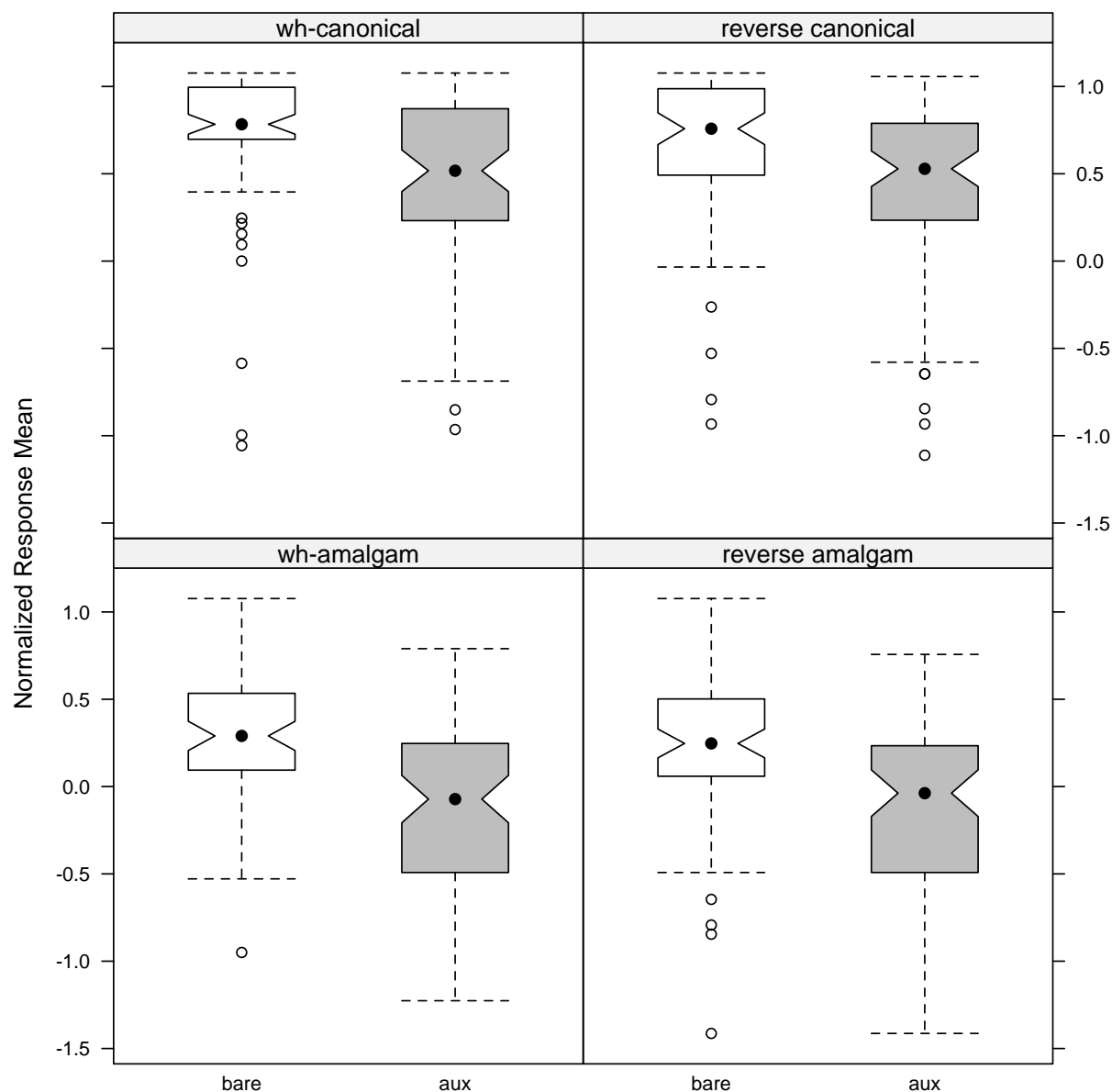
Figure 4.11: Effect of aspectual auxiliaries in specificational copular sentences

4.5.7 Restrictions on baggage: summary

This concludes the empirical profile of the restrictions on the pseudocleft copula. Other than epistemic modals, the amalgam copula must take simple form (*is, was, are*, or *were*). Most previous literature reports that the canonical pseudocleft copula, particularly in *wh*-initial pseudoclefts, must
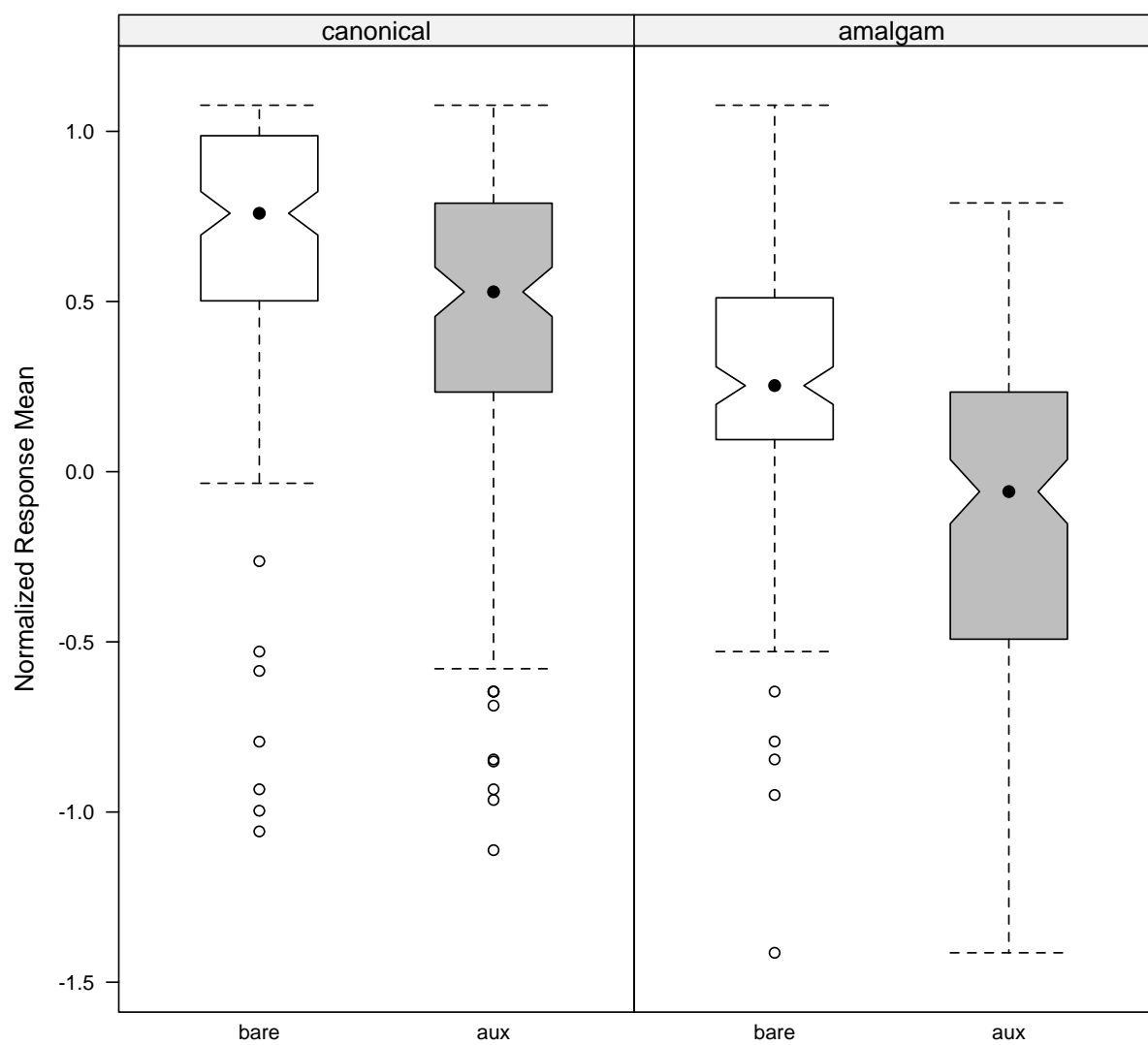
Figure 4.12: Effect of aspectual auxiliaries by canonical vs. amalgam type

remain simple (or harmonize with the lexical verb). These surveys have shown that although the simple form is generally preferred, baggage on the canonical copula is not ungrammatical. The contrast between the canonical and amalgam copulas is robust.

4.5.8 Form of copulas in double-*is*

The present section examines the form of the two copulas in double-*is* (cop_1 and cop_2), which, like their counterpart in the other amalgam sentence types, tend to remain in simple form.[25] Double-*is* warrants separate discussion for obvious reasons: it has two copulas, so it is not immediately clear which of the two should be expected to pattern with the amalgam copula.

The literature on double-*is* observes that the morphological form of both cop_1 and cop_2 is typically restricted to *is* or *was*, like the amalgam pseudocleft copula. It has been claimed that cop_2 is more restricted than cop_1: if cop_1 is *is*, then cop_2 must be *is* (Coppock and Staum-Casasanto 2004; Massam 1999). Examples like (148c) do in fact occur in COCA, but they are infrequent.

(148) a. The thing **was, is** that we had no control over the situation.

(Massam 1999:349 (39b))

b. My feeling **was, was** that she doesn't have a professional hold on the situation.

(Coppock and Staum-Casasanto 2004:3 (14))

c. *The thing **is, was** that we had no control over the situation.

(Coppock and Staum-Casasanto 2004:3 (16))

Other baggage is extremely restricted in double-*is*. Previous literature reports that it is not grammatical at all. For example, negation in combination with either cop_1 or cop_2 results in total unacceptability.

(149) a. The issue **is(*n't), is(*n't)** that I'm hungry.

b. What bothers me **is(*n't), is(*n't)** that I'm hungry.

The general finding of the corpus search and acceptability survey is that cop_2 is extremely re-

---

[25]Portions of this section are based on O'Neill (to appear).

stricted: like the amalgam pseudocleft copula, it must take simple form (although some survey participants accept epistemic modals), while cop_1 shows some flexibility.

4.5.8.1 Double-*is* in COCA

To test the restrictions on the form of the copula in real usage, I examined occurrences of double-*is* in 85,000,000 words of spoken English from 1990–2010 recorded in COCA. Since double-*is* includes a searchable string (two adjacent occurrences of the copula), it is amenable to corpus-based research.

I counted all sentence tokens with two adjacent copulas, such that one of the copulas could be considered syntactically extraneous. Strings with two copulas where both functioned as the main verbal element of a canonical clause were not counted (for example, pseudoclefts where the anchor was itself copular (150) or specificational copular sentences where the postcopular constituent was a polar question introduced by a copula (151)).

(150)   a.   What it **is is** a new invention.

            b.   What my concern **is is** that you don't know what you're doing.

(151)   a.   What I want to know **is: is** he a good guy?

            b.   Her question **is: is** this safe?

I counted double-*is* sentences regardless of weight form (DP or CP). The form of the copulas in the counted items was not restricted to *is* or *was*; it included both sentences where the copulas were bare and sentences where one or both were associated with auxiliaries and/or modals. It also included instances of non-finite *being* in the role of cop_1 (152h).

(152)   Examples of search strings

a.   is is, is was, was is, was was

b.   is are, are is, are are

c.   was were, were was, were were

d.   could be is, is could be, could be could be

e.   is isn't, isn't is, isn't isn't

f.   is has been, has been is, has been has been

g.   to be is, is to be

h.   being is

      etc.

The search yielded 514 double-*is* sentence tokens. Most had a DP anchor—typically simple, like *the problem*, but sometimes modified or including a relative clause. 62 tokens featured a CP weight (free relative or indirect question); these I will refer to as double-*is* pseudoclefts. One token, not counted in the table below, was non-specificational, having a fronted predicate in precopular position (153). Also not included in the basic count is an occurrence of double-*is* with *that's x is y* (154).

(153)   ...well, implicit in what you're saying **is is** that there is any shred professional journalism

         left.

(154)   ...that's much the problem with reporting on Pakistan **is is**...

In the sample, cop$_1$ and cop$_2$ are typically both *is* (404 or 78% of the tokens).

The counterweight was almost always a proposition: typically a finite clause introduced by *that* (337 tokens), a bare sentential counterweight (119 tokens), or another type of propositional element (*to*-infinitive, root question, CP introduced by *because*, etc.)—only 14 tokens were not

formally clausal, but these could be interpreted as remnants of elliptical clauses.

Table 4.16 summarizes the forms of cop_1 and cop_2 in the corpus sample. Like the amalgam pseudocleft copula, cop_2 is almost always bare, matching informally collected acceptability judgments (155).[26]

(155)   a.   *The issue **is, could be** he left his bag at home.

     b.   *What she thought **was, had been** that you had done it already.

     c.   *What he sees **is/are, are** he sees disasters and problems everywhere.

Cop_1, however, shows some flexibility: it is non-bare in 60 sentence tokens. 29 of these tokens feature the participial *being*, and the other tokens include modals, e.g., (156), auxiliaries, plural, and non-finite forms.

(156)   The only difference **could be is** that you want to give more state control...   (COCA)

	Bare	*being*	Modal + *be*	Aux + *been*	*are*	Other
Cop_1:	454	29	15	6	5	5
Cop_2:	512	0	0	1	1	0

Table 4.16: Form of cop_1 and cop_1 in COCA

Interestingly, the form of cop_1 only combines with baggage when the weight is a plain DP, and never when it is a CP, as in a double-*is* pseudocleft. To my knowledge, this correlation has not been noticed before. Table 4.17 and Figure 4.13 show the distribution of functional material on cop_1 by weight expression type. The next chapter will take a closer look at the relationship between double-*is* and amalgams, proposing that sentences containing a cop_1-cop_2 string can arise from two underliers.

To set the stage for the analysis in Chapter 5, consider the co-occurrence of the amalgam

---

[26]It is unclear whether the token with *are* is a disfluency; it follows another occurrence of *are*.

	Total	Cop$_1$ bare	Cop$_1$ *being*	Cop$_1$ non-bare
Total	514	454	29	31
DP double-*is*	452	392	29	31
CP double-*is*	62	62	0	0

Table 4.17: Precopular constituent and form of cop$_1$ in COCA.
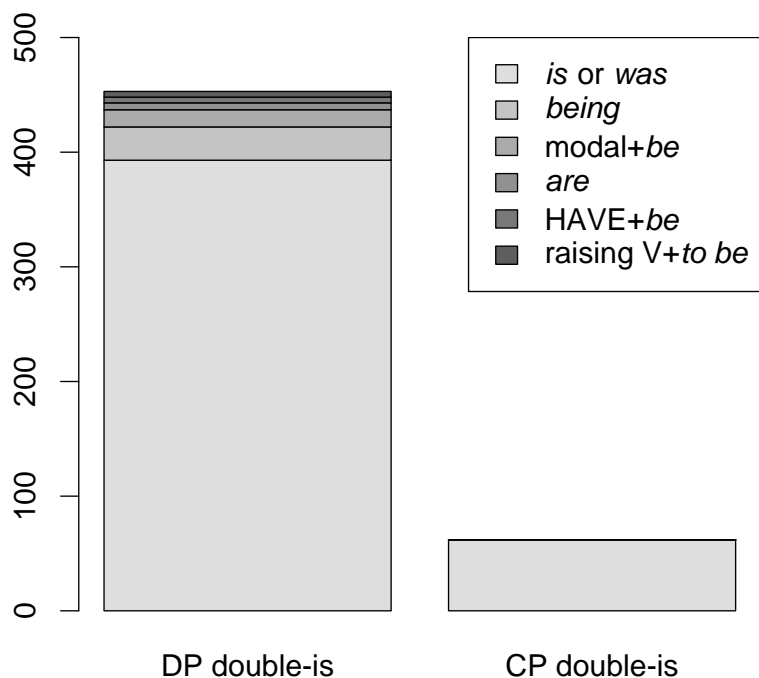
Figure 4.13: Distribution of functional material on copula 1 in COCA

sentence type and double-*is*.

(157)   a.   [What I need] **is, is** I need a cup of coffee.

b.   [That's what I need] **is, is** I need a cup of coffee.

While in the DP-initial double-*is*, cop$_1$ could be analyzed as the main verbal element of a free-relative weight with a covert *what* or null operator (see, e.g., Massam 1999), in the CP-initial version, as in (157), this parse is not plausible. The amalgam double-*is* examples in (158) are

ill-formed.

(158)   a.   *[What [what I need] **is**] **is** I need a cup of coffee.

          b.   *[What [that's what I need] **is**] **is** I need a cup of coffee.

In the 62 tokens of pseudocleft double-*is*, which all have a bare form of the copula, we therefore expect the same constraints to govern the form of both cop_1 and cop_2, especially since double-*is* occurs in combination with other (weight-initial) amalgams. While the fact that non-bare cop_1 is totally unattested in COCA with CP-initial double-*is* is suggestive, the relatively low frequency of CP-initial double-*is* makes the evidence imperfect. The next section presents an acceptability survey showing that non-bare forms of cop_1 in double-*is* amalgam pseudoclefts are not merely unattested, but actually unacceptable and ungrammatical.

4.5.8.2   Experiment 7 design

A survey compared the acceptability of non-bare cop_1 in double-*is* amalgam pseudoclefts to the acceptability of a parallel manipulation of the sole copula in an amalgam pseudocleft. The copula in the two environments is predicted to show similar behavior. Whatever limits the combinatorial possibilities of the sole copula in amalgam pseudoclefts ought to limit them in double-*is*. The experiment was designed in the same way as the others reported in this chapter.

The survey items were divided into four conditions (2x2 design illustrated in Table 4.18). Four versions of the survey were constructed, with 16 experimental items and 16 fillers counterbalanced by Latin square. Types of "baggage" considered included aspectual and modal auxiliaries. The distribution of different baggage types in the lexicalizations was proportionate to their respective distributions in the COCA sample. Items were presented in written form, so a comma was used to separate the cop_1-cop_2 string, following the conventional written form of the construction (Ander-

sen 2002). 20 participants completed the survey.

	Single *is*	Double *is*
Bare	What he likes is he likes pizza.	What he likes is, is he likes pizza.
Baggage	What he likes could be he likes pizza.	What he likes could be, is he likes pizza.

Table 4.18: Factor design for double-is survey

Normalized results were analyzed using a linear mixed effects model with random intercepts for item and subject. The model found a significant main effect of baggage overall ($p<0.001$; $t=-4.3$) and within each sentence type (see Figure 4.14).[27]

As expected, there was no interaction between the factors, because baggage has a similar effect on cop_1 ($p=0.0009$) and cop_2 ($p=0.0028$); that is, it significantly lowers acceptability in both single-*is* amalgam specificational pseudoclefts (with cop_2) and double-*is* amalgam specificational pseudoclefts (with baggage on cop_1).[28]

The results of this experiment indicate that a string containing multiple occurrences of the copula can come about in more than one way. In the double-*is* structure that co-occurs with amalgam sentences, cop_1 is not an ordinary verb; like its counterpart in amalgams, it must occur in bare form. When cop_1 combines with baggage, however, it seems to be part of an ordinary verbal projection, and the precopular constituent must be a simple DP, like an ordinary structural subject. In the next chapter, I will consider both cases, and propose two different structures: one related to the amalgam pseudocleft, and the other related to the non-reversible amalgam types.

---

[27]Acceptability ratings of baggage on the sole copula of the amalgam pseudocleft skew high, because epistemic modals on the amalgam copula are variably acceptable, in contrast to other types of baggage.

[28]Double-*is* sentences had lower acceptability overall than non-doubled sentences, but the normalized mean was positive. I attribute the lower ratings to the difficulty of processing three adjacent finite verb forms (the lexical verb of the pseudocleft, followed by two occurrences of the copula), and to prescriptive bias.
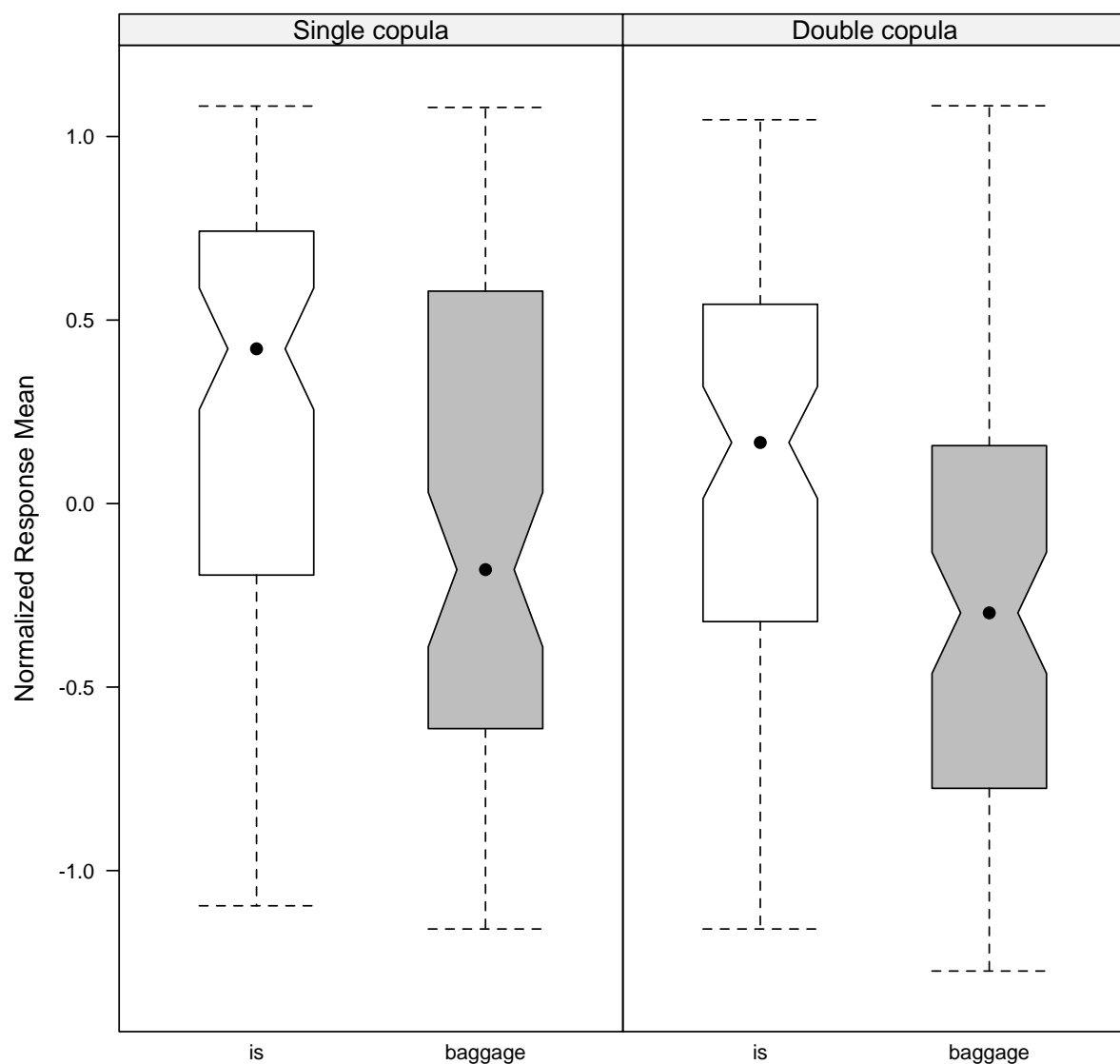
Figure 4.14: Effect of baggage on single and double-*is* pseudoclefts

4.6 Semantic ingredients of canonical vs. amalgam sentences

As we saw in Chapter 2, much of the literature on specificational copular sentences addresses semantic questions. What is the semantic type of the two elements that stand in a specification relationship? What is the semantic role, if any, of the copula that relates the elements? What licenses the special properties of specificational sentences that distinguish them from predicational ones,

e.g., allowing the more predicative expression in precopular position and licensing connectivity effects? This section presents the issues at stake, and continues the work of this chapter, comparing canonical and amalgam specificational sentences, on the semantic front.

Section 4.6.1 provides an overview of the semantic properties of the anchor/weight expression in both canonical and amalgam sentence types, concluding that it has an intensional predicate denotation in the former, and an intensional proposition denotation in the latter.[29] These denotations can both arise from DPs and from CPs; this overlap makes the task of identifying the semantic type of the weight quite challenging. In both sentence types, the anchor/weight expression is an intensional object that takes the expression on the other side of the copula as its logical subject (possibly mediated by an IDENT type-shifter; Partee 1987)—in both, it is a predicate, a function from worlds to sets of simple objects (individuals or truth values). With the pieces of the puzzle characterized, section 4.6.2 sets the stage for Chapter 5, showing that the amalgam sentence type is a species of question-answer clause.

4.6.1   Denotation of the weight

The weight expression has been a key player in analyses of the semantics of specificational sentences, for several reasons. First, its semantic type is difficult to determine; secondly, it can occupy the structural subject position even though the postcopular constituent is the more referential element in the relation; lastly, the weight seems to be responsible for connectivity effects, licensing dependent expressions in the counterweight/value without c-commanding them.

The different camps in the debate about the classification of specificational sentences tend to

---

[29]Recall that I use the term "weight" to refer to the dependent of the copula containing the "anchor", the expression whose content is specified by the value. In canonical pseudoclefts, the anchor corresponds to the entire weight, while in amalgams like *that's x is y*, the anchor is a subconstituent of the weight. While the discussion of canonical specificational copular sentences in this section most directly concerns the semantic status of the anchor—the expression whose content is specified—I will use the term "weight" throughout, for convenience, since the anchor and the weight are coextensive in the canonical sentence type.

align with different analyses of the semantic type of the weight. If predicational and specificational sentences are underlyingly the same, then the weight is a predicate expression (e.g., Partee 1986, 2000; Mikkelsen 2005; Geist 2007; den Dikken 2006). If instead, specificational sentences are equative, then the weight is most straightforwardly treated as an individual-denoting expression (Heycock and Kroch 1999). Another version of the equative analysis treats *both* the weight and the value as property expressions (Jacobson 1994; Sharvit 1999, 2009). Elaborating on the old intuition that the weight expression is somehow "less referential" than the value, Schlenker (2003), Romero (2005, 2006), Comorovski (2008) (and others) treat the weight expression as an intensional object, although they differ with respect to its specific type. For Romero (2004, 2005) and Comorovski (2008), it is an intensional individual (similar to Heycock and Kroch's 1999 analysis). For Schlenker (2003), who takes seriously the old observations of Faraci (1971) and Ross (1972) that specificational sentences are bisentential, like self-answering questions, the weight is a proposition-intension—a question.

Although there is no consensus, much of this literature converges on the idea that the weight is of a higher type than the value. This goes back to Higgins's (1979) description of the weight (which he calls "superscriptional") as analogous to the heading of a list. The copula has a colon function (cf. Koster 2000; Lambrecht 2001; O'Neill 2012; den Dikken 2013) and specifies the members of the list. This is a useful way of thinking about the meaning of specification, and of capturing, if somewhat informally, the strong intuition that the weight is not a referential expression (contra Heycock and Kroch 1999). Pursuing this intuition, I propose in this subsection that although the specific type of the weight differs between the canonical and amalgam sentence types, the relationship between the anchor/weight and its value/counterweight is the same: it is simple intensional function application.

4.6.1.1 *Wh*-**clause weights**

4.6.1.1.1 **Free relatives in canonical pseudoclefts**

The first question to confront is the category and logical type of the *wh*-weight in a canonical pseudocleft. Given its form and nominal distribution, it seems to be a free relative. Free relatives are DP-like: they are typically interpreted as definite descriptions (Caponigro 2003), which can have predicative (159) or referential readings (160). They carry a uniqueness presupposition contributed by an iota operator.

(159)   a.   What he saw was his mother.

    b.   His mother was what he saw.

(160)   a.   What he saw was beautiful.

    b.   I took a picture of what he saw.

The inanimate *what* in (159) suggests that the free relative is not referential—if it were, we would expect to find the animate *who*. Free relatives like these distribute externally like DPs: headed relatives occur in the same environments.

(161)   a.   The person that he saw was his mother.

    b.   His mother was the person that he saw.

(162)   a.   The person that he saw was beautiful.

    b.   I took a picture of the person that he saw.

I defer syntactic discussion of the DP-nature of free relatives until Chapter 7. It suffices for present purposes to point out that in the relevant *canonical* specificational environment, free relatives and surface DPs play the same semantic role.

4.6.1.1.2 *Wh*-words

Higgins (1979), Ross (2000, 2004), Caponigro and Heller (2007), and others have pointed out that there is an acceptability hierarchy with respect to which *wh*-words occur freely in the weight clause of a pseudocleft. By far the most frequently occurring (and most widely acceptable) *wh*-word is *what* (Higgins 1979; Ross 2000). A cline of acceptability, particularly in the weight-initial order, includes *where*, and then *how, why,* and *when*, and lastly, *who* and *which*. Examples in the weight-initial order are given in (163).

(163)   a.   What he ate is an apple.

          b.   Where he went is (to) the store.

          c. %How I found it is by googling it.

          d. %Why he's mad is because his sister lied.

          e. %When we're leaving is at five o'clock.

          f. %Who I saw is the teacher.

          g. *Which book I read is *War and Peace*.

Speaker variation is rampant (Ross 2000). I speculate, as I did for some of the other properties discussed above, that some of the variability is due to structural ambiguity between the canonical and amalgam source, with amalgam pseudoclefts tolerating a wider range of *wh*-words than canonicals. I will return to this point shortly, although I leave experimental exploration of the variation for future research.

Caponigro and Heller (2007) explicitly argue against treating weight clauses in any pseudocleft as indirect questions (contra Ross 1972, 2000; den Dikken et al. 2000; Schlenker 2003). They point out that the subset of *wh*-words that can occur freely in specificational pseudoclefts ((a) examples)

is the same as the subset that occurs in DP-like free relatives ((b) examples), in contrast to indirect

questions ((c) examples), where all *wh*-words can occur.[30]

(164)   what

    a.   What he saw is a nuthatch.

    b.   I saw what he saw.

    c.   I wonder what he saw.

(165)   where

    a.   Where he went is (to) the store.

    b.   I went where he went.

    c.   I wonder where he went.

(166)   how

    a.   How he found it is by googling it.

    b.   I found it how he found it.

    c.   I wonder how he found it.

(167)   why

    a.   Why he did it is because he's angry.

    b.   *I did it why he did it.   (on the DP reading)

    c.   I wonder why he did it.

(168)   when

---

[30]Caponigro and Heller's (2007) discussion includes *where, what, who, which*, and *how much*. Judgment diacritics on the (a) examples are based on theirs, and do not reflect the variability mentioned above. They do not discuss *when* or *why*.

     a. *When he will leave is at five o'clock.

     b. (i)   I will leave when he will leave.   (PP-like free relative)

           (ii) *When he will leave is too late.   (*DP environment)

     c. I wonder when he will leave.

(169)   who

     a. *Who did it is John.

     b. (i)   I saw who you saw.

           (ii) *Who did it is ugly.

     c. I wonder who did it.

(170)   which

     a. *Which book I bought is *War and Peace*.

     b. *I bought which book you bought.

     c. I wonder which book he bought.

Caponigro and Heller (2007) also present cross-linguistic arguments against treating the CP of a pseudocleft as an indirect question. They illustrate that free relatives and indirect questions are formally different in Macedonian, Hungarian, Hebrew, and Wolof, and weight CPs in these languages use the free relative form. Since CP weights do not take the form of indirect questions in many languages, they argue that a question-answer pair analysis of pseudoclefts is not well motivated. Den Dikken et al. (2000) point out, however, that these arguments need not apply to languages where pseudocleft weight form is consistent with indirect question form.

While I think Caponigro and Heller are correct that *canonical* pseudoclefts feature free relatives, their argument is clouded by some inconsistencies.[31] First of all, the status of *why* and *who* is problematic. Caponigro (2003) argues convincingly that *why* lacks a DP-like free relative denotation, and yet pseudoclefts with *why* are well-formed. *Who* can certainly introduce DP-like free relatives, and yet it is degraded when it introduces a pseudocleft. In fact, the lower acceptability of *who* is restricted to its occurrence in initial position, whether in a predicational or specificational context. This is poorly understood, and I have no explanation to add, except to point out that *who*-free relatives improve greatly in non-initial position. (The examples in (171) reflect my judgments.)

(171)   a. ??Who did it is John.   (specificational)

       b. *Who did it is friendly.   (predicational)

       c. John is who did it.   (predicational or specificational)

Most importantly, Caponigro and Heller (2007) do not distinguish explicitly between canonical and amalgam pseudoclefts, arguing indiscriminately against the indirect question analysis of pseudocleft weights, which den Dikken et al.'s (2000) only propose for so-called "Type A" (amalgam) pseudoclefts.

---

[31]The reader might wonder what blocks *-ever* relatives from occurring in pseudoclefts:

(i)   a.   *Whatever book you bought was *War and Peace*
    b.   *Whenever you arrive is at five o'clock.

These apparent counter-examples to the free relative analysis can be explained by the exhaustive nature of specification. The wh-*ever* headed clause has been analyzed as a headed relative containing a free-choice operator (Dayal 1997; Donati and Cecchetto 2011, etc.). Such quantified DPs are incompatible with the uniqueness reading of the weight term of a pseudocleft.

4.6.1.1.3   Indirect question weights in amalgams

When canonical and amalgam specificational pseudoclefts are teased apart, there is evidence that the weight clause in an amalgam can be a CP indirect question, rather than a DP free relative (see also the arguments in den Dikken et al. 2000). For instance, in the examples below (amalgam counterparts to (164)–(170)), the amalgam pseudocleft admits the full range of *wh*-words, at least for some speakers. There is variation—some speakers prefer the free relative type expressions—but informal surveys and reports in the literature (e.g., Ross 2000; Hedberg 2008) show a clear contrast between the two sentence types.[32]

(172)   a.   What he saw was he saw a nuthatch.

      b.   Where he went is he went to the store.

      c.   How I found it is I googled it.

      d.   Why he did it is (he did it) because he was angry.

      e.   When we're leaving is we're leaving at five o'clock.

      f.   %Who I saw is I saw the teacher.

      g.   %Which book I bought is I bought *War and Peace*.

Further evidence that the weight clause of an amalgam can be an indirect question is that it permits multiple *wh*-expressions, which are not allowed in free relatives, but which are allowed in indirect questions (den Dikken et al. 2000) (173).

(173)   a.   **What** we should put **where** is we should put the lamp in the corner and the desk against the wall.

---

[32](172g) is deemed acceptable in Hedberg (2008: (18)), on a par with the following Topic-Comment structure:

(i)   As for which book he bought, he bought *War and Peace*.

    b.   (i)   *I unpacked **what** we should put **where**.

         (ii) *I dislike **what** we put **where**.

    c.   I wonder **what** we should put **where**.

Not coincidentally, multiple *wh*-expressions are degraded in canonical pseudoclefts. The following examples use baggage on the copula to disambiguate in favor of the canonical pseudocleft:

(174)   a. ***Who** ate **what** had been Mary a bagel and John some pie.

        b. ***What** we put **where** will be the lamp in the corner and the desk against the wall.

    Amalgams also allow sluicing in postcopular position, provided the antecedent of the sluiced clause is recoverable (175), but the canonical weight does not.

(175)   a.   You should invite Jack is who ~~you should invite~~.

        b.   They went up the hill is where ~~they went~~.

(176)   *Up the hill will be where ~~they go~~.

The asymmetry between (176) and (175) follows from the fact that sluicing is available in indirect questions, but not in free relatives. Compare (177) and (178).

(177)   a.   She wants to invite someone, but I don't know who ~~you should invite~~.

        b. *She went where I went, and he went where ~~I went~~ too.

Sluicing is available in apparent canonical pseudoclefts where the weight is strongly presupposed, as in (178).

(178)   What am I talking about? Your major budget crisis is what ~~I'm talking about~~!

Such cases are actually partially elliptical reverse amalgam pseudoclefts, however, since they are unembeddable in ECM contexts, and do not license baggage on the copula in finite contexts (see below for more discussion of elliptical reverse amalgam pseudoclefts):

(179)   a.   What will she drink? *I expect beer to be what ~~she will drink~~.

          b.   What will she drink? *Beer will be what ~~she will drink~~.

Likewise, *else*-modification is possible in amalgams (180) (marginally, for some speakers), but not canonicals (181), in parallel to their availability in indirect questions (182), but not free relatives (183).

(180)   What else he bought was he bought a sofa and chairs.

(181)   *What else he bought had been a sofa and chairs.

(182)   I wonder what else he bought.

(183)   *I bought what else he bought.

Asymmetries between the canonical and amalgam specificational pseudoclefts thus correlate with asymmetries between free relative and indirect question type weights.

4.6.1.1.4   The weight in reverse pseudoclefts

Interestingly, in both the canonical and amalgam sentence types, when the value is in initial position, the acceptability of *wh*-words lower in the cline improves.

(184)   a.   An apple is what he ate.

          b.   (To) the store is where he went.

          c.   By googling it is how I found it.

> d.   Because his sister lied is why he's mad.
>
> e.   (At) five o'clock is when we're leaving.
>
> f.   The teacher is who he saw.
>
> g.   *%War and Peace* is which book he bought.

The improved acceptability of the full range of *wh*-words suggests that these expressions are *not* necessarily free relatives. Instead, they appear to be indirect questions. The pattern merits a closer look.

Ambiguity between the canonical and amalgam pseudocleft proves yet again to be a confound. While ellipsis down to the focus in the weight-initial amalgam is commonplace (analogous to fragment answer ellipsis), it has not yet been recognized in the reverse amalgam pseudocleft. This appears to be precisely what is going on in the surprisingly acceptable examples in (184).[33]

Take the example with *which*, the one that was most strikingly bad in weight-initial order. (The auxiliary is used to disambiguate the *wh*-initial version in favor of the canonical parse.)

(185)   a.   *Which book he bought had been *War and Peace*.

         b. *%War and Peace* is which book he bought.

Which book he bought in (185b) is pronounced with a very specific prosody: there is a sharp pitch drop before the weight clause, and it is deaccented. An example like this is likely to occur in a context where *which book he bought* has a linguistic antecedent (it is strongly T-presupposed). It thus has an almost echoic function, in contrast to the freer prosody and discourse distribution of a postcopular weight clause like *where he went*. The weight clause is always a topic, but it need not always have an immediate antecedent. This echoic use is specific to the amalgam pseudocleft, as discussed in Chapter 2. A quick syntactic test will show that, as we might suspect, (185b) has an

---

[33]Once again, I set aside the behavior of *who*, which remains puzzling.

amalgam pseudocleft underlier.

(186)   *War and Peace* had been which book he bought.

The order asymmetries in the acceptability of *wh*-words may therefore be a red herring—another consequence of the conflation of canonical and amalgam pseudoclefts.

4.6.1.1.5   The weight in non-reversible amalgams

That's x is y and question-answer amalgams pattern with their pseudocleft counterparts. They feature an indirect question CP inside the weight clause, which allows the full range of *wh*-words ((187)–(188)) and multiple *wh*-words ((189)–(190)).[34]

(187)   a.   That's what he saw, was he saw a nuthatch.

         b.   That's where he went, is he went to the store.

         c.   That's how I found it, is I googled it.

         d.   That's why he did it is, (he did it) because he was angry.

         e.   That's when we're leaving, is we're leaving at five o'clock.

         f.   That's who I saw, is I saw the teacher.

         g.   That's which book I bought, is I bought *War and Peace*.

(188)   a.   Guess what he saw, was he saw a nuthatch.

         b.   You know where he went, is he went to the store.

         c.   Guess how I found it, is I googled it.

         d.   You know why he did it, is (he did it) because he was angry.

         e.   Guess when we're leaving, is we're leaving at five o'clock.

---

[34]*Who*-indirect questions are perfectly acceptable in these contexts, where they are not initial. The limited distribution of *who*-weights seems tied to its linear position more than anything else.

      f.   You know who I saw, is I saw the teacher.

      g.   Guess which book I bought, is I bought *War and Peace*.

(189)   That's who did what, is John wrote the draft and Mary edited it.

(190)   You know who did what, is John wrote the draft and Mary edited it.

I conclude that the amalgam copular sentence can have an indirect question as its weight/variable, while the canonical pseudocleft can only have a free relative.

4.6.1.1.6   Free relatives and indirect questions as predicates

The CPs occurring in canonical and amalgam specificational sentences show parallel syntactic and semantic behavior: both are (intensional) predicates of the focused subject (value or counterweight). The canonical type CP weight is a predicate of individuals, while the amalgam type weight is a predicate of propositions. Two semantic representations for the same surface CP are given below, to clarify the role of a question-denoting expression as a predicate in amalgams.

In both cases, the CP introduces an operator binding a variable in an open proposition. If it is a free relative, the variable corresponds to the type of entity the *wh*-word ranges over (e.g., an individual); if it is a question, the variable corresponds to the proposition providing an answer in which the variable bound by the *wh*-word is replaced by its value in the actual world (Groenendijk and Stokhof 1984, 1997).

(191)   a.   What I ate $= \lambda w'[\lambda x[\text{I ate x in } w'] = \lambda x[\text{I ate x in w}]]$

      b.   What I ate $= \lambda w' \lambda w''[\lambda x[\text{I ate x in } w'] = \lambda x[\text{I ate x in w}]]$

In (191b), the question denotation is a unique proposition intension, rather than a set of propositions, while in (191a), the free relative denotation is a predicate intension. To avoid postulating

a separate equative copula just for specificational amalgams (as Schlenker 2003 does), where two proposition intensions are related, I follow the mainstream assumption that the IDENT type-shifter can convert the proposition intension into a singleton set of proposition intensions. The resulting meaning is, at an intuitive level, no different from what is expressed by (191b). Alternatively, I could adopt a Karttunen (1977) semantics for questions, which treats the denotation of a question as the set of its true answers, rather than the unique true answer. In either case, the weight in an amalgam is of a type that takes a proposition—the counterweight—as its argument.

4.6.1.2   DP-weights

DP anchors/weights occur in both canonical and amalgam specificational copular sentences. In fact, as we have already seen, free relatives occurring in canonical pseudoclefts have a DP-like distribution. In canonical specificational copular sentences, DP anchors are interpreted as intensional predicates: functions from worlds to sets of individuals. They take a world argument and an individual argument, the latter provided by the focused value. In amalgams, the same DPs are interpreted as intensional predicates of propositions, i.e., as concealed questions. They take a world argument and a proposition argument. An examination of simple form DPs in the two sentence types highlights the semantic difference between them.

4.6.1.2.1   DP predicates in canonical specificational sentences

The DP weight types that can occur in specificational sentences, (192), include definite descriptions, partitives, some indefinites (although these may require a different treatment; see Mikkelsen 2005; den Dikken 2006), but not proper names (193), deictic pronouns (Higgins 1979), or strongly quantified DPs (194) Mikkelsen (2005). The allowable DP-types have in common that they are all predicates. The neuter pronoun in the tag question is a property anaphor, and cannot take a human-denoting referential expression as an antecedent, so it is incompatible with predicational

and equative sentences where the precopular expression is referential.

(192)   a.   The president is Barack Obama, isn't it?

    b.   Her best friend is Jack, isn't it?

    c.   One of the most interesting people we met was Jill, wasn't it?

    d.   A really great doctor is John, isn't it?

(193)   a.   Mr. Brown is John, it's *it/he?

    b.   *He is John, isn't *it/he?

(194)   *Every book I read was *War and Peace*, *Go, Dog. Go!*, and *The Minimalist Program*.

While having a predicate denotation is a necessary condition for serving as a weight in a specificational sentence, it is not sufficient: not all predicative expressions can be weights. First of all, the weight must be nominal enough to serve as a structural subject. An adjectival predicate is therefore not a good pseudocleft weight, (195).

(195)   a.   *Nice is Mary.

    b.   *Very difficult is this book.

Of course, adjectival values can occur as logical subjects in pseudoclefts.

(196)   Important to himself is what John is.

In such cases, however, the adjective phrase is interpreted as a referential expression—it is the subject, not the predicate of the clause. Its nominality is likely contributed by a type-shifter (NOM, Partee 1987; Chierchia and Turner 1988), which maps a property onto its entity-correlate (see also Jacobson 1994). A non-DP predicate cannot be a structural subject.

There are also some predicative DPs that do not fare well as weight expressions. DP-weight amalgams, which will be examined more closely later in this section, are particularly picky—even some definite DPs do not make good weights in initial position. In both the canonical and amalgam type, indefinites are severely restricted in specificational sentences, while they are perfectly compatible with predicative interpretations.

(197)   a. ??A student is Mary.

        b. *A student is Mary is a student.

        c. Mary is a student.

The privileged status of definite DPs in specificational sentences is typically captured in terms of information structure (Mikkelsen 2005; Geist 2007; Comorovski 2008). Mikkelsen (2005) proposes that specificational sentences are marked in aligning the predicate expression with the subject position, which is possible because the predicate is a topic. Since it is a topic, it must have certain contextually supplied properties. The challenge is to explain why simple indefinites cannot serve as weights. For example, even when an indefinite DP is "maximally topical", it cannot be the weight of a specificational sentence.

(198)   Who around here is a doctor?

        a. Bill is a doctor. John is a doctor, too.

        b. Bill is a doctor. #A doctor is John, too.   (Mikkelsen 2005: (8.45))

Interestingly, a heavy indefinite can introduce a specificational sentence:

(199)   a. A doctor who might be able to help you is John.

        b. A great book to read when you have a lot of time on your hands is *War and Peace*.

Mikkelsen attributes the unacceptability of simple indefinite weights to an interaction between the discourse conditions on topics and the discourse conditions on indefinites. The type of topic that occurs as a weight expression must be relatively discourse-old. An indefinite, by contrast, is subject to a Novelty Condition (Heim 1982:300 fn.): it must introduce a new discourse referent. A bare indefinite, therefore, cannot satisfy both criteria at the same time.

A modified indefinite, on the other hand, has a built-in link to the discourse. Its modifier allows it to satisfy both conditions at once. Comorovski (2008) proposes that the weight expression must be "indirectly contextually anchored". While the weight itself is intensional, it contains some referential expression that is interpreted with respect to the context of utterance (e.g., the extralinguistic context or the discourse). This hybrid behavior makes the DP weight's semantic status difficult to determine.[35]

4.6.1.2.2 Quantificational DP weights

Strongly quantificational DPs fail to serve as weights of specificational sentences ((194) is reproduced below).

(200)   a.   *Most people here are John, Mary, and Bill.

         b.   *Every book I read was *War and Peace*, *Go, Dog. Go!*, and *The Minimalist Program.*

Mikkelsen (2005) attributes the unacceptability of strongly quantified DPs as weights to the fact that they do not occur naturally as small clause predicates.

(201)   a.   *I consider John, Mary, and Bill (to be) most people here.

         b.   I find *War and Peace*, *Go, Dog. Go!*, and *The Minimalist Program* ??(to be) every

               book he read.

---

[35]See Chapter 8, section 8.3 for an analysis of the hybrid behavior of tense in canonical pseudoclefts.

Nevertheless, this generalization is not very strong. In other predicational environments, these expressions fare much better.

(202)   a.   I expect John, Mary, and Bill to be most students in this class.

           b.   That makes *War and Peace*, *Go, Dog. Go!*, and *The Minimalist Program* every book he read.

The problem with strong quantifiers like these as specificational weights has to do with the exhaustive identity reading of a specificational sentence. In general, quantifiers can be type-shifted into predicates via Partee's (1987) BE, which collects the members of the singleton sets in the generalized quantifier's denotation into one set. A quantifier like *most people* contains no singleton sets. As such, specificational sentences with such quantifiers in weight position will always be anomalous, because the postcopular value cannot be a member of an empty set (Mikkelsen 2005). *Every book* contains a singleton set only if there is only one book, so the sentence, *Every book is War and Peace* is interpretable if *War and Peace* is the only book, but such a sentence is still anomalous, because it violates a presupposition associated with *every* (and general plausibility) that there exists a plurality of books.

4.6.1.2.3   Indefinite DP weights and concealed questions

Heycock (2012) points out that the restriction on simple indefinites as topical weight expressions in specificational sentences doesn't hold for strong indefinites. Consider (203).

(203)   One doctor is Bill; another is John.   (Heycock 2012:230 (39))

These sentences differ from the others considered so far in that they clearly lack an implicature or entailment of exhaustivity. Instead, they evoke the "mention one" or "mention some"-type reading associated with questions.

(204)   a.   Where can I get a cup of coffee?

           b.   What's a good local restaurant?

The answers to these questions are not assumed to be exhaustive—the asker seeks a true answer, but not a complete answer. The similarity between (203) and mention-one questions suggests a parallel between CP-weights and DP-weights: they have concealed question interpretations (Schlenker 2003; Romero 2004, 2005, 2006; Nathan 2006; Frana 2010, etc.).

    These authors have observed that the DPs compatible with the weight position of a specificational sentence can also occur in the complement of verbs that select for (concealed) questions (Grimshaw 1979). Concealed question DPs include definite descriptions, possessives, partitives, and modified indefinites, but not proper names or pronouns. (Free relatives can also have concealed question denotations, a point that is obscured by the fact that indirect questions are already question-denoting expressions.) Observe the following similarities between the two DP environments.

(205)   a.   Tell me the location of the accident.

           b.   The location of the accident is Broadway and 50th.

(206)   a.   Tell me her favorite book.

           b.   Her favorite book is *War and Peace*.

(207)   a.   Tell me one of the best restaurants around.

           b.   One of the best restaurants around is Per Se.

(208)   a.   Tell me out a good restaurant in the neighborhood.

           b.   A good restaurant in the neighborhood is Seva.

(209)   a.   *Tell me John.

         b.   *John is Mr Brown, (isn't it?)

(210)   a.   *Tell me him.

         b.   *He is John, isn't it?

Indefinites in concealed question environments are subject to the same "indirect contextual anchoring" constraint that Comorovski (2008) proposes for specificational sentences. This parallel suggests that the information structure explanation for the unacceptability of bare indefinites should be reconsidered, since concealed questions do not have to be topics. Modified indefinites also receive the same "mention-one" interpretation in concealed question environments that they receive in specificational sentences.

(211)   a.   *A doctor is John.

         b.   *A doctor is John is a doctor.

         c.   *Tell me a doctor.

(212)   a.   A doctor who might be able to help you is John.

         b.   ?A doctor who might be able to help you is John could help you.[36]

         c.   Tell me a doctor who might be able to help me.

Ideally, the analysis of concealed questions and specificational weights will take into account this parallel.

---

[36]The somewhat marked status of this sentence can be attributed to the fact that the value is in the subject position of the counterweight, which is generally dispreferred in amalgams.

4.6.1.2.4   Individual concepts

Romero (2004, 2005) argues that the property concealed questions and weights share is that they are individual concepts, intensional individuals of type $<s,e>$. Certain predicates, such as *know* and specificational *be*, select intensional expressions. They can satisfy their selectional requirements either with the extension of an intensional object, or with its intension. Romero thus predicts that a parallel ambiguity (first observed in Heim (1979)) obtains in specificational sentences and sentences with concealed question complements.

(213)   John knows the price that Fred knows.   (Romero 2005: (28))

        Reading A: John knows the extension of 'the price that Fred knows'

           = John knows the answer to the price question Fred knows the answer to

           = They both know that the price is \$X.

        Reading B: John knows the intension of 'the price that Fred knows'

           = John knows the answer to the meta-question about the prices that Fred knows

           = John knows which price question Fred knows the answer to, but he doesn't necessarily know the answer.

(214)   The price that Fred thought was \$1.29 was (actually) \$1.79.   (Romero 2005: (56))

        Reading A: The question whose answer Fred thought was '\$1.29' has as its real answer '\$1.79'.

(215)   The price that Fred thought was \$1.29 was the price of milk.   (Romero 2005: (58))

        Reading B: The question whose answer Fred thought was '\$1.29' is 'How much does the milk cost?'

Either a question (Reading A) or meta-question (Reading B) interpretation can arise depending on whether the copula selects as its subject the intension of the DP, or its extension. Either way, the weight it is an intensional object. Romero thus pins the similarity between concealed questions and weights on the selecting predicate. While the weight expression is certainly an intensional object (see discussion of temporal dependence in Chapter 8), the use of numbers as values in all the relevant examples is suspect. The ambiguity is difficult to replicate in more natural examples. The two examples in (216) do not obviously differ from each other.

(216)   a.   The person that Fred thought was Mary was (actually) Sue.

    b.   The person that Fred thought was Mary was (actually) the teacher.

An additional problem with treating all weight expressions as individual concepts, which Nathan (2006) points out, is that not all DPs with individual concept readings (and which can occur in specificational sentences) can easily serve as concealed questions.

(217)   a.   ?Tell me her mother.

    b.   Her mother is Mary.

(218)   a.   ?Tell me the large city in Vermont.   (Nathan 2006:149)

    b.   The large city in Vermont is Burlington.

Strong quantifiers also behave differently in the two environments. For example, a definite noun with an embedded strong quantifier can be a concealed question.

(219)   John knows every color of the rainbow.   (Nathan 2006:115 (66))

$\forall p[\exists x : p=[\lambda w_2 \ x \text{ is a color of the rainbow at } w_2] \text{ and } C(p)] \longrightarrow [\forall w_1 \in \text{Dox}_{john(w)}$

$[p_1(w_1)=1]]$

'For all contextually relevant true propositions of the form: there is an x such that x is a

color of the rainbow in those worlds, John knows that they are true.'

The same DP cannot be used felicitously in a specificational sentence.

(220)   *Every color of the rainbow is red, orange, yellow, green, blue, indigo, and violet.

Nathan argues that concealed questions have proposition-denotations, rather than intensional

individual denotations (contra Romero 2005, 2006). They refer to the unique proposition that spec-

ifies the answer to the question. (219) is grammatical because *every color of the rainbow* refers to

a proposition in John's set of doxastic alternatives; it fails to refer to a plural individual or set of

individuals, as in (220).

4.6.1.2.5   Relational DPs

It seems, after all, that weights and concealed question DPs in canonical sentences cannot be

completely conflated; however, the analyses of concealed question and specificational weights in

Schlenker (2003), Nathan (2006), and Comorovski (2008) converge on an important attribute that

the two expression types share. Schlenker (2003) and Nathan (2006) observe that the DPs with

concealed question interpretations are relational—they have additional arguments or modifiers that

connect them to the context. Nathan describes concealed question DPs in the following way (p.19):

(221)   "A DP can be a concealed question if:

    a.   its head noun is relational: it expresses a relationship between two individuals (e.g.

        a state and its governor, a commodity and its price), or

    b.   its head noun is nonrelational, but is modified in certain ways (e.g. with a relative

        clause)."[37]

---

[37]Note that these criteria predict that (217a) and (218a) should be acceptable—Nathan provides a way to derive

In Schlenker's (2003) analysis, it is the relational nature of the concealed question weight that is the source of connectivity effects in a sentence like (222). An elided copy of the head of the weight occurs in the "answer" expression (the value), and raises to a position where it c-commands the anaphor.

(222)   His$_i$ worry is <his$_i$ worry> himself$_i$.

In fact, Schlenker proposes that such sentences have the logical structure of DP-weight amalgams. Both the major constituents are semantically propositional, just as they are in syntactically overt amalgams. The weight refers to a proposition intension: the set of worlds where the unique true proposition answering the question *what is his worry* in those worlds is the same as the unique true answer in the actual world (Groenendijk and Stokhof 1984, 1997). Since the value is actually composed of a relation between the copy of *his worry* and the focus *himself*, it is also propositional.[38]

4.6.1.2.6   Proposition-denoting DPs

Both Schlenker and Nathan (contra Heim 1979 and Romero 2005, 2006) take a concealed question DP to refer to a single proposition. Unlike Schlenker, Nathan hesitates to explicitly give concealed question DPs question denotations, but he provides a formalization of concealed question meaning that is ultimately quite similar to Schlenker's. Relational and modified nouns can be type-shifted into propositions because of the contribution of the modifier or additional argument. The modifier or argument type-shifts the bare relational noun into the set of propositions such that the value of the noun is what it is in the actual world. The definite determiner picks out the unique true propo-

---

a concealed question reading for *the large city in Vermont*, but sets aside the oddness of *her mother* embedded under *know*. Since other relational role nouns like *her teacher* sound fine in this context, there may be a lexical quirk associated with kinship terms that makes them more difficult to shift into propositional meanings.

[38]Schlenker proposes that this semantics for questions requires an equative analysis of the specificational copula. It is equally compatible with a simple intensional function application analysis, provided Partee's IDENT type-shifter can operate on propositions as well as individuals.

sition that identifies the value of the noun. Both approaches capture the empirical generalization that only relational or modified nouns occur in specificational contexts, just as Mikkelsen (2005) and Comorovski (2008) observe.

An additional asymmetry between concealed questions and weights is that a concealed question DP can ask for a pair-list answer, while a weight in a canonical SCS cannot. (Recall that the future modal is used to ensure that the sentence is parsed as a canonical specificational sentence, rather than an amalgam with ellipsis.)

(223)   I want to know the book that every student will read.

   a.   John will read *War and Peace*, Mary will read *Go, dog. Go!*, and Bill will read *The Minimalist Program*.

   b.   *The book that every student will read will be John *War and Peace*, Mary *Go, dog. Go!*, and Bill *The Minimalist Program*.

Pair-list-seeking concealed questions seem strongly propositional—they want to be answered with complete propositions, rather than fragment pair-list answers. This calls to mind the amalgam pseudocleft discussed in the last section, which does allow a multiple *wh*-question weight and pair-list answers in the fully propositional counterweight. We might wonder at this point whether the amalgam specificational copular sentence licenses a multiple concealed question DP. Sure enough, when the pair-list answers are embedded in propositional counterweights, the multiple question DP is acceptable.

(224)   The book that every student will read is John will read *War and Peace*, Mary will read *Go, dog. Go!*, and Bill will read *The Minimalist Program*.

Let us briefly take stock of the facts before turning to an inventory of DP weight types in

amalgams. Specificational weights must be either definite or modified. All weights in canonical specificational sentences have predicate readings (either lexically or as a result of applying the IDENT type shifter), but not all predicates can be weights; all weights can have concealed question denotations, but not all concealed questions can be weights. What unites the DP types surveyed above is that they are all intensional functions: the prototypical DP weight is a function from a world to a (singleton) set of individuals.

4.6.1.2.7 DP weights in amalgams

Now consider the DP-weight of an amalgam. Paralleling what we found for CP weights, the type of the DP weight in the amalgam is an intensional set of propositions, while its counterpart in canonicals is an intensional set of individuals.

While the CPs that can occur as the weight expressions in canonical copular sentences are a subset of those that can occur in the amalgam type, the reverse is true when it comes to DP weight terms. Not all predicative DPs make good amalgam weights. Only DPs with propositional content, like *problem, issue, idea, fact*, can freely occur in the weight slot of an amalgam type copular sentence.

(225)   a.   The problem is I ran out of money.

            b.   The issue is we need a new budget.

            c.   The idea is they leave them out all night.

            d.   The fact is she doesn't know what she's talking about.

(226)   a.   I ran out of money is the problem

            b.   We need a new budget is the issue.

            c.   They leave them out all night is the idea.

    d.   She doesn't know what she's talking about is the fact (of the matter).

Headed relative clause weights are similarly acceptable, underscoring the importance of propositional content.

(227)   a.   All you need is you need to take care of your hair.

            b.   You need to take care of your hair is all (you need).

  DPs that are predicates of individuals have a much more restricted distribution in amalgams.

(228)   a. ??The winner of the prize is the winner / it is John.

            b. ??The winner of the prize is John is the winner of the prize.

The examples in (228) are not completely unacceptable if the weight is strongly discourse anaphoric; it has a similar echoic character to CP weights with *which*. I will return to this observation shortly.

    Since concealed questions have proposition denotations, we expect to find a different area of overlap between those that occur in canonical specificational sentences and those that occur in amalgams. Indeed, proposition-denoting concealed questions with *every*, which cannot occur in canonical specificational sentences precisely because they are propositional and not predicative, can occur in amalgams. The relevant examples from the preceding discussion are reproduced for comparison in (b) and (c) below.

(229)   a.   Every book he read was he read *War and Peace*, *The Minimalist Program*, and *The Cat in the Hat*.

            b.   Tell me every book he read.

            c.   *Every book he read had been *War and Peace*, *The Minimalist Program*, and *The Cat in the Hat*.

The DP weight of an amalgam thus has more in common with concealed questions than its canonical counterpart.

A reasonable hypothesis, given these observations, is that all DPs with concealed question readings should work in amalgam type SCSs. This is in fact not the case, as (228a), above, illustrates. DPs that lexically denote individuals, predicates, or relational individuals can be shifted to concealed question denotations, but they cannot so easily serve in amalgam specificational contexts. Compare (228a) to (230).

(230)   Tell me the winner of the prize.

Only those that range over propositions distribute freely in amalgams.

Comorovski (2008) notes that proposition-denoting DPs make particularly good weights (with canonical specificational sentences in mind), because their uniqueness requirement is met lexically: their values are propositions, which are notionally singular. While it is possible to generate plural interpretations from conjunctions of propositions (231c), these contexts are very rare. (Recall from section 4.5.2 above that speakers showed a preference for singular marking on the copula when coordinated CPs were related to a singular weight.)

(231)   a.   The problem is I forgot to set my alarm and I didn't plan my lesson.

         b.   The problems are I forgot to set my alarm and I didn't plan my lesson.

         c.   That I forgot to set my alarm and that I didn't plan my lesson are equally problematic.

A prediction of a concealed question analysis of the amalgam DP weight, drawing on Comorovski's (2008) observation, is that simple weak indefinites fare much better as weights in amalgams than they do in canonicals.

(232)   a.   A problem is we don't have any money.

           b.   We don't have any money is a problem.

(233)   ??A problem is our budget.

Propositional DPs do not need contextual support (e.g., from a modifier) to satisfy the uniqueness requirement, because they are lexically relational/functional. Since propositional DPs have an extra argument—the proposition—they do not need to be overtly modified in order to satisfy the requirement that they be contextually anchored. Headed relatives, whose heads are not lexically propositional, have propositional content in their relative clauses, so they are perfectly acceptable in amalgams.

4.6.1.2.8   Taking stock

The puzzle that emerges from the discussion so far is that concealed questions all have proposition denotations, and all DP weights in amalgams have proposition denotations, but not all concealed questions can easily be amalgam weights. Recall that concealed questions generated from DPs like *his favorite book* do not make good amalgam weights. Why do proposition-denoting DPs, free relatives, and indirect questions pattern together, to the exclusion of DPs with *derived* propositional meanings?

4.6.1.2.9   DP weights in non-reversible amalgams

Complicating the picture, in the context of a *that's x is y* or question-answer amalgam, all question-denoting expression types become available as anchors. The examples in (234) include *wh*-expressions with question meanings and those in (235) give DPs. (Examples below are all *that's x*, but question-answer amalgams and free-*be* amalgams show the same pattern.)

(234)   a.   That's what he did is he left early.

          b.   That's who did what is John washed the dishes and Mary dried them.

(235)   a.   That's the problem is (that) he left early.

          b.   That's one of the things I noticed is (that) he left early.

          c.   That's all I worry about is (that) he might disappear.

          d.   That's the book I read is I read *War and Peace*.

          e.   That's my favorite book is *War and Peace*.

          f.   That's a teacher who can help you is John.

On the other hand, rigid designators, pronouns, and indefinites that do not have lexical proposi-
tional content, are not licit in this environment.

(236)   a.   *That's Mr. Brown, is John.

          b.   *That's him, is John.

          c.   *That's a teacher, is John.

Given the range of facts in (235), the generalization about the weight DP can be clarified. The
whole weight expression—meaning the *that's x* clause, not just the anchor—must have proposi-
tional content. Non-lexically propositional DPs survive in amalgam specificational copular sen-
tences when they are embedded within a propositional weight; they occur with impunity only in
the non-reversible amalgam types where the anchor is a subconstituent of the weight. If the anchor
is a direct dependent of the copula, it must refer to a proposition.

Another way of looking at the unacceptability of examples like (228a), which have non-propositional
DP weights, is that the "answer" must be an identity statement—another specificational sentence.
In these cases, which can be observed in acceptable form in *that's x is y*, there is a strong tendency

for only the value to be supplied overtly.

(237)   a.   That's my favorite book is ??(it's) *War and Peace*.

      b.   That's what my favorite book is is ?(it's) *War and Peace*.

      c.   ??My favorite book is it's *War and Peace*.

To my ear, (237b) is better than (237a); I have also heard many comparable examples "in the wild".

(238)   a.   "What it is, is it's the insecurity," he said. "It's the latent insecurity."[39]

      b.   No, what it is is it's reality.[40]

      c.   What it is is it's cruel.[41]

Examples like these suggest that the culprit in (228a) and (237a) is not the counterweight clause. The reduced or elliptical version is certainly preferred, but the fully overt version is not ungrammatical. A fully overt counterweight in this environment is dependent on the presence of a proposition-seeking element.

As we saw in the preceding discussion, reverse amalgams are more flexible with respect to both properties: allowing a specificational or identity sentence as the counterweight, and allowing a non-propositional weight expression.

(239)   a.   My favorite book is *War and Peace*, is my favorite book.

      b.   My favorite book is *War and Peace*, is what my favorite book is.

As I mentioned above, the constraints on examples like (239) are reminiscent of the distribution of *which*-weights. I suspect that the strong anaphoricity associated with the afterthought-like position

---

[39]Nagourney, Adam. "In Tapes, Nixon Rails About Jews and Blacks." *New York Times*. December 10, 2010. `http://www.nytimes.com/2010/12/11/us/politics/11nixon.html`

[40]`http://www.realclearpolitics.com/video/2014/10/26/.html%fireworks_wasserman_schultz_vs_reince_priebus_on_if_obamas_policies_are_on_the_ballot.html`

[41]*Scrubs*. Television. NBC.

of the weight in these sentences lowers the cost of computing an appropriate question reading for the DP. While shifting an expression like *my favorite book* to a concealed question is low-cost where the focus alternatives are individuals, shifting it to a concealed question where the focus alternatives are propositions is more costly.

4.6.1.3 Summary

In summary, the weight expression of a specificational sentence can take the form of either a CP or a DP. The weight is always a higher-type object than the value/counterweight—it is a(n intensional) predicate containing a contextually anchored or relational element. Because of these properties, it overlaps in distribution with DPs serving as concealed questions. In canonicals, the weight is a function from worlds to sets of individuals, while in amalgams, it is a function from worlds to sets of propositions. In both structures, the relationship between the weight and the value/counterweight is predicational.

4.6.2 The question-answer clause

Now that I have provided a detailed inventory of the properties of the amalgam specificational sentence and its differences from its canonical counterpart, I am ready to lay the groundwork for a specific analysis of the construction. Amalgam sentences are a species of question-answer clause: the copula relates a question to its propositional answer.[42] The question-answer meaning of the amalgam is behind the obligatory focus interpretation of the counterweight—since the counter-

---

[42]Question-answer clauses have been shown to exist in American Sign Language (Caponigro and Davidson 2011). The ASL construction allows not only *wh*-questions, but also polar questions. It differs from the English amalgam in this respect. The ASL question-answer clause is also not reversible: it only occurs in question-initial order, and seems to have a similar function to what I have been calling the "question-answer" amalgam:

(i)   You know what I'm looking for is (I'm looking for) the unicorn with the longest horn.

The English amalgam pseudocleft, as I will show in the next chapter, has a different syntax than its ASL counterpart, but the constructions are clearly related in their interpretation and information function.

weight clause always functions as the answer to a question, it cannot be backgrounded.

4.6.2.1 Connectivity effects

Question-answer clauses provide an obvious source for connectivity effects observed in specifi-
cational sentences (Akmajian 1979; Ross 2000; den Dikken et al. 2000; Schlenker 2003). Con-
nectivity effects, where the value is in a syntactic dependency with an element embedded in the
weight (and which therefore does not c-command it), have always posed a puzzle in the literature
on specificational sentences. For example, an NPI can appear in the value of a *wh*-initial pseudo-
cleft without being c-commanded by any overt negative element. The apparent licenser can instead
appear in a non-commanding position embedded within the weight clause.

(240)   What nobody bought was any wine.   (den Dikken et al. 2000)

Assigning (240) an amalgam parse, where *any wine* is the remnant of an elliptical counterweight,
makes the source of NPI licensing straightforward: it is the elliptical *nobody* c-commanding *any
wine* (den Dikken et al. 2000).

    Amalgam specificational sentences allow the full range of connectivity effects (although re-
strictions on backward ellipsis make NPI connectivity difficult to recover in the reverse pseudocleft
in (241d)).

(241)   a.   That's what nobody bought is any wine.

           b.   Guess what nobody bought is any wine.

           c.   John didn't buy the most important thing, is any wine.

           d. ??Any wine is what John didn't buy.

    Other connectivity effects involve reflexive binding, reciprocal binding, bound variable bind-
ing, opacity, and selection. These are normally syntactic dependencies that are strictly local or

defined in terms of c-command, so the following examples, where there is no overt local licenser,

suggest the presence of an elliptical licenser.

(242)   a.   That's who he is always talking about, is himself.

         b.   That's what they should do, is kiss each other.

         c.   That's who each kid called, is their own mother.

         d.   That's what I'm looking for, is the unicorn with the longest horn.

         e.   That's not what I counted, is any socks.

Even when only a DP value is pronounced in *that's x is y*, a full clausal counterweight is underly-

ingly present.

4.6.2.2   Two speakers vs. one speaker

Since the question-answer pair of an amalgam is asserted by a single speaker, the pragmatic con-

gruence requirements are even more rigid than in cross-speaker discourse (den Dikken et al. 2000).

Compare (243) and (244).

(243)   What did Joe eat?

         a.   He didn't eat spaghetti.

         b.   I don't know.

(244)   a.   ?What Joe ate was he didn't eat spaghetti.

         b.   #What Joe ate was I don't know.

It is possible to find counterweights that would be rather uncooperative answers in a real discourse,

but such examples have a very marked rhetorical function, and require the weight to be strongly

anaphoric. Consider the following example culled from television dialogue.

(245)   My value here is I have no value.[43]

In context, the speaker is trying to determine why she has been chosen for a particular task, concerned that she is being used; after failing to obtain an answer to her question from a supervisor, she settles on the proposition in (245).

Speakers can also provide answers that remove one proposition from the set of possible answers, as in (244a). Examples like these are quite common, especially in *that's x is y*, which can facilitate the negation using the *that's x* clause.

(246)   That's not why I chose it is because of the fat content.

These examples are a rough counterpart to the "mention-one" question—they mention one possible answer that is not the true one.

The speaker can also hedge by embedding the answer under a bridge verb, as in (247a), or deny knowledge of the answer, as in (247b).

(247)   a.   I think what we should do is I think we should get together the Heritage Foundation, CIS and several other institutions in Washington and we should look forward to having a conference perhaps in the fall and we should call that conference "Taboo Subjects in Immigration".[44]

         b.   You know what he should do is, I'm not really sure what.

Den Dikken et al. (2000) propose that indirect answers like these, where the counterweight clause includes more or different information than the question is seeking, are not possible; for them, this is another indication that the counterweight is a non-root clause.

---

[43]*The West Wing*. Television. 2000. NBC.
[44]`http://cis.org/PanelTranscripts/illusionary-allure-panel`

(248)   What John bought was I believe that he bought a book.   (den Dikken et al. 2000:49

       (23a))

I disagree with their empirical judgment, however, and find such examples quite productive in natural contexts and fully acceptable. Slightly less common, but still acceptable, are counterweights embedded under third person or quantificational subjects, (249).

(249)   What John bought was everybody knows that the bought a book.

The answer is still offered from the speaker's perspective.

4.7   Conclusion

The amalgam copular sentence type exhibits a different pattern of syntactic and semantic behavior from its counterpart, the double-DP specificational copular sentence or canonical pseudocleft. The canonical pseudocleft permits some types of A- and A′-movement, in contrast to the amalgam, which is completely "frozen". Similarly, the canonical sentence types can be embedded in non-finite contexts, while the amalgams cannot.

A series of surveys testing the form of the copula in the two sentence types yielded a sharp contrast. The canonical pseudocleft allows the copula to combine with a range of functional "baggage", including negation, modals, and auxiliaries; the amalgam copula resists such material. For some speakers, it can combine with epistemic modals, but in general, it must remain bare. The same pattern is observed in non-reversible amalgams, like *that's x is y*, question-answer amalgams, "free-*be*", and double-*is*.

Section 4.6 examined the weight expression in the canonical and amalgam sentence types, and found that they have different semantics. The canonical copula relates an intensional predicate

of individuals and an individual-type expression, while the amalgam copula relates an intensional predicate of propositions and a proposition-type expression. In short, the amalgam copular sentence is a type of question-answer clause. It is a typologically rare construction, where a root clause "answer" serves as a logical subject.

Now that all the salient properties of the amalgam sentence type have been catalogued, I turn to a syntactic analysis that unites all of these properties and distinguishes the amalgam from the canonical copular sentence.

Chapter 5

The structure of copular amalgams

5.1   The copular amalgam clause puzzle

Specificational copular amalgam sentences do not fit into the structural model of ordinary copular sentences. In this chapter, I provide amalgams with a model that captures their unusual properties, as well as their configurational similarities with canonical clause types. I propose that copular amalgam sentences are projected from the left periphery of the clause. This model is unconventional, since it does not make use of the projections of V or T.

As Chapter 4 showed, the copular amalgam sentence family, exemplified in (1), exhibits a range of peculiar properties when compared to its better-behaved counterpart, the canonical specificational copular sentence (2).

(1)   a.   What she needs is she needs a break.

       b.   She needs a break is what she needs.

       c.   That's what she needs is she needs a break.

       d.   You know what she needs is she needs a break.

       e.   The thing is is she needs a break.

       f.   I really think she could use that pretty soon is she could use a break.

(2) a. What she needs is a break.

  b. A break is what she needs.

Drawing on the methods discussed in Chapter 3, I examined acceptability contrasts between canonical and amalgam sentences in different syntactic environments. While both sentence types bear all the hallmarks of the specificational copular sentence, the syntactic behavior of the amalgam pseudocleft is severely restricted with respect to its counterpart.

The main finding of the experiments is that the copula in the canonical pseudocleft behaves like an ordinary verbal copula, while the copula in amalgams does not. The amalgam copula associates with a reduced set of functional elements ("baggage"): it cannot combine with negation, future *will*, or aspectual auxiliaries.

Further comparison of the two sentence types demonstrated that while the weight expression in canonical pseudoclefts is a free relative clause (interpreted as a predicate intension), the weight in an amalgam is an indirect or concealed question expression (interpreted as the intension of a predicate of propositions). The logical subject of the canonical pseudocleft is therefore an individual or a property, while the logical subject of an amalgam is a proposition. Putting the pieces together: the copular amalgam is a type of "question-answer clause", where the copula relates a propositional subject and predicate (Ross 1972; den Dikken et al. 2000; Schlenker 2003; Caponigro and Davidson 2011).

I propose a syntactic account of amalgams that takes seriously both their bisentential character and their similarities to ordinary copular sentences. The amalgam pseudocleft, like its canonical counterpart, is derived from a predicational small clause, but instead of merging with T or V, it is headed by Fin, the lowest head in the C-domain (assuming a split-CP model of the clause; e.g., Rizzi 1997, 2004; Cinque 1999). The non-reversible amalgams, likewise, consist entirely of left-

peripheral functional material.

The peculiar properties of amalgams result from the fact that their Fin does not take TP as its complement—it takes its propositional subject and predicate directly, as in (3).

(3)   Structure of copular amalgams

```
                        ForceP
                       /      \
              Op-ASSERT        Force′
                              /      \
                          Force       ...FinP
                                     /        \
                                   CP          Fin′
                               /      \       /     \
                         I want a break  Fin        CP
                                          |       /    \
                                          is   what I want
```
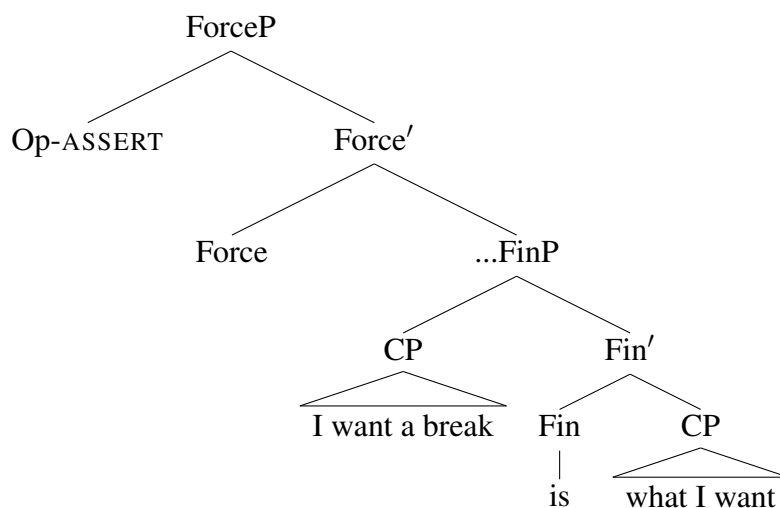
The absence of the lower T and V domains in the spine of the amalgam clause limits the functional material that can associate with the copula, and renders most movement operations unnecessary and impossible. The unusual structure in (3) also allows a bare, root-like clause to serve as a structural subject—the occupant of Spec,FinP. In canonical clauses where the structural subject position is a derived specifier (Spec,TP), the subject must be nominal.[1]

The proposal that the copula can be inserted directly into the left periphery is not new. Cross-linguistically, copular elements are readily co-opted by the grammar to signal focus and topic relations. Even the English verbal copula *is* has been argued to occur as a focus marker (e.g., Massam 1999, Massam 2013; McConvell 1988) and a topic-marker (den Dikken et al. 2000; O'Neill to appear). Den Dikken (2006) proposes that these functions form a natural class with that of the verbal copula: they all fall under the rubric of predication. The copula's status as a default element

---

[1] I defer detailed analysis of subject licensing in copular amalgams until Chapter 7.

allows it to occur in a range of syntactic environments (see, e.g., Bjorkman 2011), including the clausal left periphery. What is novel in the present proposal is that the copula in the left periphery is formally and functionally finite, although it never passes through a V or T position.

The function of the amalgam pseudocleft copula as a relator of propositional constituents lends itself to an analogy with "specifying coordination" structures (see, e.g., Koster 2000; de Vries 2006, 2009; Kluck 2011; Ott and de Vries to appear). The non-reversible amalgam types, in particular, where both the weight and the counterweight are syntactically root clauses, are configurationally similar to coordinate structures. Coordination, like predication, is an asymmetrical syntactic relation mediated by a simple functional head. The analogy between specificational and coordination is enhanced by the fact that in appositive modification structures, coordinators can have specificational semantics. In amalgams, unlike in ordinary sentential coordination, the copula serves as a *finite* sentential connective.

This chapter begins with an argument that the finite copula in amalgams is located in the C-domain of the clause, rather than in V or T. Next, section 5.3 illustrates the derivation of amalgam pseudoclefts, highlighting the structural similarities between amalgam and canonical copular sentences. Section 5.4 shows how variations in the C-domain of the clause can capture differences between the copular amalgam sentence types. In section 5.4.3, I discuss the functional and configurational similarities between copular amalgams and so-called "specifying coordination" structures. Lastly, 5.5 sets the scene for Part 3 of the dissertation, which examines the implications of this model for finiteness.

5.2 The finite copula in the left periphery

The amalgam copular clause is unusually restricted in its syntactic behavior. While the canonical copula is free to combine with a complete verbal functional complex, the amalgam copula's combinatorial possibilities are much more limited. This asymmetry between the two clause types finds a natural explanation if the copula in amalgams is inserted in the left periphery. I will argue in this section that the copula is not merely inserted in a high position—the copular amalgam clause is projected directly from the C-domain, rather than from V or T.

The first argument I present is that the copula is *spelled out* in a C-domain head position. Each layer of clause structure serves a dedicated set of abstract functions (Hale 1986; Chomsky 1986, 2001; Abney 1987; Grohmann 2003; Sigurðsson 2004; Borer 2005; Wiltschko 2014). Because the domain functions are fixed (supplied by Universal Grammar and independent of language-specific categories), they can be used to diagnose the syntactic position of language-specific functional elements like the copula (Wiltschko 2014). The copula is spelled out in the environment of C-domain material.

Not only is the copula spelled out in the C-domain; it *originates* in the C-domain in amalgams. Evidence for the copula's base position comes from its association with the abstract functions of the C-domain of the clause. Given its morphological features and position relative to other elements, I propose that it originates in Fin in amalgam pseudoclefts, but is spelled out in Top in *wh*-initial amalgam pseudoclefts and the other members of the amalgam copular sentence family.

Another component of this proposal raises more serious issues. I argue that in amalgams, the V and T-domains are not projected at all. This kind of clause, where a finite C-domain is projected independently of lower domains, is not predicted to be possible in a bottom-up model of the clause, where each layer of structure implies the presence of the lower ones (e.g., Grimshaw 1991). The

model of the copular amalgam clause that best accounts for its unusual properties requires that finiteness (encoded syntactically in Fin) be syntactically and semantically independent of Tense. The implications of this proposal are discussed in Part 3 (Chapters 6–8).

5.2.1 Domain functions as position diagnostics

It is a remarkable discovery in generative syntax that all languages include many of the same syntactic design features. Specifically, the same hierarchy of abstract structural domains underlies the nominal and verbal clauses cross-linguistically. Even languages with radically different surface morphosyntax and different inventories of functional categories show the hallmarks of the same underlying clausal architecture.

Since the 1980s, scholars have generally partitioned the clause into three, or sometimes four layers, representing these universal structural domains (Hale 1986; Chomsky 1986, 2008; Abney 1987; Pollock 1989; Grimshaw 1991; Bittner and Hale 1996a;Grohmann 2003; Sigurðsson 2004; Borer 2005;McFadden 2013; Ritter and Wiltschko 2014; Wiltschko 2014, among others). The four domains that make up the functional "spine" of the clause in Wiltschko's (2014) model are illustrated in Figure 5.1.
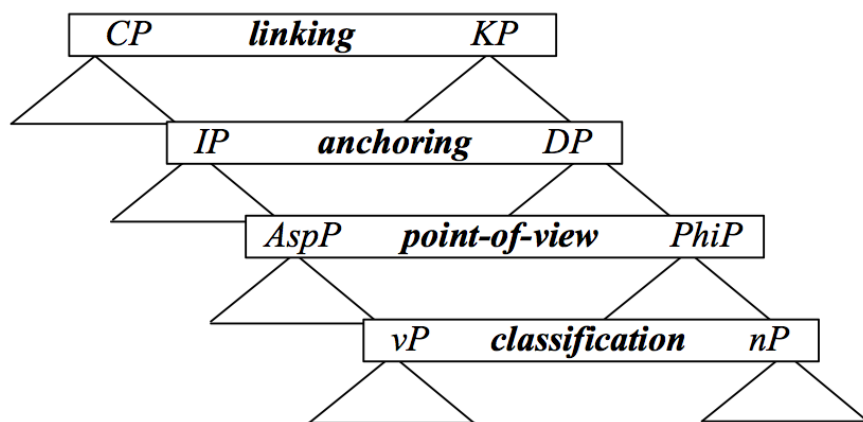
Figure 5.1: Universal hierarchy of functional domains (Wiltschko 2014:77)

The hierarchical arrangement of the structural domains is fixed. Expressions related to the core properties of the event or individual are situated low in the structure, while operators and discourse-related elements are situated high. I illustrate this clause model using examples drawn from English and Garifuna, an agglutinating Arawakan language with a typologically rare V(Aux)SO order.[2]

Despite the word-order difference between the two languages in declaratives, a subject question like (4) shows the same hierarchical domain order in Garifuna and English.

(4)

```
                        Linking
                       /       \
                    ka          Anchoring
                    who         /       \
                    Who     ba=sa        Point-of-view
                            PROS=Q       /          \
                            will      ñoun           Classification
                                   repeatedly             |
                                   repeatedly       abachar-ua-da
                                                    drunk-INCHO-VRLZ
                                                       get drunk?
```
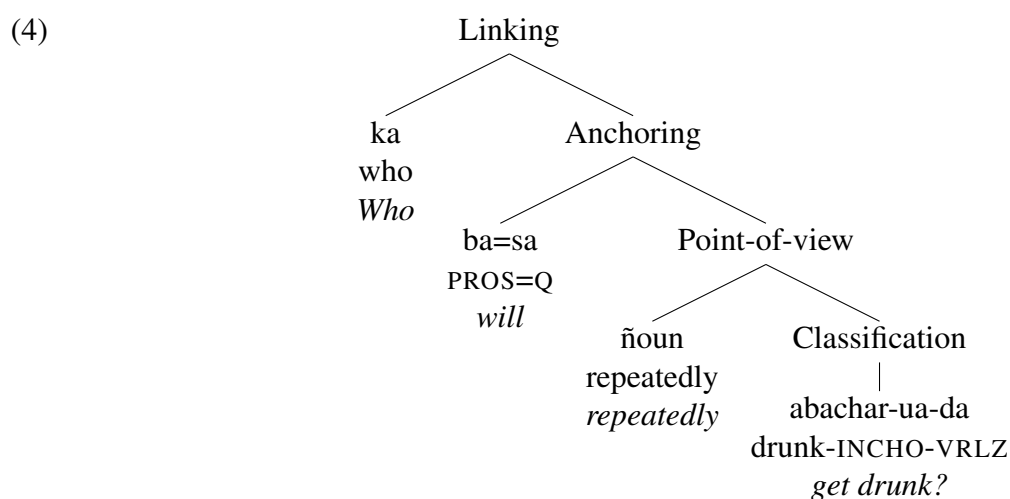
The operator and question-marker are introduced in the highest domain—the *wh*-operator *ka* occupies a peripheral, scope-taking position (the optional question marker *sa* is a second-position enclitic). The prospective marker *ba* is next. It anchors the proposition to the utterance context and hosts agreement markers in other contexts, like occupants of T/Infl in English. Immediately below this position is an aspectual modifier *ñoun* 'repeatedly'. Then, event-classifying valence morphology (*ua* 'INCHOATIVE') occurs immediately outside the verbalizing suffix *da*, which categorizes the root. This example shows that despite the typological differences between English and Garifuna, their clauses are structured by the same universal spine.

The labels of the domains in the structure above are functional—they do not refer to any par-

---

[2]Garifuna examples are from my own fieldwork, unless otherwise noted.

ticular semantic or syntactic category, but rather, to more abstract categories belonging to the language faculty, or perhaps to some other cognitive system (Wiltschko 2014). The domains of structure are categorially independent; that is, the elements that occupy them have different formal and semantic properties in different languages.

Given the robust cross-linguistic variation with respect to the semantic content and formal properties of functional elements, linguists face a nontrivial challenge in identifying functional categories. It is a matter for debate whether any categories are actually universal (Matthewson 2006; Haspelmath 2007; Déchaine and Tremblay, in progress).

Let us briefly consider the category Tense, since the question of its universality is relevant for copular amalgams. Following discussion in Ritter and Wiltschko (2014) and Wiltschko (2014), we might ask: do temporal markers always associate directly with an anchoring domain head T? A complementary question is: must the anchoring function of the clause be instantiated by an expression with tense content?

The answer to both questions, according to Bohnemeyer (2002), Shaer (2003), Bittner (2005), Tonhauser (2011), Ritter and Wiltschko (2014), Wiltschko (2014) and others, is "no". Temporal markers have different formal properties cross-linguistically (Bybee et al. 1994); similarly, the anchoring domain of the clause can have different content in different languages. In languages where temporal markers are grammaticalized and occur in the structural environment of Infl, they should be analyzed as instances of the category Tense. In other languages, however, some other category may occupy this structural position, and so the anchoring function is interpreted differently (for more detailed discussion of the diagnostics, see Shaer 2003, Ritter and Wiltschko 2014, and Wiltschko 2014).

In English canonical clauses, for instance, Tense can be identified as a head in the clausal

spine.[3] Temporal markers in English occupy a syntactic position that is just below the left periphery, (5a); they contrast with each other in a paradigm (including zero-forms), (5b); they are unique (one per clause), (5c); they are obligatory, (5d); and they are bleached of their ordinary interpretation in specific syntactic environments (e.g., counterfactuals), (5e). These five properties are exhibited by the past-tense markers in the sentences below.

(5)   a.   What **did** he do?

      b.   They like-**0** / lik**ed** coffee.

      c.   He is eat**ing** / *****eats** pizza.

      d.   Last year, he *****eat /ate** pizza.

      e.   If I **left** tomorrow, I wouldn't get there until evening.

In Garifuna, by contrast, past-tense markers are optional second-position clitics. They can co-occur with other finiteness markers; and they are always interpreted—there are no "fake" uses of these past markers.

(6)   a.   Haru-tu(=**buga**)   muna to.
           white-3.F.SG(=PAST) house DEF
           'The house was white.'

      b.   n-aliha   **ba**-i=**buga**,   pero m-aliha   n-umu-ti(=**buga**).
           1.SG-read FUT-3.SG.M=PAST, but   NEG-read 1.SG-NON.FUT-3.SG.M(=PAST)
           'I was going to read it, but I didn't.'

      c.   Afiñe-tina   gagumula-ha-ha-di-ti=(**meha**)=(**buga**)   (la).
           believe-1.SG smoke-VRBL-PERF-HABIT-3.SG.M=PAST=PAST COMP
           'I used to believe that he used to smoke.'

      d.   Ahein(=**meha**) hamuga ga-sani-tina...
           if(=PAST)   OPT   have-child-1.SG...
           'If I had (had) a child...'

---

[3]In Chapter 8, I will propose that the tense morpheme in T actually lacks semantic content; semantic tense is introduced by a time expression ZP (e.g., Zagona 1990) that is syntactically dependent on T. Here, I use the term "temporal marker" as a proxy for the T-head and its time argument.

Temporal markers in these two languages thus overlap in semantic content, but not in syntactic distribution. While temporal markers in English map to the clausal spine, they are adjuncts in Garifuna.

In fact, the counterpart of English Tense, with respect to hierarchical position and the properties illustrated above, is Aspect.[4] What do Tense in English and Aspect in Garifuna have in common? Both categories occupy the syntactic domain immediately below the left-peripheral scope-taking domain; both encode contrasts for finiteness: elements of their paradigms are obligatory in finite clauses, and blocked in certain non-finite clauses; and both are responsible for *displacement*: the relationship between the main event and the utterance time (see also Sigurðsson 2004). Aspect in Garifuna is the *anchoring* category, just as Tense is in English. Other languages which, like Garifuna, use Aspect to anchor, include Yucatec Maya (Bohnemeyer 2002), Guarani (Tonhauser 2011), Chol, and Tagalog (Coon 2013).

Miniature case studies like this comparison of English and Garifuna show that the universal functions of clausal domains can be "triangulated". It is impossible to distill the contribution of the abstract domains merely by analyzing the semantic and morphosyntactic properties of language-specific functional elements. This approach has an additional conceptual advantage of being able to model similarities between the nominal and verbal domains (e.g., Chomsky 1970; Abney 1987; Borer 2005), which despite morphological and semantic differences serve the same abstract functions (see Chapter 7 for further discussion). Careful attention to the relationship between a given functional element, like a temporal marker, and the abstract functions of the clausal domains provides a valuable methodological tool: it allows that element's syntactic position to be identified.

A complication enters the picture when we examine language-internal variation. What can we

---

[4]The prospective auxiliary *ba* in the example above appears in several different syntactic environments in Garifuna, and remains quite puzzling (Ekulona 2000). It is not clear whether it belongs to the same paradigm as the aspectual auxiliaries. In either case, Garifuna is an aspect-mood language, rather than a tense-aspect language.

learn from the distribution of an all-purpose functional element like the copula? Since the copula

is semantically bleached, it can occur in a variety of environments.

(7)   a.   She **was** a teacher.

    b.   He has **been** relieved of duty.

    c.   Don't **be** a fool.

According to Wiltschko (2014), "multi-functionality" is a typical property of categories that map

to the spine of the clause, since these elements realize different functions depending on which

domain they associate with.[5] The domain functions, as they are realized in English, can diag-

nose the syntactic position of the copula. In (7), the English verbal copula occurs in T supporting

tense/anchoring inflection, in the role of aspectual auxiliary, expressing point of view, and in the

verbal domain, where it introduces an event and an agent argument. When the copula is spelled out

in different domains, its function can reveal its base position.

5.2.2   The four structural domains

In the following subsections, I will briefly characterize the functions of each domain, with specific

attention to its function in copular sentences. I will use the abstract categories in the top row of Ta-

ble 5.1, but they have conceptual counterparts in other models. Occasionally, I will use traditional

category labels corresponding to the main element of each domain, as in the second row.

5.2.2.1   Classification

The bottom domain is dedicated to the classification of events and individuals: it introduces the lex-

ical core. In Sigurðsson's (2004) framework, this constitutes the event domain; similarly, Grohmann

---

[5]A classic example is the English function word *that*, which can associate either with D in the nominal domain or with T/Fin in the verbal domain. I take a closer look at the subordinating complementizer *that* in Chapter 7, which examines its role in the anchoring domain of the verbal and nominal spines.

	Context	Displacement	Point-of-View	Classification
	CP	TP	AspP	VP
Wiltschko (2014)	Linking	Anchoring	Point-of-View	Classification
Chomsky (2001, 2008)	CP	TP		v*P
Borer (2005)	E(vent)P	TP	AspP	VP
Sigurðsson (2004)	speech/Λ	grammatical/ϕ		event/θ
Grohmann (2003)	discourse/ω	agreement/ϕ		thematic/θ

Table 5.1: Labels of the clausal domains

(2003) refers to it as the θ-domain. V and N provide the lexical semantic content/restriction of the individual or event (e.g., telic vs. atelic; count vs. mass). This low layer of structure also introduces arguments, which further restrict the properties of the event or individual.[6]

Copular sentences present a special case. Although it is verbal, the copula itself does not participate in the classification function. It is the non-verbal predicate (N, A, or P) that determines the properties of the event. The non-verbal core of the copular sentence, a predicational small clause, encodes the functions of the classification domain. The lexical properties of the event are specified by the predicate.[7] The subject argument is introduced not by the copula, but by the head of the small clause (pace Rothstein 2001): the copula itself assigns no theta roles.

Consider the embedded small clause in (8).

(8)   I consider [John intelligent].

The small clause is maximally small, because it includes only the classification domain: none of the higher domains are projected in (8). The small clause introduces the core event (state) whose properties and participants are established. In its instantiation of the classification domain, the small clause in (8) does not differ from a finite copular clause, e.g., (9).

---

[6]In a Neo-Davidsonian event structure, the lexical properties of the verb and the arguments are all predicates of the event.

[7]I use the term "event" loosely here—later, I will follow Maienborn (2005) in treating the spatio-temporal entity in a copular clause as a state.

(9)   John is intelligent.

The verbal projection beginning with the copula lacks a distinct classification domain. Its function is purely grammatical.

5.2.2.2   Point of view

The middle field of the clause in many languages, English included, houses Aspect. AspectP syntactically mediates between the VP/vP (containing basic event and argument structure) and TP (containing its tense specification).

(10)   a.   She is singing.

        b.   She has sung.

        c.   She had sung.

Consider the functional and formal properties of Aspect.

The *function* of Aspect is to provide a point of view on the temporal structure of an event or situation. It mediates between the lexical and thematic properties of the event, and the event's location in time, by locating it with respect to a time that is under discussion (Topic Time, Klein 1994). Simplifying somewhat, in imperfective aspect, the event is described from within, while in perfective aspect, the event is described from the outside (e.g., Comrie 1976; Hale 1986; Verkuyl 1993; Klein 1994; Demirdache and Uribe-Etxebarria 2000, 2007; Arche 2006, and many others).

(11)   a.   I am writing the book.

        b.   I wrote the book.

The formal properties of temporal aspect vary cross-linguistically (see discussion in Wiltschko 2014). Nevertheless, languages that lack a counterpart of English Aspect do not fail to establish a

point of view on events—they merely use categories with different formal properties to do so, and map temporal aspect markers to different parts of the clausal spine.

In Blackfoot (Algonquian), for instance, temporal distinctions analogous to English aspects are conveyed by affixes that attach to the verbal stem at a hierarchical position that is below the level of "low" nominalization, below vP (Bliss et al. 2011). The expected position of (outer) Aspect—between vP and IP—is instead occupied by direct/inverse markers ("theme" markers in the Algonquianist literature; e.g., Bliss et al. 2011; Ritter and Wiltschko 2014), which indicates a person-based point of view on events: whether the subject is a first or second person participant in the discourse (internal point of view) or a non-participant/third person (external point of view). Wiltschko (2014) argues that it is point of view, rather than Aspect, that is the universal category, universally occupying the domain immediately above the vP and below the anchoring/displacement domain.[8] By abstracting away from category-specific properties, like the temporal content of Aspect in English, these apparently missing categories or exotic categories in the verbal spine of some languages become much less surprising.

In other frameworks (e.g., Grohmann 2003; Sigurðsson 2004), the viewpoint and displacement domains are conflated into one grammatical domain. Both domains encode grammatical relations: accusative arguments are displaced to the lower middle domain, near Aspect (or what was called AgrOP in Chomsky (1986)), while nominative arguments (and possibly absolutives) are displaced to the anchoring domain (to AgrSP, or TP). Grammatical relations are structurally distinguished, but not functionally distinguished. From a functional perspective, it is difficult to tease apart viewpoint and displacement: they tend to rely on the same semantic content; that is, a language that anchors to the utterance context using Tense will also have temporal viewpoint Aspect. Many lan-

---

[8]In the nominal domain, Wiltschko (2014) proposes that number marking provides the analogue of verbal viewpoint, establishing a point of view on the individual as either internal or external.

guages, like Garifuna, anchor events to the utterance context using aspectual categories, so it is not entirely clear that the viewpoint function can always be distinguished from the displacement function.

5.2.2.3   Displacement

The displacement domain serves to locate the central event or individual with respect to the utterance context.[9] I use the multipurpose term "displacement" intentionally. The displacement domain is the locus of the displacement property of natural language—it is what allows us to formulate propositions about events and individuals outside of our immediate experience. It is also the target of A-displacement/movement encoding canonical grammatical relations. The conceptual and syntactic functions of the grammatical anchoring domain both involve displacement.

In Sigurðsson's (2004) model, the function of the displacement domain is to indicate relations between participants in the event domain and participants in the speech act. For example, if the event agent is identified with the "logophoric" agent (the speaker), then first-person features are positively specified.

The most well-studied grammaticalization of displacement in the clause is Tense. Temporal displacement, familiar from finite clauses in most Indo-European languages, situates an event as overlapping the utterance context (present), preceding it (past), or following it (future). Not all languages anchor through tense, however (e.g., Bohnemeyer 2002; Bianchi 2003; Shaer 2003; Bittner 2005; Tonhauser 2011; Amritavalli 2013; Ritter and Wiltschko 2014). Languages can use times, locations, individuals, and (possible) worlds to locate eventualities with respect to the reference context.

I will occasionally describe events as "anchored to" a reference context, using Wiltschko's

---

[9]Displacement and deictic anchoring are discussed in detail in Chapters 7 and 8, so the present discussion remains somewhat superficial.

(2014) term, but the structural distinction between anchoring via displacement and anchoring via deixis will be central to my analysis of amalgam pseudoclefts.

5.2.2.4   Context

The context domain is the locus of information related to the linguistic and extralinguistic context of the proposition. Elements associated with the clausal left periphery include force and clause-typing markers, markers of information structure, evidentials, some modals, speaker and addressee-oriented adverbial modifiers (Giorgi 2010), possibly sentential connectives (Jayaseelan 2013), and finiteness features (see e.g., Rizzi 1997, 2004; Cinque 1999 for models of an exploded left periphery). Among these elements, the common denominator is the function of relating the proposition to a context (e.g., Rizzi 1997; Wiltschko 2014).

The function of the context domain in a bottom-up model can be conceptualized as follows. The lower domains encode basic event structure, point of view, and spatio-temporal location of the proposition relative to a context. What remains to be characterized is the context itself. Is the utterance an assertion, a question, or a command? What information does the speaker present as given and what is provided? Is the proposition being used by the speaker to carry out some speech act, or is someone else's speech act being reported? What is the speaker's attitude toward the proposition? Is the proposition evaluated with respect to the context of utterance, or with respect to an anaphoric context?

Both the context and displacement domains are integral to encoding finiteness in ordinary clauses. The relationship between the two domains is signaled by an Agree relationship between their T and [ɸ]-features (Rizzi 1997, 2004; Bianchi 2003; Adger 2007, etc.). Since finiteness ultimately reduces to the clause's ability to serve as an independent utterance, the context domain is the fundamental locus of finiteness, while the displacement domain allows for different specific

interpretations of that finiteness/non-finiteness.

The ability of a clause to serve as an independent utterance depends on two things: the presence of Force, whose specifier harbors illocutionary operators (e.g., ASSERT, Meinunger 2006), and the presence of a deictic variable in Spec,Fin that refers directly to the utterance context (Bianchi 2003). The latter allows the proposition to be evaluated independently of superordinate structure, and the latter establishes a relation between the proposition at the evaluation context and the speaker. Displacement from the context requires an additional predicate (see Chapter 8 for detailed discussion).

The nature of the syntactic dependency between the top two domains is a topic of some debate in Minimalist syntax. Since the displacement domain plays a mediating role in clausal architecture, the degree to which the two domains are functionally independent of each other is often unclear. Similarly, the context and displacement domains in the noun phrase are intimately related. In Wiltschko's (2014) approach, the function of the nominal context domain, abbreviated as KP in Figure 5.1, is to link nominal expressions to the linguistic or extralinguistic context. In dependent contexts, this function is instantiated by prepositional case-assigning elements, like *to* and *of*, which link DP to the larger structure in non-nominative (verbal anchoring domain) environments (see also Bittner and Hale 1996b). It is not coincidental that these prepositional elements can also occupy in the verbal context domain in non-finite clauses that are dependently anchored.

5.2.3   The copula in Fin

There are two components to the proposal that the copula in amalgams occupies the context domain of the clause. First, it is spelled out in a C-domain position. This can be established using its position relative to elements that are independently established occupy this domain. The second component of the proposal is that, unlike other copular clause types, the amalgam copula is base-

merged in a context domain position. It does not move into the left periphery in the course of the derivation, as it does in subject-auxiliary inversion. The base position of the copula can be determined by the domain functions it associates with, and also those it fails to associate with (which are discussed further in section 5.2.4).

It is unremarkable to find copular elements inserted directly in the context domain of the clause. For example, copulas in some languages distribute as topic and focus-markers.[10] (I will argue that the English copula is merged in Top in the double-*is* construction.)

(12)   Ka=sa   aliha ba-i   John?
        what=Q read   FOC/COP-3.SG.M John
        'What did John read?' / 'What is it that John read?'   Garifuna

(13)   Mwàlà làgá àn   kwàsà tsír   ní
        woman some COP/FOC eat   beans DEF
        '[A woman]$_F$ ate the beans.'   Bura (Hartmann and Zimmermann 2012: (3))

In these cases, the copular element in the left periphery co-occurs with a main verb, and does not realize the primary morphological finiteness features of the clause.

What sets the amalgam copula apart from other left-peripheral copulas is that it is *morphologically* and *functionally* finite. It associates directly with Fin, where it relates the proposition to the utterance context and to a Force operator. Encoding finiteness is the copula's "spinal function" in Wiltschko's terms.

5.2.3.1   Spell-out position of the copula

Chapter 4 showed that the amalgam copula's combinatorial possibilities are strictly limited. While the canonical copula combines with a range of functional material, the amalgam copula is almost always bare.

---

[10]In Suazo's (2002) pedagogical grammar, the extraction/focus marker *ba* is taught as the equivalent of *el verbo 'ser'* (the verb 'to be').

There was a notable exception: the amalgam copula can (marginally) combine with epistemic modals and speaker-oriented adverbs. For example, (14a) is judged more acceptable than the other examples in the set.

(14) a.   He drinks coffee **could possibly be** what he drinks.

b.   *He drinks coffee **must obligatorily be** what he drinks.

c.   *He drinks coffee **will eventually be** what he drinks.

d.   *He drinks coffee **isn't** what he drinks.

This apparent exception reveals that the structural position of the copula is higher than T. While the other functional elements associate with the structural position of T or lower, epistemic modals can map to a position above T (e.g., in the cartographic tradition). In the analyses of Condoravdi (2002) and Hacquard (2009, 2010), among others, epistemic modals are merged in a position above T, in the left periphery of the clause.

This structural position is suggested by the domain function that epistemic modals associate with: linking to context. Their modal perspective is anchored to the speaker indexical in the extra-linguistic context. The modal time—the time of the possibility—is always the local speaker's context time. In this respect, they contrast with lower "root" modals, which occur in the scope of Tense (and Aspect), and can have a past interpretation.[11] Since epistemic modals associate with contextual/deictic anchoring, rather than displacement, I take them to occur in the context domain, i.e., the left periphery.

In addition, speaker-oriented datives (15) and declarative slifts (16) are also relatively accept-

---

[11]Demirdache and Uribe-Etxebarria (2014) argue against this account, presenting evidence that epistemic modals can occur in the scope of tense and aspect. They account for these data by proposing head-movement (raising and lowering) of tense and aspect expressions to reverse scope, rather than allowing modals to map to two different positions. Amalgams provide evidence that modals on the speaker-anchored epistemic construal can occur in a structure that lacks the T-domain entirely. If indeed modals only merge in one position, and interactions with tense and aspect are derived by scope-reversing movements, then perhaps the real base position of modals is the higher one.

able in amalgams, in contrast to third-person epistemic modifiers. The fact that epistemic modifi-

cation is limited to first person shows that baggage that is interpreted with respect to the context

domain is licit in amalgams, while baggage interpreted with respect to the displacement domain

is not (e.g., for Sigurðsson 2004, the grammatical domain must be projected to specify that the

experiencer participant is negatively identified with the speech act participant).

(15)   a.   He made a new friend seems to me to be what he did.[12]

         b.   What happened seems to me to be he broke up with his girlfriend.

(16)   a.   He made a new friend (I think) is (I think) what he did.

         b.   What happened (I think) is (I think) he broke up with his girlfriend.

(17)   *He made a new friend seems to her to be what he did.

(17)   ?He made a new friend she thinks is what he did.

The position of the amalgam copula can identified by association: since the only licit baggage

on the copula consists of context domain material, the copula itself occupies the context domain.[13]

5.2.3.2   Domain function of the copula

I have shown that copula is able to associate with functional material from the context domain, but

it does not follow that the copula originates in this domain. It is unremarkable to find instances of

T in C, for instance, in questions. What sets the amalgam copula apart from these cases is that it

associates with the functions of the context domain—and only the context domain.

First of all, it associates with information structure marking. All copular amalgam sentences

---

[12]Recall that raising in copular amalgams is rejected by many speakers. Nevertheless, the speakers I surveyed informally found these cases marginal, reporting a clear contrast between the first person and third person experiencer dative. I discuss the syntax of such examples further in section 5.3.3.2.

[13]I will consider the structural and morphological status of non-finite *be* in combination with modals further in section 5.3.3.2.

have a fixed information structure: the counterweight is always a focus and the weight is always a topic (recall discussion in Chapter 2).

(18)   a.   [I need A BREAK]$_F$ is what I need.

      b.   What I need is [I need A BREAK]$_F$.

      c.   That's what I need is [I need A BREAK]$_F$.

      d.   You know what I need? is [I need A BREAK]$_F$.

      e.   I'm hoping we can talk about what's bothering me these days, is [I need A BREAK]$_F$.

Although specificational copular sentences in general are marked by this fixed information structure, value-initial specificational sentences have a predicational counterpart where the value-subject is the topic and the predicate is focused:

(19)   a.   [A break]$_F$ is what I need.

      b.   A break is [what I need]$_F$ (, not just what I want).

The amalgam pseudocleft has no such counterpart.

(20)   *I need a break is [what I need]$_F$ (, not just what I want).

The syntactic environment of the amalgam copula is associated with a fixed information function, owing to its location in the context domain.

It is not unusual for a copular element to mark focus and topic relations, as we saw in Chapter 2. In these cases, however, the copular form either originates in a lower domain, or it is not the primary bearer of finiteness. The copula in amalgams is distinct from these other left-peripheral copulas in that it originates in the left periphery, and it can associate with the two basic components of finiteness that are ordinarily marked on a verb: anchoring and force.

The copula's morphology indicates that it is finite: it is inflected for [tense] and [ɸ]. Finite inflection indicates that the proposition is anchored a (local) context. The [tense] and [ɸ] specifications of the clause are marked fully on verbal elements, but rudimentary finiteness specifications can also be found on complementizer elements in many languages. In Irish, for example, the finite complementizer shows a simple [tense] distinction, (21).

(21)   a.   Deir   sé go dtógfaidh sé an peann.
              say.PRES he that take.FUT he the pen
              'He says that he will take the pen.'

        b.   Deir   sé gur   thóg   sé an peann.
              say.PRES he that.PST take.PST he the pen
              'He says that he took the pen.'   Irish (Cottell 1995, cited in Adger 2007)

The same alternation in the amalgam copula is thus consistent with its status as a finiteness-marking element of the C-domain.

In amalgam pseudoclefts, although the copula resembles an agreeing complementizer in its inflection and position, it associates with assertive force, like a verb. Force is the quintessential context domain function, as it establishes the relationship between the proposition, the speaker, and the context of use. In a dynamic semantics framework (e.g., Krifka 2014), Force actually has an effect on the context.

Recall from Chapter 2 that the copula contributes an assertion in specificational sentences. In Lambrecht's (2001) view, this assertion is the main focus—the specificational relationship between the subject and predicate expressed by the copula is the new information conveyed by the sentence. A subordinating complementizer typically has the opposite role: it indicates that the proposition it introduces is presupposed. Although both can associate with formal finiteness features, unlike the copula, the complementizer does not associate with assertive force. Full finiteness, a property of the copular amalgam clause, requires association with force.

This treatment of the specificational assertion as a focus is consistent with Meinunger's (2004, 2006) structural treatment of (declarative) illocutionary force as a focus-sensitive operator ASSERT in Spec,ForceP. This operator selects for a proposition, and binds a focus within it. In a specificational sentence, the associate of ASSERT is the relation expressed by the copula itself. Fin in amalgams suffices to support full finiteness, because its occupant, the copula, is formally finite, anchored to the utterance context, and bound by the Force operator.

The assertive component of the copula can be detected under focus. For example, when the copula is stressed, it is interpreted as bearing verum focus—its assertive component is contrastively focused (Klein 1998; see Lohnstein 2015 for a review of the literature on verum focus).[14]

(22)   a.   What she likes (really) IS she likes coffee.

      b.   They lived in Canada (really) WAS where they lived.

Although this type of verum focus in English is signaled by a pitch accent on the verbal element, it is in fact a context domain domain phenomenon, hence *do*-insertion in verbal clauses in English, (23). For comparison, note that the same meaning is supported by *sí*-insertion in Spanish, (24).[15]

(23)   He DID lose my keys.

(24)   Juan SÍ perdió   mis   llaves.
        John YES lose.3.SG.PRET my.PL keys
        'John DID lose my keys.'

The association of the copula with the context domain finds further support from the fact that

---

[14]Some speakers prefer (22b) to (22a); nevertheless, both orders allow illocutionary adverbs like *totally* and *really*.

(i)   a.   What she likes is totally/really she likes coffee.
      b.   She likes coffee is totally/really what she likes.

The role of verum focus in diagnosing Force will be taken up in section 5.4.

[15]In canonical (English) sentences, the assertive component and the temporal anchoring component are fused, since Fin is in a dependency with both T and Force. In Chapter 8, I will argue that the amalgam copula does not contribute temporal anchoring.

in the absence of a fully specified context domain, the copular amalgam clause is ungrammatical. Amalgams cannot be embedded in non-finite envirobments. Once again, they differ in this respect from their canonical counterparts (see overview of these facts in Chapter 2).

(25)   a.   *With what she liked being she liked coffee...

          b.   *I'd hate for what she liked to be she liked coffee.

(26)   a.   *With she liked coffee being what she liked...

          b.   *I'd hate for she liked coffee to be what she liked.

(27)   a.   With what she liked being coffee...

          b.   I'd hate for what she liked to be coffee.

Non-finite clauses lack a deictic context variable, so they are dependent for anchoring on superordinate structure. They are anchored by way of a relation between their anaphoric reference time (or the topic time in a truncated non-finite clause) and an event in a higher clause. Without a deictic context variable providing independent anchoring, the copular amalgam is ungrammatical: it cannot occur in a truncated non-finite structure. The copula must associate with the context domain.

This section has illustrated that the syntactic and functional properties of the copula in amalgams cluster with left-periphery phenomena; thus, I conclude that the copula in amalgam pseudo-clefts spells out the context head that normally bears inflectional features: Fin.

5.2.4   Evidence for the absence of functional structure below Fin

This section presents evidence that copular amalgams actually lack clausal structure below Fin. The null hypothesis, based on canonical clauses, is that the copula occupies the anchoring domain

(T).[16] In what follows, I will argue that the null hypothesis must be rejected.

5.2.4.1   The English displacement domain

In the split-INFL model I use in the present work, the displacement domain is rich, consisting of projections of tense, negation, prospective/future modality, deontic modality, and certain temporal adverbials.[17] Because I treat time expressions as arguments of a temporal displacement predicate T (see Chapter 8), I must also assume a separate projection above T to accommodate the nominal subject. Let us call it Subj, after Rizzi and Shlonsky (2007), and assume that Subj and T are fused in the course of the derivation (Halle and Marantz 1993).

In ordinary finite clauses, independent anchoring is signaled by the presence of tense in the displacement domain, associated with a verb or auxiliary.

(28)   a.   She bake**d** a cake.

         b.   They ha**d** written a letter.

While English lacks overt V-movement for lexical verbs (so inflection lowers from T onto the verb, rather than the verb raising to T), the auxiliaries *be*, *have*, and *do* can occupy T, as can the copula *be*.[18] The traditional evidence for this is their position with respect to adverbs and negation, and their ability to move to C in questions, in contrast to lexical verbs (e.g., Pollock 1989). The canonical copula, as an occupant of T or V, can co-occur with other anchoring domain elements (29), as well as elements from the point of view (Aspect) and classification (V) domains (modals, auxiliaries, and adverbs). Examples with negation are provided in (29) to illustrate the contrast.

---

[16]That the finite copula occupies T in ordinary English sentences is uncontroversial, even though its origin (in V or T) is not settled.

[17]I remain agnostic as to which adverbial categories project as part of the functional spine of the clause, as in Cinque (1999), and which are adjoined. The details are not relevant to the current proposal.

[18]For now, let us assume that modal auxiliaries like *must* behave similarly to agreeing auxiliaries in this respect. I will reconsider this view later.

(29)   a.   She **has** not eaten meat.

        b.   She **is** not eating meat.

        c.   She **does** not eat meat.

        d.   She **is** not hungry.

        e.   *She **eats** not meat.

The displacement domain is also responsible for subject agreement and the nominative case relation. Although English has an impoverished agreement inflection system by comparison to its Indo-European relatives, the inflection it has is fusional, indicating a close relationship between tense and agreement. Indeed, the distribution of nominative case-marked subjects depends on a conspiracy of tense and agreement (e.g., Chomsky 1981; Landau 2004).

(30)   a.   He wants some coffee.

        b.   *She wants he to drink some coffee.

        c.   *He to drink some coffee would surprise me.

        d.   *It seems he to enjoy coffee.

        e.   *She considers he a great teacher.

Nominative subject licensing is thus another formal property of the clause that is dependent on the displacement domain.

5.2.4.2   The amalgam copula and lower domain elements

Now consider the relationship of the amalgam copula to displacement domain functions.

5.2.4.2.1   [+Finite]

First, the amalgam copula bears tense and agreement inflection, features associated with both T and Fin.

(31)   a.   He wants some coffee is what he wants.

      b.   He wanted some coffee was what he wanted.

      c.   The main problems are he is broke and he can't find a job.

      d.   *He wants some coffee be what he wants.

      e.   *What he wants to be he wants some coffee.

While these facts are consistent with the copula in a T-position, they do not require it—they are also consistent with the copula in a Fin-position, since $[\phi, T]$-features originate in the left periphery (e.g., Chomsky 2001, 2008; Richards 2007).

5.2.4.2.2   Subject licensing

The second displacement domain function that apparently associates with the amalgam copula is its ability to license a grammatical subject (albeit an unusual one—the grammatical subject of the amalgam is often a root finite clause). The correlation between finiteness features and (nominative) subject licensing, like the features themselves, depends on both the T- and Fin-areas of the clause (Bianchi 2003; Sigurðsson 2004; Landau 2004; Adger 2007), so once again, the facts are consistent with situating the copula in either T or Fin. Furthermore, since the grammatical subject of the copular amalgam can be a root clause, which cannot occur in Spec,TP, the unusual type of grammatical subject licensing the copula participates in must obtain in a position other than T.

5.2.4.2.3   Negation

The acceptability surveys presented in the previous chapter tested the effect of functional elements from the displacement and point-of-view domains on the amalgam copula, in contrast to their effect on the canonical copula. For example, the amalgam copula cannot combine with sentential negation, unlike the canonical copula. A comparison between canonical and amalgam pseudoclefts

using a linear mixed effects model showed a significant interaction between the sentence type

and the effect of negation (p<0.001). Not only is negation in amalgams judged below a baseline

acceptability rating (derived from a normalized mid-point between acceptable and unacceptable),

it is judged as significantly more deleterious to amalgams than it is to canonical pseudoclefts. The

reported judgments are represented using traditional notation in (32) and (33):

(32)   a. *He wants some coffee isn't / is not what he wants.

        b. *What he wants isn't / is not he wants some coffee.

        c. *He doesn't want coffee isn't / is not what he wants.

        d. *What he doesn't want isn't / is not he wants some coffee.

(33)   a.   Coffee isn't / is not what he wants.

        b.   What he wants isn't / is not coffee.

This survey illustrated the same contrast for *that's x is y* sentences, which were additionally ma-

nipulated by negation on the copula in the *that's x* clause, to control for potential pragmatic effects

of negating the specification relation. In both conditions (with and without the second negation),

the sentences were judged as unacceptable—they were significantly worse than both the baseline

and the canonical copular sentence condition. The following attested examples illustrate negation

in the precopular clause.

(34)   a.   That's not why I wanted spinach and artichoke dip, is / *isn't because of the fat

             content.

        b.   That's not a way to convince me, is / *isn't by saying that your fourth-graders listen

             to it.

Importantly, this acceptability contrast is syntactic: the amalgam copular sentence structure lacks the position of negation. Non-syntactic negation, where the copula is bare and the proposition is negated by a different expression or in the context, is acceptable in amalgams, (35). It is also acceptable to negate only the counterweight, especially in a *wh*-initial amalgam pseudocleft, providing what in a question-answer discourse would be an uncooperative answer, (36). ((36b) is acceptable if the weight is strongly anaphoric.)

(35)   a.   It's not true that what he wants is he wants coffee.

       b.   He quit his job is what he did! It totally isn't... I'm just kidding.

(36)   a.   What he did is, he didn't quit his job.

       b.   She doesn't want coffee is what she wants.

Syntactic negation remains unacceptable in amalgams, even when its function is metalinguistic (non-truth-conditional sentential negation; Horn 1985) in contrast to canonical specificational sentences.

(37)   a.   *What she is isn't / is not she's happy—she's overjoyed.

       b.   *She found the answer isn't / is not what she managed to do—she had no trouble at all.

       c.   *What they're afraid of isn't / is not they're afraid of a nuCUlar disaster—they're afraid of a nuclear disaster.

(38)   a.   What she is isn't / is not happy—she's overjoyed!

       b.   Find the answer isn't / is not what she managed to—she had no trouble at all.

       c.   What they're afraid of isn't / is not a nuCUlar disaster—they're afraid of a nuclear disaster.

These examples show that the incompatibility of the amalgam copula with negation (as in (32)) is syntactic, not semantic: syntactic negation is not compatible with the copula of an amalgam.

5.2.4.2.4 Future *will*

Another element of the displacement domain that is closely tied to its temporal anchoring function is the future marker *will*. Chapter 4 presented experimental evidence that while the canonical copula is compatible with future *will* (particularly when the weight clause also includes an occurrence of *will*), the amalgam copula is not. Just as with negation, the future marker significantly decreases the absolute acceptability of the amalgam pseudocleft, with a normalized acceptability ratings concentrated below the baseline. Future marking also has a significantly *worse* effect in amalgams than in canonical pseudoclefts. Once again, representative examples are reproduced with traditional grammaticality marking in (39) and (40).

(39)   a. *What they will do will be they will quit.

        b. *They will quit will be what they will do.

        c. *What they do will be they quit.

        d. *They quit will be what they do.

(40)   a. What they will drink will be coffee.

        b. Coffee will be what they will drink.

        c. What they drink will be coffee.

        d. Coffee will be what they drink.

The acceptability of the sentences in (40) shows that there is neither a syntactic nor a semantic problem with future in a specificational sentence.[19] The unacceptability of (39) therefore supports

---

[19]Additional informal tests using sentences with connectivity effects confirmed that the judgments hold for unambiguously specificational sentences. Not all of the experimental items included connectivity effects, to allow for

the present analysis, where the clause structure of the amalgam simply lacks the position for future marking.

5.2.4.2.5 Deontic modals

Similarly, the amalgam copula cannot combine with deontic modals (or their adverbial modifiers), in contrast to its canonical counterpart:

(41)   a. *What they must avoid must (necessarily) be they must avoid corruption.

        b. *They must avoid corruption must (obligatorily) be what they must avoid.

        c. *What they avoid must (legally) be they avoid corruption.

        d. *They avoid corruption must (legitimately) be what they avoid.

(42)   a.   What they must avoid must (necessarily) be corruption.

        b. ?Corruption must (legally) be what they must avoid.

        c.   What they must avoid must (legally) be corruption.

        d. ?Corruption may (legitimately) be what they avoid.

The challenge in evaluating the effect of deontic modals in a survey, as Chapter 4 discussed, is that the English modals are ambiguous between the deontic and epistemic readings, so in an acceptability task, participants would likely parse the sentences with the more acceptable epistemic reading, and give higher acceptability ratings. Informal surveys, where I pronounced the sentences aloud and used contexts that privileged the deontic reading, yielded low acceptability ratings. A confound is that the obligation associated with a deontic modal is associated with the subject—when the grammatical subject is an inanimate value, the deontic interpretation is less natural. Nevertheless, my informants find the contrast between, e.g., (41a) and (42a) to be robust, so I take this as

_____

sufficient lexical variation among items.

further evidence that the structural position for deontic modals is absent in amalgams, while it is present in canonicals, although it is not always plausible or appropriate to fill it.

Indirect corroboration of this result is available from the survey of the effect of modals and auxiliaries on the second copula of double-*is* sentences, which compared them with *wh*-initial amalgam pseudoclefts as a baseline. While a statistical comparison of amalgam and canonical pseudoclefts for this feature is not available, modals were found to significantly lower the normalized acceptability score of pseudoclefts, when compared with the bare copula version.

5.2.4.2.6 Aspectual auxiliaries

Functional elements belonging to the point of view domain similarly combine only with the canonical copula, and not with the amalgam copula. Such elements include, most notably, aspectual auxiliaries and adverbs:

(43) a. *What they have (always) done has always been they have (always) cried until they feel better.

b. *They have (always) cried until they feel better has (always) been what they have (always) done.

c. *What they do has (always) been they cry until they feel better

d. *They cry until they feel better has (always) been what they do.

(44) a. What they have (always) done has (always) been (to) cry until they feel better.

b. Cry until they feel better has (always) been what they have (always) done.

c. What they do has (always) been (to) cry until they feel better.

d. Cry until they feel better has (always) been what they do.

The discussion of the experiment in the previous chapter mentioned that aspectual auxiliaries had a negative effect on the acceptability of canonical pseudoclefts, which has sometimes been reported as categorical in the previous literature. The effect of aspectual auxiliaries in amalgams has this categorical character, while the relatively milder negative effect on canonical pseudoclefts is likely due to the meaning of perfective aspect. Without a supporting context to establish a relevant distinction between the event time and topic time of the copular predication, perfective is somewhat odd in a specificational sentence. Consider, for example, the effect of context on the following non-pseudocleft example:

(45)　?#The teacher has been John.

(46)　　Ever since I've been here... / For many years...

　　　the teacher has been John.

In informal surveys, which allow for more experimentation with context, respondents consistently preferred the bare copula in amalgams, while accepting the auxiliary in canonical pseudoclefts. Since the amalgams were not ameliorated by context, I take the results above as evidence that the aspect domain is absent from the structure of the amalgam clause.

5.2.4.2.7　Subject-auxiliary inversion

Additional syntactic evidence that the amalgam copula does not originate in a lower domain position comes from the fact that it cannot undergo inversion. If the copula were in T, for example, we would expect polar questions with inversion to be grammatical, contrary to fact.

(47)　　a.　*Is what she did she left early?

　　　b.　*Is she left early what she did?

　　　c.　*What what she drank she drank coffee?

     d.   *Was she drank coffee what she drank?

     e.   *Is that's the problem she quit her job?

Canonical pseudoclefts with subject-auxiliary inversion, on the other hand, are fully acceptable.

(48)   a.   Is what John is important to himself?

        b.   Is important to himself what John is?

        c.   Is what she likes coffee?

        d.   Is coffee what she likes?

These facts are perfectly compatible with the structure of the canonical pseudocleft, where the declarative version features either the value or fronted weight clause in the structural subject position, and the copula in T. In a polar question, the copula simply inverts to Fin (or some position in the C-domain endowed with a [Q]-feature). Under the present analysis of the amalgam pseudocleft, where the copula is base-generated in Fin, there is no need for it to move—it is already in the highest domain.

    For similar reasons, counterfactual inversion is acceptable in canonical but not amalgam pseudoclefts:

(49)   Had what she liked been coffee...

(50)   *Had what she liked been she liked coffee...

5.2.4.2.8   Taking stock

The contrasts between the canonical and amalgam copula with respect to their interactions with the anchoring and point-of view domain functions are robust. The canonical copula, as expected, can occupy V when a higher auxiliary occupies T, or it can occupy T directly, preceding sentential

negation and a variety of adverbial modifiers. The amalgam copula cannot combine with these elements: although it alternates in [T] and [φ]-features, it is generally bare. Since the amalgam copula cannot combine with material independently shown to occupy the anchoring and point of view domains, the copula itself does not occupy those domains either.

5.2.4.3   The amalgam copula and lower domain functions

Further evidence that the copula does not occupy the T-domain comes from its lack of interaction with the core function of displacement. Although the copula bears [T] and [φ] inflection, this inflection does not encode the same functions associated with an element originating in the T-domain; that is, the copula's [T]-feature is "fake" tense and its [φ] does not facilitate ordinary nominative subject licensing. In Chapter 8, I will argue that while the copula in T associates with the displacement function in canonical specificational sentences, anchoring in amalgams is non-temporal, relying instead on deixis to the utterance context.

Preliminary evidence offered here is that temporal adverbials are impossible in amalgam pseudoclefts. The data here are drawn from informal surveys, rather than formal surveys.

(51)   a.   *Yesterday, what she liked was she liked coffee.

       b.   Yesterday, what she liked was coffee.

Context helps to make the contrast clear, because just as with aspectual modification, it is pragmatically odd to restrict the Topic Time of a specification relation without context. The adverb in this test must be placed in a position where it cannot be construed with the counterweight clause. This confound is difficult to control for, because temporal adverbs tend to be pronounced in peripheral positions.

(52)   It's hard to keep track of what kind of phone he has. He's always getting new ones. As far

as I know...

    a.   Yesterday, what he had was an iPhone.

    b. ??Yesterday, what he had was he had an iPhone.

The atemporality of the copular clause itself is also demonstrated in the following contrast. (53b) can only be salvaged if *not anymore* is coerced to modify *he had an iPhone*.

(53)   a.   Before, an iPhone was what he had, but not anymore / it isn't anymore.

           b.   *Before, he had an iPhone was what he had, but not anymore / it isn't anymore.

The absence of the displacement function, taken together with the unacceptability of the amalgam copula with functional material from the middle structural domains, show that these domains are in fact absent from the structure of the copular amalgam clause.

5.3   Deriving the amalgam pseudocleft

Despite the major difference in the functional structure of the canonical and amalgam copular sentence types, their basic configuration is the same. Both sentence types feature a simple predication relation between a subject and (non-verbal) predicate. In Chapter 2, I introduced den Dikken's (2006) model of predication, where all instances of predication involve the projection of a functional category that accommodates the subject and predicate in its minimal domain. In copular sentences, this functional category is ultimately spelled out by the copula itself. The following tree, based on den Dikken (2006), illustrates the small clause underlier of copular sentences.

(54)

```
                              RP
                    ┌─────────┴─────────┐
                 Subject               R′
              ┌────┴────┐         ┌────┴────┐
      Counterweight/Value       R⁰      Predicate
                                        ┌────┴────┐
                                     Weight/Variable
```
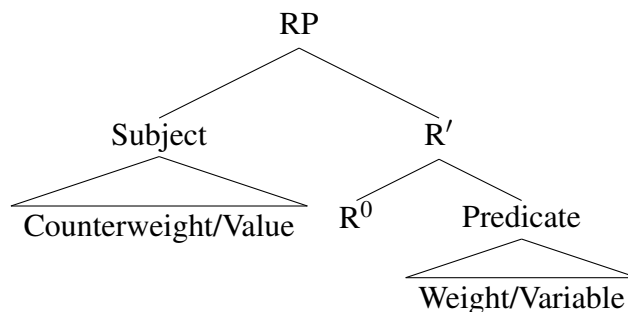
The Relator is a meta-category: any functional head mediating a predication relation is a Relator. Functional projections low in the V-domain, like *v*, for example, relate a tenseless verbal predicate and an agentive subject. T relates a property of times and a time argument (see Chapter 8 for a structural analysis of Tense/tense). In ordinary finite clauses, Fin relates a property of a context and a context (represented as a Reference Time in a neo-Reichenbachian framework). Top, similarly, relates a subject and the predicate, although the logical relationship between a topic and comment expression is more flexible than the other relations listed above, one of "aboutness" rather than set membership. Not coincidentally, the copula can be found in all of these relator positions.

Recall that the central question of this dissertation is: what is the *minimum* amount of structure required to form a finite clause? Copular sentences provide a useful testing ground for potential answers to the question, because, by definition, their clausal spine is stripped of lexical material. Taking the small clause RP as a base, I will show that the copular amalgam sentence instantiates the minimal finite clause. While formal properties of nominal subjects require the projection of intermediate layers of structure between the domain of Finiteness and the RP in canonical clauses, these needs are obviated in the amalgam, where the RP itself is a projection of Fin.

5.3.1 The classification domain in copular sentences

In copular sentences, the RP of a copular clause plays the same role as a VP in a non-copular clause type; it instantiates the classification domain. Copular sentences include an eventuality-like entity

whose lexical properties and participants are specified. Although the R head itself is not a theta-role assigner, the RP domain, like the VP/vP domain of ordinary clauses, serves to introduce an event/state participant into the structure. The small clause thus satisfies the functional description of the classification domain.

The Relator Phrase includes a *state*. For example, the small clause [*John fast*] introduces a state that is characterized by John having the *fast* property (Maienborn 2005). Even the totally verbless embedded small clause includes such an entity whose temporal location can be specified.[20]

(55)   We're much more likely to win the game, with [John so fast] lately.

The possibility of temporally locating and modifying the small clause predication proves that the projection of a classification domain does not require the verbal copula itself. Another way of putting it is that the primary occupant of the classification region of the clause need not be a lexical V. The association of both verbal and non-verbal predicates with the classification domain provides an argument against the V-requirement, which states that all clauses must include the projection of a lexical verb. Déchaine (1995), Rothstein (1999), Schütze (2004), Coon (2013), and others treat the predicational copula as a lexical verb (an element of category V), while I argue, following Heycock and Kroch (1999), Doron (1986), Déprez (2003), Mikkelsen (2005), den Dikken (2006), Benmamoun (2008) and others, that the copula is functional. The function of the lexical verb in ordinary verbal sentences—to restrict an event and introduce selectional requirements for arguments—is carried out instead by the non-verbal predicate.

If there is something like a semantic V-requirement, it does not privilege the lexical category V,

---

[20]Additional examples are discussed in Schütze (1997:279 (7)):

(i)   a.   With John out of town last week, we'll have a lot to catch up on next week.
        b.   With John out of town next week, we had a lot to prepare last week.
        c.   With John out of town this week, we have a lot to do today.

but rather, minimal domains that introduce states and events. If there is a syntactic V-requirement, it would entail that the V in copular sentences associates with a higher domain, e.g., displacement. In this case, it would seem to be a T-requirement, rather than a V-requirement. I conclude that the classification domain in copular sentences is the small clause itself, while the verbal copula is required, for formal reasons, in T.

5.3.2   The displacement domain in copular sentences

The small clause underlier of a canonical copular sentence consists of only the classification domain: it establishes no point of view on the eventuality or anchor to the utterance context. Because of its limited functional structure, the bare RP small clause is clearly non-finite—it cannot be an independent utterance. Since it is non-finite, it also fails to anchor a dependent (nominal) subject argument. The extremely reduced functional spine of the RP thus depends on external structure. This is where the verbal copula enters the picture.

In canonical copular sentences, RP merges with a head in the displacement domain, which introduces the structure that mediates argument licensing and anchoring to the utterance context. In an ordinary English clause, this head is T.[21] T introduces a valued [tense]-feature and a time argument into the structure, which allows for temporal displacement. This region of the T-domain is also responsible for argument licensing. Subj, a head immediately above T, inherits probing [ɸ]-features from Fin in the left periphery. Its specifier is the target of EPP-driven A-movement of the structural subject. The projections of Subj and T simply mediate between the left periphery and the classification domain. Because these projections are required to satisfy the needs of higher heads and the needs of the nominal subject, ordinary copular sentences exhibit the full range of displacement domain effects, in addition to classification domain effects.

---

[21]In copular sentences, imperfective aspect is the default, and no extra projection for aspect is needed, but RP can also merge with Asp in perfective contexts.

Curiously, in amalgams, as I demonstrated in section 5.2.4, anchoring domain elements are absent, but the classification functions are realized. The copular amalgam sentence includes an eventuality of sorts, and specifies its properties and participants—there is some propositional content to the clause, after all. This eventuality is not an ordinary Davidsonian event, however, because it cannot undergo spatio-temporal modification. Consider the examples in (56), where locative and temporal modification fails, in contrast to a sentence with an ordinary event (57). (Note that (56) is marked ungrammatical on the reading where *in her office* and *suddenly* modify the predication relation between *what she drank* and *she drank coffee*. It is of course acceptable on the reading where *in her office* modifies the *drank*-clause.)

(56)   a.   *In her office, what she drank is/was she drank coffee.

     b.   *Suddenly, what she drank is/was she drank coffee.

(57)   a.   In her office, she drinks coffee.

     b.   Suddenly, she drank coffee.

Instead, the object whose properties and participants are specified in (56) patterns with Asher's (2000) "facts", rather than with eventualities. Fact are world-dependent objects, while eventualities are concrete slices of actual space-time. While I leave this notion somewhat informal for the present discussion, a fact is similar to an intensional proposition denotation—it is a property of a world. Recall the conclusion of Chapter 4: copular amalgam clauses are question-answer pairs. Rather than asserting that a property holds of some *individual*, they assert that a property holds of some *proposition*. The fact of *she drinks coffee* being *what she drinks* could be compared to the fact of *two plus two* being *four*.[22]

---

[22]Of course, we assume that the mathematical fact holds of all worlds, while her drink choices may vary.

(58)   a.   *Two plus two is four in her office.

     b.   *Suddenly, two plus two was four.

Facts cannot easily undergo temporal modification, but they do participate in anaphora.[23]

(59)   [He really let me down is why I fired him]$_i$, and you know it$_i$.

(60)   [Two plus two is four]$_i$, and it$_i$ blew this four-year-old's mind.

There is no reason to treat fact-type objects as exceptional in the context of the functional model of the clause. Although a fact is not an eventuality, it is subject to the same kind of classification: its properties and participants are specified in a small clause (the participant is the speech act associated with the counterweight clause). The amalgam RP introduces a fact and its properties, so the RP of an amalgam instantiates the classification domain of structure.

Since the classification domain of canonical sentences introduces an eventuality where an individual bears some property at that eventuality, while the classification domain of amalgams introduces a fact where a sentence bears some property, it is not surprising from a semantic perspective that point-of-view and anchoring domain modifiers should be inappropriate in amalgams. The incompatibility between the fact in amalgams and these modifiers is also syntactic. Amalgams contrast in this regard with other sentences expressing properties of facts, because copular amalgams completely lack the displacement domain of the clause. Facts in sentences like (58) must include the displacement domain for formal reasons; TP is projected to license their nominal subjects and check the EPP-specified probes in their left peripheries. Since the grammatical domain, headed by

---

[23]Importantly, facts expressed by syntactic structures including a full Tense-domain can be asserted to hold of particular Topic Time intervals: there must be some process by which a fact-type object is evaluated with respect to an individual's beliefs about a particular interval of time, because they can be tensed, and their temporal interpretation can be anaphoric in context. See (61), below. What is relevant in (58) is that out-of-the-blue temporal modification of a fact is nonsensical. I will show in Chapter 8 that in amalgams, which lack T-domains, even contextually facilitated temporal restriction fails.

Tense, is projected, there is a time argument present in such sentences, which can be detected in contexts where temporal modification is pragmatically coerced.

(61)   Although he was ultimately brainwashed to believe that two plus two was five, **the whole time he was in the torture chamber**, two plus two **had of course been** four.

Although *two plus two is four* is not time-dependent, the Topic Time for which the assertion holds is restricted to *the whole time he was in the torture chamber*.

In sum, the copular amalgam is unusual because the functions of classification, anchoring, and deixis to the context are collapsed into a single small clause domain. While in the canonical copular sentence, syntactic constraints motivate the projection of separate domains for event classification and anchoring/grammatical relations, in the amalgam type, only one structural domain is needed. Because the impetus for separating the RP from the grammatical domain is lacking in amalgams, I take the copula in the amalgam to instantiate R directly. There is no need for A-movement of the subject in copular amalgams. The logical and grammatical subject is generated in the specifier position of the RP, which places it in a spec-head relationship with Fin, the bearer of probing clausal features. Base-generation of the sentential subject is in fact expected: bare sentences do not generally undergo A-movement (e.g., Webelhuth 1992; Moulton 2013; see Chapter 7 for more discussion). The context domain head Fin takes the clause's primary subject and predicate in its own local projection. We now have evidence supporting the proposal that the copular amalgam is a uniquely "minimal" type of clause.

5.3.3   Copula insertion

Before implementing the RP model of amalgams for the different sentence types, I will address the question of why the amalgam pseudocleft, like standard copular sentences, requires a verbal

copula at all. This is a puzzle, because as I argued in previous sections, the obligatory presence of the copula does not seem to follow from either semantics or syntax. Indeed, the syntactic V-requirement is directly at odds with the analysis of clause structure in this thesis, and with the analysis of the copula as a functional element. Higher domains of clausal structure do not imply lower ones. In other words, a finite clause does not need a Verb in its spine.

Nevertheless, the counterweight-initial amalgam does require an overt copula, and there is clearly something verbal about its function. The copula in the amalgam pseudocleft marks the clause as finite. It inflects like a verb for tense and agreement. It also expresses the assertive force of the proposition; when stressed, it projects verum focus alternatives (62).

(62)   He needed a break WAS$_F$ what he needed.

When the copula is omitted in the counterweight-initial amalgam pseudocleft, the ungrammaticality is on a par with copula omission in canonical clauses, for English speakers who do not allow null copulas.[24]

(63)   a. *She went to the store what she did.

         b. *She crazy.

The weight-initial members of the amalgam family do permit the copula to be absent, but the colon intonation and semantic connection between the two clauses remains.[25]

(64)   a. That's my issue (is) I need more time.

         b. Guess what he did (is) he threw the plate across the room.

         c. I'd consider that to be our main problem (is) we don't know what we're looking for.

---

[24]Further research is needed to explore the distribution of copular amalgam clauses in zero-copula English dialects.
[25]Chapter 4 showed that the copula may only be absent in a weight-initial pseudocleft if there is hanging topic intonation, and the weight is strongly discourse anaphoric.

    d.   What he did %(is) he threw the plate across the room.

I will discuss this pattern further in the next section, but for now, it suffices to illustrate that the "V-requirement" pattern in amalgams is not uniform. What is responsible for the obligatory presence of the copula, even when it is not an instantiation of T?

5.3.3.1   Against the V-requirement

I argued that there is no semantic requirement that a lexical Verb appear in every (non-small) clause (contra, e.g., Rothstein 1999). Rothstein's proposal is that a lexical Verb is required, because it is responsible for introducing the eventuality or state object that can be located by Tense; this locatability is a prerequisite for finite sentence-hood. Without a Verb, there is no locatable eventuality. She observes differences in the interpretation of temporal modification in bare vs. copular small clauses, and proposes that the copula's role is to introduce the eventuality. I showed that a locatable state can be introduced by a bare finite clause in the absence of the verbal copula, casting doubt on this proposal for canonical copular sentences. (65) is repeated from (55), above.

(65)   We're much more likely to win the game, with [John so fast] lately.

Since the presence of an eventuality in a copular sentence is independent of the verbal copula, and the copula can be present without a temporally located eventuality (in amalgams; see details in Chapter 8), Rothstein's (1999) analysis of the V-requirement in copular sentences cannot be maintained.

    I have also argued against a syntactic implementation of the V-requirement, based on den Dikken's (2006) analysis of predication structures as projections of functional heads. The syntactic V-requirement (e.g., Déchaine 1995; Schütze 2004; Cowper 2010) states that every clause must include a projection of a lexical Verb. For these authors, the requirement that V occur in the

environment of T is cashed out in terms of a selectional property of the higher functional head. In Baker (2003), the verbal copula in a language like English is similarly constrained by syntactic selection. T must attract a *lexical* element to host tense morphology. In English, verbal predicates and non-verbal predicates alike are generated as functional categories in a PredP (similar to an RP; see also Bowers 1993), where Pred is a verbalizer. If the predicate conflates with Pred, it is spelled out as a lexical verb, and can combine with Tense; non-verbal predicates do not conflate with Pred, so they remain functional, and thus cannot combine with Tense. The copula is inserted in this context, providing a licit lexical V host for tense.

An independent problem with treating copulas cross-linguistically as lexical V hosts for Tense is the existence of pronominal copulas. When inflection fails to combine with a lexical verb in languages like Hebrew and Arabic, a *nominal* category is inserted to support it, not a verbal category (Doron 1986; Benmamoun 2008). If the T/Infl domain of the clause is the source of the copula-requirement, then it is not cross-linguistically a lexical V requirement.

In fact, the English copula does not behave like a lexical Verb. It does not assign theta roles or case. When it is finite, it is spelled out in T, while lexical verbs remain in V. It can undergo T-to-C movement, unlike lexical verbs. An analysis of the copula that predicts its distribution in functional positions is preferable.

In light of the syntactic analysis of the copular amalgam sentence, a selection-based account of the V-requirement must be rejected. The proposals above all hinge on a close selectional relationship between the T- and V-domains of the clause, but these domains are completely absent from the amalgam copular sentence. Higher heads in the functional sequence do not necessarily select lower ones. If they did, a clause could not be finite without T(ense). Since amalgams lack the domain of Tense, but are functionally finite, T-V selection-based approach to the obligatoriness

of the copula will fail.

I have argued against both semantic and syntactic implementations of the V-requirement, which are inconsistent with copular sentences. If the copula is not required by semantics or syntax, then the next suspect is morphology. Such an account of auxiliary *be* is proposed by Bjorkman (2011), who argues that auxiliary *be* appears in the post-syntactic morphological component whenever inflectional features of heads in the clausal spine fail to combine with the main verb. Similar proposals are implemented for pronominal copulas in the analyses of Semitic mentioned above. Since the copula is the least specified verbal form, it is inserted as a default element in environments where affixal features need to be supported. The function of the finite copula in copular sentences is thus like the traditional analysis of *do*-support—it provides a spell-out host for inflection on T in the absence of a lexical verb.[26]

If affixal features on T can be stranded, and rescued by copula-insertion, then we expect that stranded [φ] and [T]-features on Fin can be rescued in the same way. In an ordinary derivation, Fin selects and Agrees with T.[27] I leave detailed discussion of the relationship between Fin and T until Chapter 7, but let us assume for the purposes of the present discussion that T can remerge in Fin to satisfy an EPP property of Fin's [T]-feature. When this happens, whatever supports T's features (e.g., *that*, *to*; Pesetsky and Torrego 2001) provides PF support for Fin's. When Fin fails to select TP in a copular amalgam sentence, there is no host for its features. In these cases, the copula is inserted as a last resort to host [φ] and [T].

---

[26]Bjorkman points out that if the distribution of *be* is predicted by the need to support morphology, then the analysis of *do*-insertion as support must be revisited, since *be* and *do* occur in different environments. Her proposal is that *do* is inserted when a head *v* needs to be spelled out overtly, because it is severed from V. The existence of *do*-insertion therefore need not constitute an argument against *be*-support. Insertion of semantically light or empty verbs is a general process.

[27]Here, assume that T includes both T and Subj, since the position of temporal argument licensing is different from the position of nominal argument licensing. The conflation of T and Subj is a matter of convenience for the present discussion.

The English verbal copula can be inserted in any head along the clausal spine where support for stranded features is needed; it is not restricted to V or T. If it occurs in the V or Asp layers, it takes the bare form *be* (Rothstein 1999; Becker 2004); in the T-layer, where it supports [T] and [φ]-features, it takes finite form (e.g., *is*, *are*). It also takes finite form in the C-layer in amalgams, because it occurs in the "birthplace" of finiteness.

5.3.3.2 Some problem cases

Now let us consider a potential problem for this analysis. In amalgams that include a modal or a semi-copular raising verb, a non-finite occurrence of *be* is required, (66). Recall that although epistemic modals and raising constructions are less acceptable than the simple-copula versions of the amalgams, they are not as unacceptable as the other "baggage" types discussed in section 5.2.4.

(66)   a.   That's her issue could *(be) she needs more time.

        b.   What he wants might *(be) he wants a raise.

In this respect, the amalgam sentence type precisely parallels the canonical copular sentence, (67).

(67)   a.   Her issue could *(be) that she needs more time.

        b.   What he wants might *(be) a raise.

All copular amalgams pattern with ordinary copular sentences in requiring *be* in a sequence with an epistemic modal.

The obligatory presence of *be* following modals and certain raising predicates in ordinary sentences in English poses a challenge for a morphological support approach to copula-insertion. If the copula is inserted to support affixal features, then it must be explained why all of the sentence types (canonical and amalgam) behave on a par with respect to requiring non-finite *be* after a modal. Why is the modal alone insufficient to support finiteness morphology?

(68) a. That guy must *(be) tired.

b. The teacher might *(be) John.

c. What he ate could *(have) *(been) a bagel.

(69) a. %He ate a bagel could *(have) *(been) what he ate.

b. %What he ate might *(have) *(been) he ate a bagel.

c. %That's what he did would *(have) *(been) he ate a bagel.

If the copular clauses in (69) include only Force–Fin, then we might expect the modal to be inserted directly in Fin, and the weight or counterweight to follow directly. Nevertheless, if the canonical string of auxiliaries including *be* is absent, the sentence is ungrammatical, just as in ordinary verbal clauses. Where is the modal, and why is there a head spelled out below the modal in amalgams at all?

Non-finite *be* is required in other syntactic environments that, like copular amalgams, are structurally reduced. Schütze (1997:199) discusses so-called Headlinese (abbreviated English; see Stowell 1991, 1996), where even though finiteness marking (including the finite auxiliary and pred-icational copula) can be omitted, non-finite *be* is not. He cites the following examples, based on Stowell's observations. (The use of all-caps indicates their status as headlines.)

(70) a. HILLARY TO *(BE) INDICTED FOR FASHION FAUX PAS

b. TAX CUT CAN *(BE) PAID FOR, DOLE INSISTS

c. DEFICIT SHOULD *(BE) PRIORITY, PEROT URGES PENSIONS WILL *(BE) SAFE UNDER NEW PLAN

d. CANDIDATES MIGHT *(BE) LYING, ANALYSTS POINT OUT HILLARY WANTS TO *(BE) PRESIDENT, FRIENDS CONFIDE

Schütze (1997, 2004b) and Becker (2004) also observe that English-speaking children and people with Specific Language Impairment can omit the tensed copula, but not the non-finite form *be*. Schütze's proposal is that T must select V, and when a V node is empty and spelled out in isolation (i.e., not head-moved to T), it must be spelled out as *be*. Whether or not head-movement to T takes place depends on the feature specifications of different vocabulary items. Certain modals, imperative, infinitival *to*, etc., when merged with T, block V-raising, so in these contexts, *be* must be inserted into V.

Non-finite *be* is categorically required after a modal across a heterogeneous set of English speakers, varieties, and registers, but this requirement cannot be due to selection. Since copular amalgams lack T, the putative selector of lexical V, a syntactic account like Schütze's (2004b) will not work. A semantic account, which requires verbal *be* to be present in order to introduce an event, is incompatible with the fact that a bare RP (one without an overt copula) can introduce an event/state. (Moreover, in the present account of copular amalgams, this RP introducing the world-dependent fact, the counterpart of the state in canonical sentences, is in fact FinP, and not VP.) I will explore a potential solution that follows the traditional intuition that the modal interferes with the process that transmits morphological [T, φ]-features. Since the syntax of amalgams and canonical clauses is different, but the morphology of the copula in the environment of the copula is the same, an account appealing to the domain of the morphology is desirable.

Following Condoravdi (2002), Cinque (2004), and Hacquard (2009, 2010), I proposed that epistemic modals associate with the context domain—they merge above Fin, where they are anchored deictically to the speaker, rather than to a Topic Time near T. An analogous structural proposal is offered by Giorgi (2010). She observes that inflected verbs with evidential meanings show up in curious syntactic environments (see also Rizzi (2004) on *sembra* 'it.seems'). Giorgi

proposes that the apparently verbal *credo* 'I believe' in Italian is inserted directly in the context

domain of the clause when it is an evidential modifier, as in (71).

(71)   Credo A   PARIGI Maria sia   andata (non a   Londra)
           believe.1SG.PRES TO   PARIS Maria   is   gone (not   to   London)
           'I believe that it's to Paris that Maria went, not to London.'   (Giorgi 2010:70 (35))

Notice that there is no complementizer introducing the clause with the focalized constituent *A*

PARIGI. In an ordinary context, an embedded clause with a focus must include the complementizer.

Contrast (71) with (72).

(72)   Gianni crede   *(che) A   PARIGI Maria   sia andata (non a Londra)
          Gianni believes *(that) TO PARIS   (Maria) is gone   (not to London)
          'Gianni believes that it's to Paris that Maria went, not to London.' (Giorgi 2010:70 (36))

Giorgi argues that the absence of the complementizer (among other properties) indicates that (71) is

monoclausal: *credo* is not a lexical verb with a CP complement, but rather, an evidential modifier in

a high C-domain position. It incorporates first person agreement features (not the features marked

on the main lexical verb), because it is locally valued by the pronominal expression referring to the

speaker (see further discussion in Chapter 8).

The analogy between Giorgi's (2010) left-peripheral evidential heads and epistemic modifiers

in copular amalgams is robust. The copular amalgam supports the semi-copular *seem* modifier, for

example, where there is no structure to support a bona fide lexical verb.

(73)   a.   What he did seems to me is he stormed out.

         b.   He stormed out seems to me is what he did.

(74)   a.   *What he did proves is he stormed out.

         b.   *He stormed out proves is what he did.

In these examples, the copula remains morphologically finite, like *sia* in the environment of *crede* in the Italian example (71), because *seems* is located high in the periphery, where it does not block association of [tense] and [φ] with Fin. The copula is still obligatory, however, because Fin's features need phonological support.[28] Unlike Giorgi's evidential head, however, *seems* bears third person, and not first person. It is likely that *seems* is adjoined, serving as a parenthetical modifier, rather than as an embedding predicate.

Consider now the modals, which are uninflected.

(75)   a.   What he did must be he stormed out.

         b.   He stormed out could be what he did.

Now the copula's form is bare. Unlike with *seems*, the copula cannot take finite form in combination with modal auxiliaries.

(76)   (77)   a.   *What he did must is he stormed out.

                b.   *He stormed out could is what he did.

This asymmetry indicates that the modal is in a head position that interferes with the valuation of Fin's features. I assume, therefore, that there are multiple positions for epistemic modifiers in the left periphery of the clause. Suppose the feature [+fin] is inherent to Fin, but the features [T, φ] are born in Force. Because of the binding relationship that obtains between Force and Fin, the [T, φ] are transmitted to Fin.[29] When the modal head occupied by epistemic *must* and its relatives (call it

---

[28]If the copula is omitted in (74a), the sentence is acceptable for speakers who allow a hanging topic-type weight-initial pseudocleft. The *seem*-modifier is still permitted before the intonation break associated with the topic.

[29]This downward feature-transmission process could also be envisioned as feature spreading (see Tortora 2014 for a formalization of [fin]-feature spreading) from Force onto lower heads in the left-peripheral functional sequence. The presence of the modal above Fin blocks subsequent spreading of the features onto Fin. In Tortora's model, the structures where spreading is blocked, e.g., at the boundary between the T domain and the participial domain in certain modal constructions, are analyzed as "lightly bicluasal".

It is not immediately clear whether the model of light biclausality can be translated into the left periphery. First of all, as I will propose in section 5.4, when the copula is inserted directly into Top, its [φ]-features do not participate in

Mood), has content, it expresses the main assertion, and thus serves as the associate of the focus-sensitive Force operator.[30] Mood becomes the beneficiary of [T, ɸ]-feature transmission, and acts as an intervenor for association of Force with Fin. (78) illustrates the relationship between Force, Mood, and Fin. Lower specifiers are omitted, for simplicity.[31]

(78)

ForceP

Op-ASSERT$_i$ Force′

Force$_{[+decl,uT,u\phi]}$ MoodP$_i$

Mood$_{[uT,u\phi]}$ FinP

might   Fin$_{[+fin]}$

be
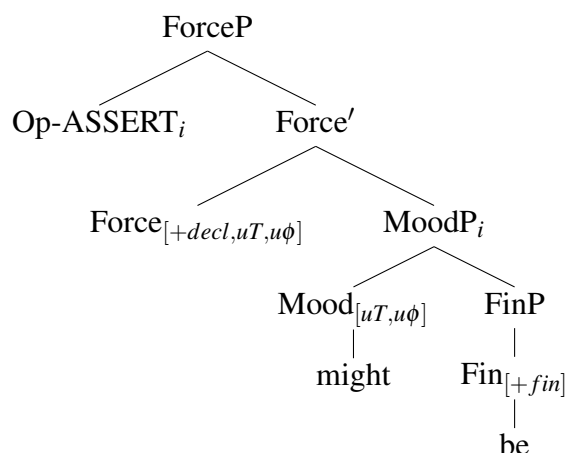
In this case, the finiteness features [+fin], [T], and [ɸ] are expressed discontinuously: [T, ɸ] on Mood, and [+fin] on Fin.[32] Fin must then be spelled out by a form that is even less specified than *is*: the bare copula *be*. The claim that *be* is [+fin] is supported by its occurrence in finite contexts, as in (79). (Although these clauses are not declarative, they are anchored to the utterance context.)

(79)   a.   Be nice!

---

agreement. More importantly, it is already unconventional to treat FinP as a clause independent of a VP. Non-trivial questions are raised by taking this proposal even further and considering ForceP to be a light clause independent of FinP. What does the "clausality" of ForceP consist of, and how can it be delineated from the "clausality" of FinP? Shlonsky (2010) argues that the cartographic enterprise ultimately requires treating ForceP and FinP as separate phasal domains, given the problems it poses for the locality of selection, among other things. If indeed ForceP and FinP project separate domains, then a lightly biclausal model of the amalgam with a feature-spreading boundary between Mood and Fin is plausible, but many questions remain. I leave further discussion and formalization for future research.

[30]Later, I will propose that the first copula in double-*is* occupies a spinal position between Force and Fin, and yet it does not block [T, ɸ]-feature transmission from Force. This is because Top is not a potential associate of Force—it does not express a focused assertion.

[31]Given the analysis of subject licensing I will develop in Chapter 7, I will assume that the overt counterweight is base-generated in Spec,MoodP, binding an empty category in Spec,FinP.

[32]In a further complication, when *seems* can also be followed by a non-finite occurrence of *to be*; it cannot be followed by *be* alone. A cartographic approach to this puzzle, which may not be optimal, would propose that *seems* is also an intervenor in Force association. It blocks transmission of [T, ɸ] to Mood, which simply bears [+realis]. Mood's [+realis]-feature is then supported by *to*, and Fin's [+fin], by *be*.

b.   They require [that you be available on weekends].

The association of [T, φ] with Force is corroborated by the fact that different Force specifications correlate with different [T, φ] features on the verb: imperative and subjunctive clauses, as (79) illustrates, have *bare* finite verbs.[33]

   To summarize, the modal + *be* sequence occurs in amalgams where Fin does not receive [T] and [φ]. The epistemic modal is inserted in a position above Fin where [T] and [φ] are spelled out, as a result of transmission from Force, and Fin's [+fin] and [bare]-features are supported by the auxiliary *be*.

5.4   Deriving the copular amalgam family

Now let us apply the minimal finite clause model of copular amalgams to the different amalgam sentence types. Counterweight-initial amalgam pseudoclefts are base-generated Fin-headed clauses lacking the typical pattern of A-movement of the subject into the grammatical domain. Chapter 2 showed that the weight-initial pseudocleft is like its canonical counterpart in having a fixed Topic-Comment information structure. In this respect, the weight-initial amalgam pseudocleft is also like the other members of the amalgam copular sentence family, which are not reversible. There are at least two immediately plausible ways to generate this order, in both pseudoclefts and non-pseudocleft copular amalgams. The first possibility is that the Topic-Comment order is derived by predicate inversion, just as in the canonical pseudocleft. The second possibility is that it is base-generated, as den Dikken et al. (2000) propose for weight-initial amalgam pseudoclefts. The evidence suggests that both structures are available. The basic amalgam pseudocleft includes Fin,

---

[33]In canonical declarative clauses where [T, φ] is separated from [+fin], the bare [+fin] feature is ordinarily spelled out on the lexical verb or highest auxiliary, rather than on Fin. It seems, therefore, that the feature [+fin] is only spelled out once per clause. In amalgams, [+fin] is spelled out in its basic position, Fin, but in canonical clauses, the Fin values a lower head in the functional sequence for [+fin], and it is this lower occurrence of the feature that has phonological content.

while non-reversible amalgams are projected from Top.

There is a grammaticalization path between T/Infl and Foc/Top; this path can also be observed in acquisition, where English language learners use inflected copular forms as topic-markers (Kim 2011). It is unsurprising therefore that the finiteness-marking and information-structure-marking functions are difficult to tease apart in the different members of the copular amalgam sentence family (see also discussion in Massam 1999, 2013). Since the predicate-initial amalgams do not form a completely homogeneous class with respect to these functions, and since there is robust speaker variation, a heterogeneous analysis is unavoidable. The existence of a variety of clause types with different "sizes" of spine is actually predicted by the present approach, which allows higher domains to be independent of lower domains.

5.4.1   Speaker variation

Previous literature and within-subjects comparisons reveal rampant speaker variation with respect to the different copular amalgam sentence types. For some speakers, all of the copular amalgam sentence types discussed in this thesis are acceptable, while others accept only certain members of the family. Based on informal surveys, I have identified the types of speakers represented in Table 5.2.

	Rev. amalg. pseudocleft	*Wh*-initial amalg. pseudocleft	That's x is y
Type 1	✗	✗	✗
Type 2	✓	✓	✓
Type 3a	✓	✓	✗
Type 3b	✓	✓(only when V=do)	✗
Type 4a	✗	✓	✓
Type 4b	✗	✓(only when V=do)	✓
UNATTESTED	✓	✗	✓

Table 5.2: Speaker types

An interesting pattern emerges in my informal and formal survey responses: the weight-initial

amalgam pseudocleft seems to be a structural chameleon. There are speakers who accept weight-initial amalgam pseudoclefts and reverse amalgam pseudoclefts, but not *that's x is y*, and there are speakers who accept weight-initial amalgam pseudoclefts and *that's x is y*, but not reverse amalgam pseudoclefts. This pattern is compatible with the findings for other languages as well. For example, German speakers I surveyed accept weight-initial amalgam pseudoclefts and *that's x is y*, but I have not been able to identify any speakers who have the reverse amalgam pseudocleft (or DP-weight amalgam).

Now consider the pattern in the responses to the debrief questions in the formal acceptability experiments I conducted. After completing acceptability surveys including baseline amalgam pseudoclefts (but not *that's x is y*, because otherwise the surveys would have been too long), participants rated their perception of how likely they were to use sentences like each of those in (80), below, as "likely", "possibly", and "unlikely".

(80)   a.   What I need is I need a vacation.

        b.   I need a vacation is what I need.

        c.   That's what I need is I need a vacation.

The responses should be taken with a grain of salt, since we know that speakers do not accurately assess their own use of this construction. In addition, *that's x is y* and the weight-initial pseudocleft allow for ellipsis in the counterweight clause—since speakers evaluated these sentences in writing, the overtness of this material might have lowered their responses.[34] The pattern is suggestive, nonetheless. Out of 159 responses (not all survey participants answered the debrief questions), 99 participants reported that their use of the weight-initial pseudocleft was either "possible" or

---

[34]If I had used examples with *do* in the weight clause, for instance, I suspect the self-reports would be more favorable, since *do* in these amalgams is highly frequent. Similarly, if I had used reduced counterweights in *that's x is y*, ratings would likely improve, but the parallelism with the counterweight-initial version would be lost.

"likely". 88 gave these ratings to the reverse amalgam pseudocleft, and 78 did so to *that's x is y*. Only 22 participants said they were "unlikely" to use the weight-initial amalgam pseudocleft but would at least "possibly" use *both* the reverse amalgam pseudocleft and *that's x is y*. (14 said they were "unlikely" to use all three.)

These numbers are not definitive, given the confounds of optional reduced counterweights and bias against amalgams in written form, but the pattern of variation suggests that there are multiple grammars are work. My analysis of weight-initial amalgam pseudoclefts takes this into account.

5.4.2 Weight-initial amalgam pseudoclefts: two sources

The weight-initial amalgam pseudocleft has a fixed Topic-Comment information structure. This information structure can be derived from two different syntactic derivations. The first involves predicate inversion, as in the canonical pseudocleft, and the second involves base-generation of a Topic-Comment structure (den Dikken et al. 2000). In the former structure, the weight-initial amalgam pseudocleft is derivationally related to the counterweight-initial version. Speakers who have the reverse amalgam pseudocleft—base-generated FinP with the counterweight as structural subject—are predicted to allow predicate inversion around Fin. In the Topic-Comment structure, which may coexist with the inversion option in the same speaker's grammatical repertoire, projection takes place directly from Top, rather than from Fin.

It is not obvious how to distinguish these two sources. In canonical specificational copular sentences, several tests are available, since the two structures are associated with different predictions: the subject either occupies an A-position or an A'-position. This in turn has consequences for restrictions on movement and the form of the copula. In amalgams, however, which are frozen and restrict the form of the copula, these tests are unavailable. Formal differences between the two structures are subtle.

5.4.2.1 Predicate inversion in amalgam pseudoclefts

There are three reasons to assume that predicate inversion is available in amalgam pseudoclefts.

5.4.2.1.1 Parallelism with the canonical copular sentence

The first reason is conceptual: if the counterweight-initial structure is a close analogue to the value-initial structure in a canonical sentence, and the mechanism of predicate inversion is not sensitive to the A/A$'$ distinction, then there is every reason to expect that predicate inversion around Fin should be possible. Recall from Chapter 4 that the weight is a predicate in both sentence types: in canonicals, it is a predicate of individuals, and in amalgams, it is a predicate of propositions. The consequences of predicate inversion are the same in the two sentence types.

In canonical specificational copular sentences, either the subject or the predicate of the small clause ultimately ends up in Spec,TP, the structural subject position.[35] The EPP-specified [φ]-feature on T (inherited from Fin) attracts the small clause subject to its specifier. In this configuration, information structure is free: the sentence can receive either a predicational or specificational interpretation, and either the subject or the predicate can be focused.

The other option, predicate inversion, requires an extra syntactic step. This extra syntactic step involves domain-extending head movement (den Dikken 2006), which allows the predicate, originally in the complement position of a phase head, to access a derived "escape hatch". When the phase head remerges, a new edge position becomes available. The predicate can then invert around its subject, and raise into the higher T-domain. Predicate inversion can apply freely, provided the predicate can satisfy the EPP property of the [φ] probe.

In amalgams, we expect the same process to be available. If Fin remerges, a derived specifier

---

[35]Recall that I use Spec,TP as a shorthand for both the structural subject position and the tense position. I assume that these are separate, to make room for a Topic Time argument (e.g., Rizzi 2004; Rizzi and Shlonsky 2007), but since conflating them has no consequences for the present discussion, I use the familiar labels.

position becomes available, making the weight free to move around the counterweight. This derived position is, conveniently, Spec,TopP. The complex head Fin+Top has the inflectional features of Fin, and so the same finite copular form spells it out.

(81)

```
                          ForceP
                         /      \
                  Op-ASSERT      Force′
                                /      \
                            Force       TopP_t
                                       /      \
                                   CP_i        Top′
                                  /    \       /    \
                            What he did  Top+Fin     FinP
                                          |          /    \
                                          is        CP     Fin′
                                                  /   \    /    \
                                            he left early <Fin> <t_i>
```
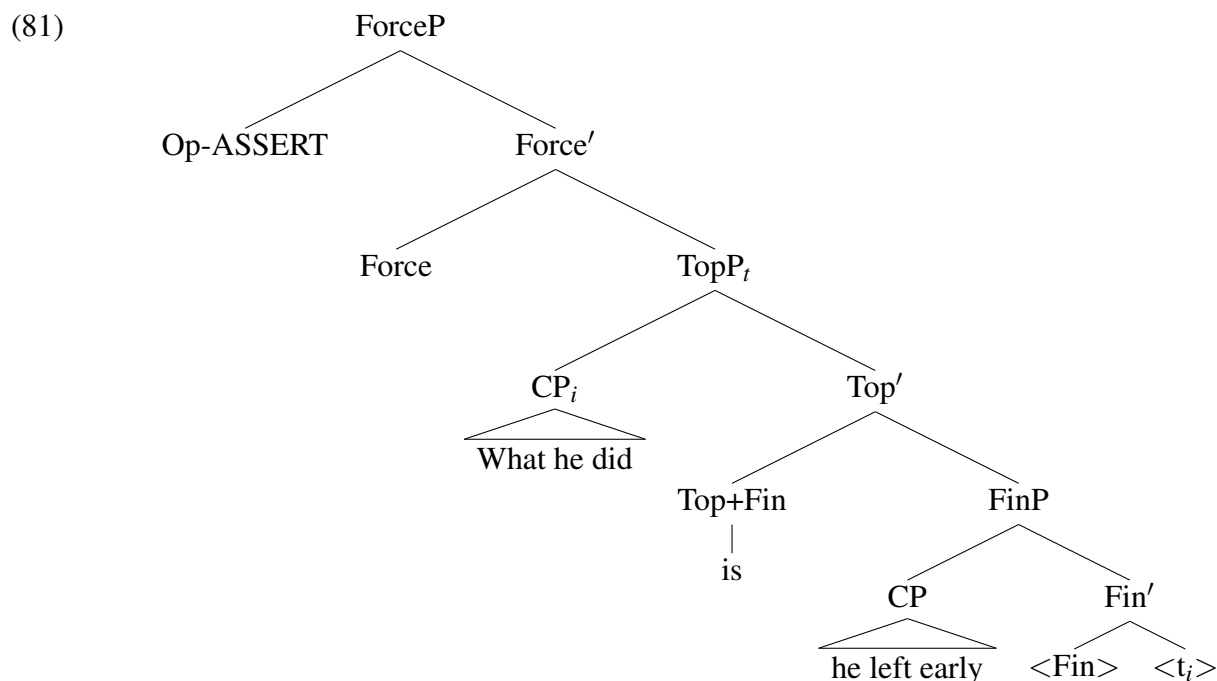
The resulting configuration is the desired Topic-Comment configuration. While the parallel between the canonical predicate inversion structure and the weight-initial amalgam structure is consistent with the existence of predicate inversion in amalgams, it is also consistent with a base-generated Topic-Comment structure.

5.4.2.1.2 Agreement features

Support for the inversion option comes from the distribution of number agreement morphology. In canonical specificational sentences, the copula's [φ] inflection reflects the Agree relation that holds between T and its structural subject. Agreement with a postcopular subject in a predicate inversion sentence is possible in Italian (Moro 1997), Faroese (Heycock 2012), and other languages, but predicate inversion in English normally renders the small clause opaque for Agree relations.

(82) a. My favorite person is/*are you.

   b. You are/*is my favorite person.

   c. My biggest problem is/?are the kids.

   d. The kids *is/are the biggest problem.

A similar Agree relation holds in amalgam pseudoclefts, but it is triggered by the features of Fin rather than T. While it is typically singular, the copula can take plural form when the weight is a plural DP. (Recall discussion in Chapter 4.)

(83) a. **The main problems are** he misunderstood the budget and the committee decided against him.

   b. He misunderstood the budget and the committee decided against him **are the main problems**.

The experimental results reported in Chapter 4, supplemented by informal judgment surveys, indicated that the plural form of the copula cannot be valued by a conjunction of counterweight clauses—only a plural weight can be responsible for the occurrence of *are* in amalgams.

The copula in both canonical and amalgam pseudoclefts shows the same pattern of number agreement in the environment of a plural weight clause in its specifier. This similarity is expected if the copula spells out heads with similar feature composition, and unexpected if it spells out two very different heads. Since Fin and T have similar feature composition, while Top and T do not, the number agreement facts suggest that the amalgam copula in these weight-initial pseudoclefts spells out a head that includes Fin, rather than Top alone.

5.4.2.1.3 Force

The third piece of evidence that the counterweight-initial and weight-initial pseudocleft can be derivationally related comes from the distribution of Force. Full finiteness requires a specified Force head and a (positively) specified Fin head. The structure in (81), therefore, includes a projection of Force above TopP, and an ASSERT operator. This operator binds the proposition associated with FinP. If the structure I propose for counterweight-initial pseudoclefts can underlie weight-initial pseudoclefts, then we expect to find independent illocutionary force associated with the copular clause in both sentence types. If there is only a base-generated Topic-Comment structure, then the full finiteness profile is not expected—the copular clause should lack its own illocutionary force. The former relation occurs in weight-initial pseudoclefts for some speakers.

In the amalgam pseudocleft, the counterweight clause is certainly the most informative proposition, but it is not the only locus of illocutionary force. The copular clause associates with force as well. It can combine with illocutionary adverbs like *really*, and it can support a verum focus interpretation when prosodically stressed.

(84)   What she needs really IS she needs a break.

In most cases, both the counterweight and the amalgam clauses are assertions, and so it is difficult to distinguish the two Force operators.

(85)   a.   What I need is I need a break.

        (i)   I need a break.

        (ii)   The answer to the question of what I need is: I need a break.

    b.   I need a break is what I need.

        (i)   I need a break.

(ii)   The answer to the question of what I need is: I need a break.

Pragmatic constraints on the question-answer clause make it difficult to show that there are two separate assertions here. For example, modifying them with different speaker-oriented adverbs or different epistemic modality is typically infelicitous: the speaker cannot easily have a different degree of commitment to a focused assertion's truth than she has to that focused assertion's truth as the answer to a presupposed open question. (86) seems acceptable, but it is difficult to construct additional examples, and difficult to ensure that the second speaker-oriented adverb is not a parenthetical modifying the copular clause.

(86)   ?Surprisingly, what he did was unfortunately, he stormed out without saying anything.

When the counterweight clause is interrogative, however, the two Force operators can be distinguished. The primary speech act is the assertion associated with the copular clause. In (87), for example, the speaker can use a tag question, compatible with declarative force, and the interlocutor can felicitously respond by affirming or denying the relationship between the weight and the counterweight, as in (87a and b) (both diagnostics for the presence of an ASSERT operator, Meinunger 2004, 2006).

(87)   What he wants to know is what is he going to do (right? / eh? / huh? / isn't it?)

    a.   I know! He has no idea, poor guy.

    b.   No, I think he has it pretty well figured out.

The speaker can be interpreted as *reporting*, rather than *performing* the interrogative in the counterweight.[36]

---

[36]The speaker can also be interpreted as performing the illocutionary act associated with the counterweight. Although traditionally, associating an embedded clause with Force was thought impossible, this notion has been reconsidered, beginning with the analysis of embedded root phenomena in Hooper and Thompson (1973). More recently,

Recall from the discussion in Chapter 2 that the independent status of the copular clause's illocutionary act in amalgam pseudoclefts contrasts with other constructions relating root-like propositions, like interrogative slifts, where the speaker is necessarily taken to be asking the putatively embedded root question, and there is only one Force operator (Haddican et al. 2014). Crucially, in slifting, no relation between the two CPs is asserted: there is no copular clause putting the pieces together here, only a Topic-Comment-like evidential relationship.

(88)   How old is she, did she say?   (Haddican et al. 2014:100 (88))

      a.   45.

      b. *She did, 45.

      c.   45, but she didn't say it.

      d. *Yes/No.

In the amalgam, the interlocutor can felicitously respond to the copular clause's force, rather than the counterweight, while in the slift, the interlocutor must respond to the main information question. Since the copular clause can be associated with an independent force value in the amalgam pseudocleft, regardless of linear order, the amalgam pseudocleft includes both Fin and Force in its spine.

5.4.2.2 Base-generated Topic-Comment pseudoclefts

Now consider the structure in (89), adapted from den Dikken et al.'s (2000) proposal for the weight-initial amalgam pseudocleft.

Krifka (2014) has proposed that one Force operator can be embedded under another: the embedded root is not merely in the scope of the matrix Force—it has its own Force. The copular amalgam offers additional evidence that embedded roots can bear Force, particularly embedded roots in *subject* position, which have not been studied previously.

(89)

```
                        ForceP
                       /      \
                 Op-ASSERT    Force′
                             /      \
                         Force      TopP_t
                                   /      \
                               CP_i        Top′
                              /    \       /    \
                      What he did  Top    FinP
                                    |     /    \
                                    is   he left early
```
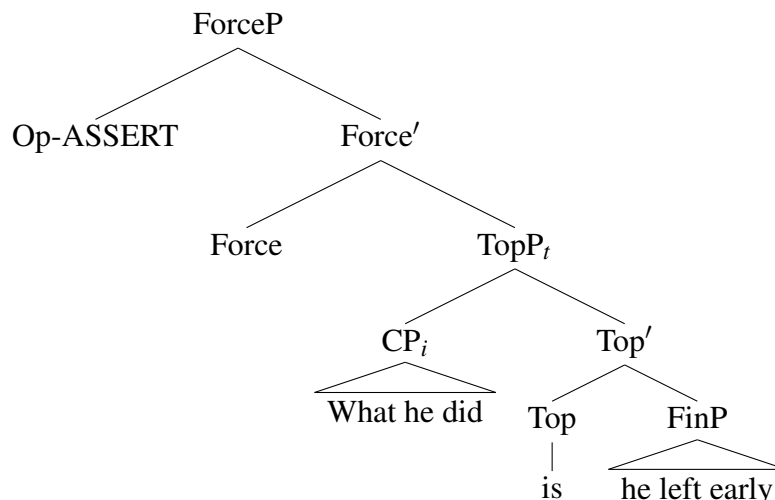
In this structure, the weight clause is base-generated as a Topic, the copula spells out a simple topic-marker, and the counterweight clause is the matrix clause. The Force layer above TopP is the "extended" projection of the counterweight clause. (89) is therefore a canonical clause type; it is unusual only in that its topic-marker is morphologically finite, and has the form and predicational function of the verbal copula.

This structure is consistent with the now familiar empirical profile of the amalgam pseudocleft. The copula must be in a simple morphological form, these sentences are not embeddable in ECM contexts, information structure is fixed, and the structure is frozen, resisting A- and A′-movements. The copula lexicalizes a simple functional head, Top, and so it cannot host other morphology or auxiliaries. The structure is frozen because extraction of or over a criterial topic is generally impossible (Rizzi 2006).

There is evidence, however, that even Topic-Comment amalgam pseudoclefts are not canonical clause types: the counterweight is not the matrix clause. The amalgam pseudocleft in this case is projected directly from Top. While the complement of Top is a clause with fleshed out V and T-domains (as in a monoclausal structure where FinP or FocP occurs below Top), it can also be a fully specified root, (90).

(90)   a.   What disturbs me is what did you think you would accomplish?

        b.   What she needs is, for her own sake, she needs to quit.

The presence of root phenomena in the counterweight clause argues for a bisentential model of the Topic-Comment amalgam pseudocleft, with the "stunted" spine of the copular clause projecting independently of the counterweight, from Top.

Evidence that the base-generated Topic-Comment amalgam structure is available in addition to the inversion structure comes primarily from the existence of speakers like those reported in den Dikken et al. (2000), who do not have counterweight-initial pseudoclefts. In languages that lack the counterweight-initial option, I assume that the amalgam copula has been grammaticalized as a topic-marker.

Because the copula is a simple topic-marker in this structure, it is not expected to inflect for [ϕ]-features. The [3sg] form of the copula is insensitive to the features of the weight. The [3sg] feature is probably a default, a vestige of the copula's form in the inflectional environment. It is not in an EPP-specified Agree relation with the phrase in its specifier, in other words, in contrast to the Top+Fin head in the predicate inversion structure. While speakers who have the predicate-inversion amalgam pseudocleft structure allow the amalgam copula to inflect for plural, speakers who only have the structure proposed by den Dikken et al. (2000) do not accept plural agreement on the copula. In fact, these authors use the impossibility of plural agreement to argue that the copula is only a simple topic head (see also discussion in Chapter 4). The fact that speakers with both orders allow the plural form of the copula, while speakers with only the weight-initial version do not, constitutes strong evidence that there are two different sources for the weight-initial amalgam pseudocleft string.

There is something "squishy" about the copula's distribution as a topic-marker, however, because it shows tense inflection (something den Dikken et al. 2000 struggle to explain) and only occurs where it links propositional elements in a specificational relationship. It is not a run-of-the-mill topic-marker:

(91)   a.   *As for Mary, is, she will make a great doctor.

           b.   *To Mary, is, we should give a book.

The examples in (91) are not acceptable to me, but I speculate, after several years of tuning my ear to such things, that they either have occurred, or will occur in someone's English. The Topic-Comment tendencies of English, its ubiquitous copula, its impoverished agreement system, and the common grammaticalization trajectory from copula to discourse particle (and the acquisition trajectory in the other direction; Kim 2011) could easily conspire to yield (91).

5.4.3   The Colon Phrase: coordination meets specification

In many respects, the syntax of copular amalgam sentences parallels that of coordinate structures. If, as suggested in Heycock and Kroch (1999) and den Dikken (2006), there is a deeper similarity between predication and coordination, then it is unsurprising to find a type of Relator Phrase that shows mixed properties. Coordination and predication structures locate two expressions in a strictly local relationship mediated by a functional head, both are recursive, and both allow for non-constituent ellipsis under identity. Sentential coordinations and copular amalgams both accommodate root-clauses in specifier positions. Both relate "likes": two propositional elements. They are both subject to extraction restrictions: the frozenness of the amalgam is reminiscent of the Coordinate Structure Constraint (Ross 1967).

The heads of the two structures also share some properties. Both express simple logical rela-

tions (essentially, intersection), and occupy functional heads in the context domain of the clause.[37]

While the amalgam copula patterns in some respects like a verb, its functions also overlap with the functions of conjunctions like *and*.

5.4.3.1 Specifying coordination

Coordinate structures and specificational copular clauses can have the same specificational interpretation. While prototypical coordinate structures are headed by the basic conjunctions *and* and *or*, coordinate constructions with specialized semantics are prolific in natural language. De Vries (2009) describes a large inventory of semantically distinct coordinators that all show similar syntactic behavior. The present proposal adds the English copula to this inventory.

Koster (2000) proposed that appositive modification (and extraposition) is mediated by structural coordination. Extraposed constituents serve as specificational modifiers of their "anchors" in the main clause. In later work, de Vries (2009), Kluck (2011), and Ott and de Vries (to appear) expand this proposal to account for non-restrictive relative clause modification, right and left dislocation, and parenthesis. The relationship between the extraposed constituent and the anchor is mediated by structural coordination. Koster (2000) describes the type of coordinate structure at work in extraposition and apposition as a "Colon Phrase", e.g., in (92).

(92)   John built something beautiful: a golden igloo.

This terminology is highly reminiscent of the "colon function" ascribed to pseudoclefts and specificational copular sentences since Higgins's (1979) seminal work (e.g., Declerck 1988; Lambrecht 2001; den Dikken 2005b, 2013), although the authors working under the coordination hypothe-

---

[37]In Dravidian, for example, coordinators and finiteness/force markers compete for the same syntactic position (Jayaseelan 2013). Of course, the same is not true for English, but it is not unreasonable to associate sentential conjunction with a high left-peripheral position, and therefore to suppose that conjunctions and the amalgam copula may be more closely related than they appear at first blush.

sis for apposition do not associate the Colon head with an overt copula, or with the pseudocleft construction.

Appositives, as in (93), whether they are asyndetic or mediated by an adverbial or overt coordinator, have a similar colon function to specificational sentences.

(93) a. I met John, your neighbor.

    b. Joop lives in The Netherlands, **or** Holland, as it is often called.   (de Vries 2009:13 (32a))

    c. I have a better idea, **namely**, that we should call ahead.

(94) a. The winner is: you!

    b. Here's what I need, **is** I need a new job.

The similarity between coordination and specification goes two ways: the copula can have a coordinating function.

The two constructions also share some marked properties. The structure they occupy is "frozen", recalling the Coordinate Structure Constraint (Ross 1967). The two major constituents linked by the copula in amalgams must be parallel in the (semantic) sense of the Law of Coordination of Likes: both refer to propositions (see discussion in Chapter 4). In both canonical coordinate structures and amalgams, the conjuncts can be a question and an answer.

(95) a. [You can lead a horse to water], **but** [you can't make it drink]

    b. [You can lead a horse to water], **but** [will it drink?]

(96) a. [That's what you can do] **is** [you can lead a horse to water]

    b. [You know what you can try?] **is** [you can try leading a horse to water]

Recursion is available in both coordinate structures and copular amalgams, but notably, not in canonical specificational copular sentences, (97).

(97)   a.   I saw [Joe and [John and Mary]].

        b.   [[That's what I need is I need a coffee] is what I'm saying].

        c.   *[[What I like to eat is fruit] is apples].

Coordinate structures and amalgams both license non-constituent ellipsis, (98). The fragment answer ellipsis operation available in amalgams is analogous to conjunction reduction.[38,39]

(98)   a.   She plans to give John a book and then ~~she plans to give~~ Mary a gift card.

        b.   That's what she thinks of, is ~~she thinks of~~ her first train ride.

The novelty of this proposal analogizing copular specification and coordination is that the copula connects two propositional elements while also encoding finiteness morphology. If the amalgam copula were merely a grammaticalized conjunction, its inflection would be quite surprising. Not only does the copula inflect for finiteness; it also interacts with illocutionary force in amalgam pseudoclefts. Conjunctions clearly lack these properties.

(99)   a.   She went to Paris and(*__ed__) he went to Rome.

---

[38]The analysis of ellipsis or constituent sharing in coordinate structures is a contentious one, and for reasons of space I cannot do justice to the debate in this thesis. Since what matters for the present purpose is simply the parallel between coordination and the copular amalgam structure, I leave the analysis of pseudo-gapping, etc., open.

[39]One feature of coordination that does not hold for specificational sentences is iterativity.

(i)   a.   *That's my feeling, is I'm too tired to go out, is I'd rather stay in.
        b.   *The teacher is Mary is/and John.

While coordination in its most basic form *and* can iterate, not all coordinators (e.g., *but*) can iterate freely.

(ii)   ?John is very pretentious, but he's a linguist, but he's not interested in grammar.

Iteration in coordinate structure is constrained by semantics and pragmatics (de Vries 2009:10). Since specification carries an exhaustivity implicature, the weight cannot be specified by two different values in two different clauses. Since the possibility of iteration depends on the semantics of the coordination relationship, the fact that *specification* is not iterative does not undermine the parallel between coordination and specification.

    b.   She likes coffee and(*s) he likes tea.

(100)   a.   *She went to Paris **really AND** he went to Rome.

          b.   *She likes coffee **sure AND** he likes tea.

In addition, the non-optionality in pseudoclefts would also be surprising if it were *just* a conjunction. Recall that a reverse amalgam pseudocleft with a dropped copula feels distinctly like an ordinary predicational copular sentence in English with illicit copula drop (101), and not like the optionally asyndetic Colon Phrase discussed above (102), or like stylistic asyndetic coordination (103).

(101)   a.   *I need a new job why I'm upset.

          b.   *The weather why I'm upset.

(102)   a.   I met your neighbor, John.

          b.   I have a better idea: we should call ahead.

(103)   I came. I saw. I conquered.

Since the context domain of the clause can be divorced from the lower domains, we can reconsider the types of relations that context-domain elements can encode. There is no reason why a functional head like Top or Fin should not have a coordinating function in addition to a discourse-marking function. Conversely, there is no reason why a coordinator should not become grammaticalized as an instance of Fin, as it has in Dravidian. Jayaseelan (2013) shows that conjunctions in Malayalam are in complementary distribution with Fin: only non-finite clauses can be coordinated.

(104)   a.   *John wannu-um, Mary pooyi-um.
             John came-CONJ Mary went-CONJ
             Intended meaning: 'John came and Mary went.'

b.   John war-uka-(y)um   Mary pook-uka-(y)um cey-tu.
John come-INF-CONJ Mary go-INF-CONJ   do-PAST
'John came and Mary went.' (Lit. 'John to come and Mary to go, did.')

(Jayaseelan 2013)

Jayaseelan suggests that Malayalam's left periphery is not as articulated as it is in other languages, leading to competition between these elements. English offers enough room for both coordination and finiteness, but in amalgams, they seem to be in close proximity. The hybrid coordination and verbal properties of the amalgam copula are summarized in Table 5.3.

	Verb *is*	Amalgam *is*	*And*
Colon Function	✓	✓	✓
Nominal "subject"	✓	✓	✓
Tense inflection	✓	✓	✗
Agreement inflection	✓	✓	✗
Force	✓	✓	✗
Negation	✓	✗	✗
Temporal auxiliaries	✓	✗	✗
Tense meaning	✓	✗	✗
Root clause specifier	✗	✓	✓
"Likes"	✗	✓	✓
Recursion	✗	✓	✓
Non-constituent ellipsis	✗	✓	✓

Table 5.3: Verbal and coordinator properties of the amalgam

5.4.3.2   Appositive-like copular amalgams

Let us now consider the weight-initial copular amalgam sentence types in the light of the coordination analogy sketched above. The primary sentence type to be considered here is *that's x is y*, but also relevant is the question-answer amalgam (see also Lambrecht and Ross-Hagebaum 2006) and the free-*be* amalgam. In each of these sentence types, the precopular weight clause—not just the counterweight clause—is root-like.

(105)   a.   [That's what I'm worried about], is [I'm worried about the cost].

      b.   You know what you should do, is [you should apply again in March].

      c.   [I made the point once before], is [(that) we have to work on this committee]. (Mc-
          Convell 2004)

5.4.3.2.1   Afterthought function of appositive-like copular amalgams

Coordination is a simple and appealing model for such structures: they include two complete sentences (note the bracketing above). In addition, they often have an "afterthought"-like quality to them. The speaker makes an assertion, and then tacks on an apposition to resolve an ambiguity or reinforce part of the assertion. Coordination, considered from an online perspective, is a very useful strategy for expressing afterthoughts. Both structures are fundamentally additive.

In this respect, the non-reversible amalgam sentence types pattern with the counterweight-initial amalgam pseudocleft, rather than with the weight-initial version. The weight-initial pseudocleft begins with a sentence fragment, so speaker and hearer are fully aware that more structure is coming. The root-clause initial amalgam types, however, are frequently used to express afterthoughts. Speakers are keenly aware of this: many of my informants make meta-linguistic comments to this effect when I ask them for judgments. This may be part of the reason these sentence types suffer from prescriptive stigma. In carefully monitored speech (and writing), forethought is considered valuable. The afterthought use is apparent in the prosodic contour of these sentences, which features an optional pause before the copula, in contrast to the weight-initial amalgam pseudocleft, which more often places the pause after the copula.

Viewed from this on-line perspective, the counterweight-initial pseudocleft and the non-reversible *that's x is y*-type amalgams are basic, while the weight-initial amalgam pseudocleft requires more planning. In the former sentence types, an object that would be happy to serve as a totally indepen-

dent clause is snatched up by another structurally reduced, yet formally finite clause, to serve as its structural subject/first conjunct. The copula links this initial clause to its apposition and projects either a full context-domain spine from Fin to Force, or a stunted spine, lacking either Fin, Force, or both. In the stunted, non-reversible amalgam types, the copular clause does not have independent Force, similar to a coordinate structure.

5.4.3.2.2   Force in *that's x is y*

While the previous discussion showed that in amalgam pseudoclefts, the copular clause and the counterweight can have two independent specifications for Force, and express two independent finite clauses, this is not the case in *that's x is y*. In *that's x is y*, the precopular clause can be interrogative, expressing a polar question. When this happens, there is no way for the copular clause to be interpreted as declarative. It is interpreted as part of the question, and pronounced with polar question intonation, (106), like a non-restrictive apposition to a DP, (107).[40]

(106)   Is that the best way to get off of that stuff? Is just to sleep it off?

    a.   Yeah, it is.

    b.   #Yes, that's the best way, but it's not to sleep it off.

    c.   #I don't know, but thanks for telling me that it's to sleep it off.

(107)   a.   Is that who you saw, John?

    b.   Did you see my neighbor, that nice guy I told you about?

In contrast to the amalgam pseudoclefts, verum focus and Force adverbs are incompatible with the copular clause in *that's x is y*.

---

[40]De Vries (2009) and Kluck (2011) show that clausal appositions themselves, like parentheticals, can have different Force specifications than their hosts, but what is relevant in the present discussion is whether the relation between the apposition and the host can have its own Force.

(108)   a.   *That's what I did IS I quit my job.

        b.   *That's what I did really is I quit my job.

The same distribution of Force is observed in single-question slifting, which likewise has an afterthought quality (Haddican et al. 2014). Recall (88), reproduced below.

(109)   How old is she, did she say?

A single interrogative Force operator scopes over the whole structure (Haddican et al. locate it in a Speech Act projection above the whole sentence), binding only the first clause, *how old is she?*. The second clause, *did she say*, is formally interrogative, but has no independent Force. In slifting, of course, the relationship between the two clausal constituents is not mediated by a finite element, so it is not expected to have Force.

5.4.3.2.3   Force in question-answer amalgams

The question-answer amalgam includes a mix of epistemic and specificational properties, representing something of a middle ground between slifting and *that's x is y*. (Square brackets in the examples below remind the reader that these are not embedded amalgam pseudoclefts.)

(110)   a.   [Guess who I saw], is (I saw) John.

        b.   [You know who I saw], is (I saw) John.

Like slifting, the evidential/epistemic clause may include only a small class of verbs. It also shows person restrictions: the weight clause in a question-answer amalgam must be addressee-oriented, like an imperative. In the question-answer amalgam, the precopular clause prompts the hearer to consider the question. It identifies the question as given, and activates the possible answers

to the question in the hearer's mind.[41] Because of its information function and imperative-like

Force, other epistemic verbs are not licit. (Note that the following are ungrammatical with the

precopular string construed as a constituent; if the string corresponding to an amalgam pseudocleft

is embedded under the epistemic verb, e.g., preceded by *that*, the structures are grammatical. The

difference is subtle, but the parses can be disambiguated prosodically.)

(111)   a.   *[You believe who I saw], is I saw John.

         b.   *[We guess who I saw], is I saw John.

The postcopular clause in the question-answer amalgam is declarative. The weight and coun-

terweight thus have different specifications for Force, but the copular clause itself seems to lack

Force. For example, it is impossible to trigger verum focus alternatives (consistent with assertion)

by stressing the copula in a question-answer amalgam (112). (To help disambiguate the question-

answer amalgam parse from the embedded amalgam pseudocleft parse, I will write a question mark

after the weight clause to signal the peculiar imperative Force that accompanies these *guess*-type

clauses.)

(112)   a.   *Guess who I saw? IS I saw John.

         b.   *You you know who I saw? IS I saw John.

The use of force adverbials is similarly bad (113).

(113)   a.   *Guess what I did? (sure) is (indeed) I quit my job.

         b.   *You know what I did? (sure) is (indeed) I quit my job.

Tag questions on the copular clause also fail, (114).

---

[41]The imperative-like Force associated with *guess*-clauses is odd—it has some interrogative properties. Indeed, English speakers often use a question mark when writing *guess*-imperatives. I am not aware of any analysis of this type of Force. It is formally imperative, but its information-seeking function should not be dismissed.

(114)   a.   *Guess what I did? is I quit my job, isn't it?

         b.   *You know who I saw? was I saw John, wasn't it?

Since an independent assertive interpretation of the copular clause is not available, it lacks independent Force.

5.4.3.3   The clausal spine of appositive-like copular amalgams

We have seen in this section that the appositive-like copular amalgams, which are not reversible, are similar to coordinate structures. They occupy a middle ground between the fully finite relation encoded in the reversible amalgam pseudocleft, and the totally non-finite relation encoded in the canonical coordinate structure. Although their copula is formally finite, its projection lacks independent Force. The Force of the sentence instead comes from the weight and/or counterweight clauses. The clausal spine of these copular amalgam sentence types is stunted, by comparison to the reversible pseudocleft type. The question is: does it lack structure at the top (Force) or at the bottom (Fin)?

The first possibility is that the copular clause lacks a projection of Force—the proposition it expresses (associated with a variable in the environment of Fin or Top) is not asserted. Its reduced finiteness could be compared to that of a non-V2 embedded clause in Germanic, which is fully specified for morphological finiteness and is anchored to the utterance context, but which is not asserted (e.g., Haegeman 2004; Meinunger 2006).

(115)   Ich habe gehofft, [dass sie mich in den Urlaub   mitnehmen würde].
         I   have hoped   that she me   in the vacation with-bring would
         'I hoped that she would bring me along on vacation.'

If this is correct, then the precopular clause which does have Force is structurally subordinate, as it is in a coordinate structure, or the more controversial model of embedded Force in (Krifka 2014).

Consider the possible structure for *that's x is y* in (116).

(116)

```
                                    FinP
                        _____
                       /                        \
                   ForceP                       Fin′
              _____              _____
             /              \            /              \
        Op-ASSERT         Force′        Fin             CP
                      _____        |         _____
                     /           \       is        I want a break
                 Force        ...FinP
                          _____
                          That's what I want
```
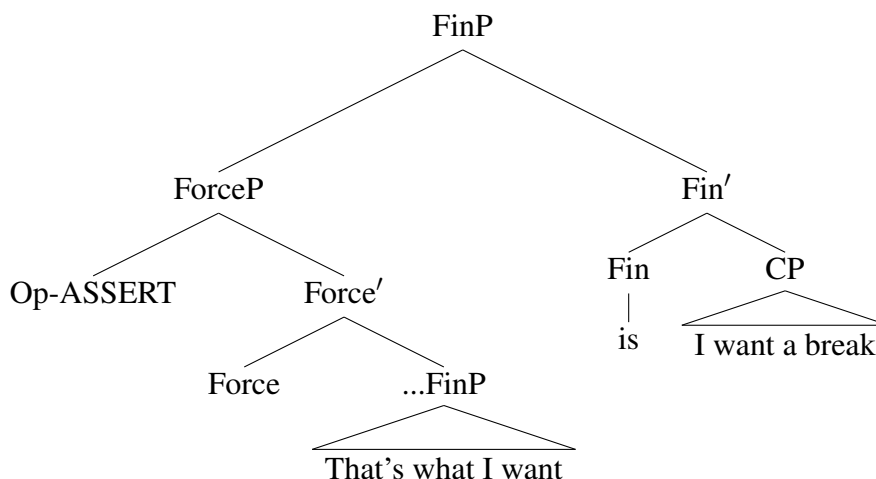
This structure predicts the absence of Force in the amalgam copular clause, but it does not

predict the fact that a single instance of Force can scope over the whole proposition. Recall that in

that's x is y, when *that's x* is a polar question, interrogative Force associates with the counterweight

(*y*) as well. ((117) repeats (106)).

(117)   Is that the best way to get off of that stuff? Is just to sleep it off?

Haddican et al. (2014) argue that in dual-question slifts, a matrix Force operator binds both propo-

sitions.

(118)   How did she do it? did she say?

Sentences like (118) can exhibit binding and scope interactions indicating that they are integrated

(i.e., not two separate sentences), but they are bound by one Force operator, and hence, are both

interpreted as questions.

An alternative to (116) suggested by these observations is that it is Fin, not Force, that is absent

in the copular amalgam clause. The matrix Force operator binds the propositions expressed by

the weight and counterweight clauses in *that's x is y*. The copular clause projection includes no

instance of Fin, so it cannot be bound by Force. It is likely that these sentences have the base-generated Topic-Comment structure proposed for weight-initial pseudoclefts. The copular clause itself is the matrix clause, but it projects from Top, rather than Fin (119). In this respect, it is more "stunted" than the clausal spine of the reversible amalgam pseudocleft.

(119)
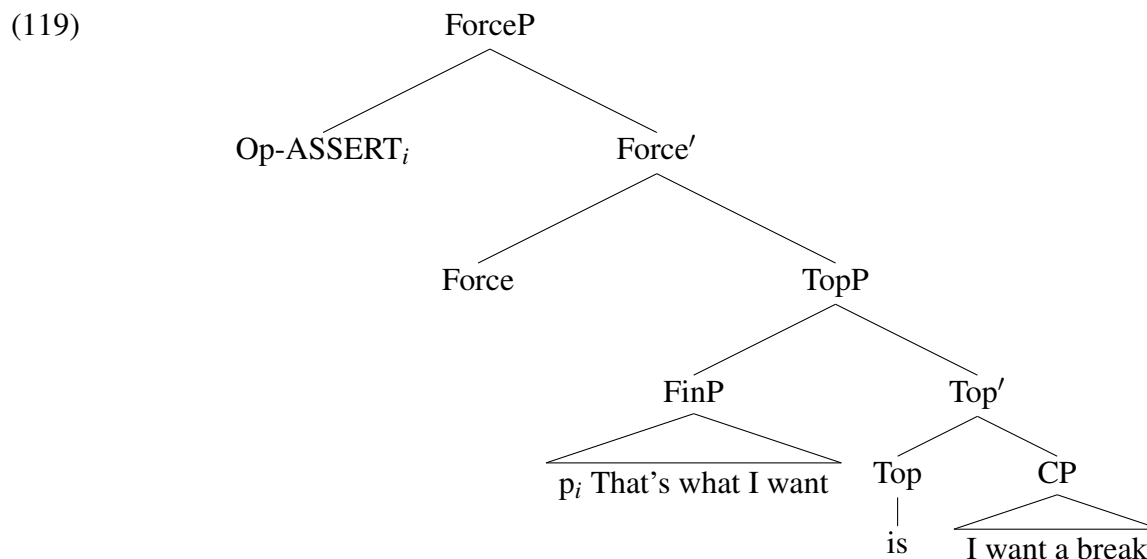
Base-generation of the weight clause in a Topic position correctly rules out the ungrammatical predicate inversion structure where the counterweight clause precedes the copula.

(120)   *I want a break is that's what I want.

Just as Haddican et al. (2014) propose for slifting, the precopular weight clause must be in the left-peripheral position in order to be bound by matrix Force. I label the copula's position Top, because of the information structural profile of the sentence types, but the more generic : (Colon) is just as appropriate. I am not aware of any explicit test to determine the categorial identity of this position.

5.4.3.3.1   Bare Topic-Comment amalgams

Now consider question-answer and free-*be* amalgams, (121) and (122).

(121)   You know who I saw? is I saw John.

(122)   Our kids are great on vacations, but when we come back, is they need to play. (Massam

1999)

The functional role of the copula in these two amalgam types is even more limited than it is in the

pseudoclefts and *that's x is y*: it is more like a coordinator and less like a verb. While question-

answer amalgams must be specificational, in free-*be*, the relation can be much looser. It is a type

of non-logical predication. In (122), the copula simply mediates between a Topic and a Comment.

Unlike in *that's x is y*, the weight and counterweight in question-answer and free-*be* amalgams

can have distinct illocutionary force. In these cases, therefore, I propose that the copular clause

includes only a projection of Top, and both the weight and the counterweight are fully specified

for Force.

5.4.3.4   Summary

The parallels between coordination and specificational amalgams are intriguing. The range of

copular amalgam constructions shows that like the cross-categorial predicational RP, the cross-

categorial :P (specifying coordination phrase) has special properties when it relates sentences.

When a root-like sentence is related to a non-root indirect question or concealed question DP, the

copula is merged in Fin, and projects a fully specified context domain, including an illocutionary

Force operator. When a root-like sentence is related to another root-like sentence, however, the

structure projected from the copula's position has more properties of coordination and fewer of

finiteness. The clausal spine can thus be more or less finite, depending on how much structure in

the context domain is projected. Crucially, structure can be absent at *either end*: an embedded finite

clause can be Finite but fail to express Force, or a matrix finite clause can have Force but lack Fin.

5.4.4 Double-*is*: two sources

In this section, I extend the Topic-Comment model to the remaining copular amalgam type, double-*is*.[42] In double-*is*, two occurrences of the finite copula (cop_1 and cop_2) occur in a specificational sentence, often a weight-initial amalgam.

(123)   a.   The issue **is, is** we're out of money.

         b.   What I'm worried about **is, is** we're out of money.

         c.   That's my concern **is, is** we're out of money.

         d.   You know what I'm worried about **is, is** we're out of money.

While double-*is* has been analyzed in previous literature (Jehn 1979; McConvell 1988, 2004; Massam 1999, 2013; Andersen 2002; Coppock and Staum-Casasanto 2004; Brenier and Michaelis 2005; Coppock et al. 2006; Curzan 2012), its relationship to and co-occurrence with the other amalgam sentence types has not yet been accounted for.

There are a number of syntactic similarities between true double-*is* and copular amalgams, which has suggested a superficially attractive analysis where the misbehaving copula in both sentence types is the same (Massam 1999, 2013; McConvell 2004). I will show here that double-*is* cannot be completely assimilated to one of the structures proposed in the preceding discussion. The most obvious reason for this is that double-*is* actually co-occurs with weight-initial amalgams, an observation that is original, as far as I know. (124a) is a naturally occurring example of double-*is* in *that's x is y*, and (124b and c), with double-*is* in the amalgam pseudocleft and the question-answer amalgam, are loosely based on it.

(124)   a.   [That's the key] **is, is** [you have to basically give up your life to save their lives][43]

---

[42]Much of this section is based on O'Neill (to appear).

[43]http://sixtyminutes.ninemsn.com.au/article.aspx?id=1026247

b.   [What you have to do] **is, is** [you have to give up your life to save their lives]

c.   [You know what you do] **is, is** [you take on a lot of risk]

Double-*is* sentences must therefore involve some additional structure.

5.4.4.1   True double-*is*

There are two potential sources for double-*is*. The first I will refer to as "true" double-*is*. These true double-*is* sentences are Topic-Comment structures (see also Brenier and Michaelis 2005). Cop$_1$ is a high (hanging) topic-marker, and cop$_2$ is the "colon", heading the lower Top position. The portion of the structure corresponding to the projection of Top: is identical to the other non-reversible amalgams. The Topic phrase (the weight) licenses a coreferential empty topic that is specified by the focused counterweight clause.[44] The tree in (125) illustrates this structure.

---

[44]This empty topic is base-generated in the lower Topic position, so it is not like a moved empty topic associated with a gap in a topic-drop construction, e.g., in other Germanic languages.

(i)   e$_i$ heb ik al   t$_i$ gelezen
        have I   already   read
      Context: "Are you familiar with this book?"

Thanks to Marcel den Dikken for pointing out the difference between the two topic types, and for providing the Dutch example above.

(125)

```
                        ForceP
                       /      \
              Op-ASSERT       Force'
                             /      \
                          Force      TopP
                                    /     \
                                 CP_i      Top'
                           /          \    /    \
                    The issue/What he did  Top   Top:P
                                           |     /    \
                                           is   e_i    Top:'
                                                      /      \
                                                   Top:       CP
                                                    |        /   \
                                                    is    he left early
```
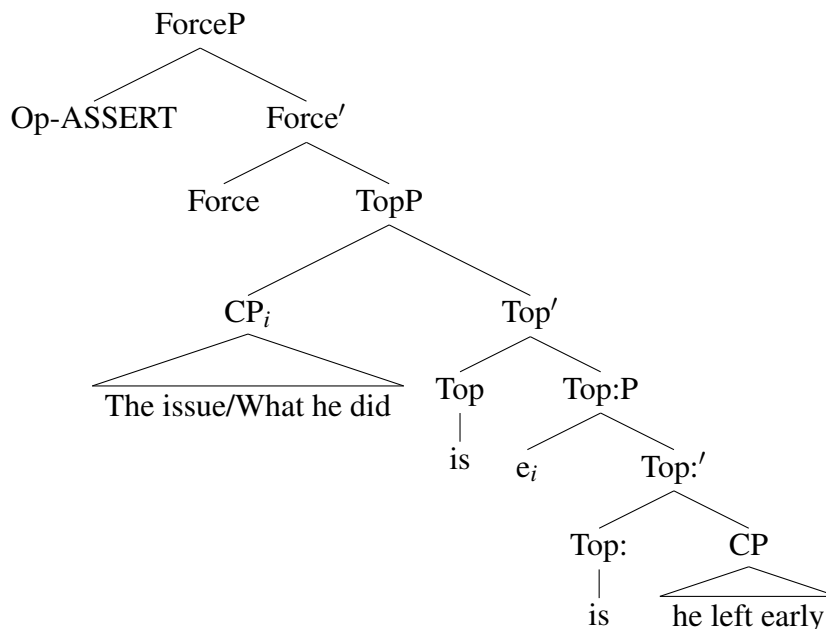
A variant of true double-*is* includes the predicate-inversion derivation of the amalgam pseudocleft. In this case, the empty topic is the underlying predicate of a FinP small clause. In its derived position in the specifier of Top:+Fin, it Agrees with [ɸ] on Fin, which licenses it in the absence of phonological content.

The left-peripheral position of the two copulas in true double-*is* is confirmed by the robust syntactic constraints on the construction. Double-*is* sentences and the other amalgams show the same empirical profile. For example, cop_1 and cop_2 must generally be in simple form.

(126)   a.   *The issue **isn't, is** he left early.

   b.   *The issue **is, isn't** he left early.

(127)   a.   *The issue **will be, is** he left early.

   b.   *The issue **is, will be** he left early.

The copulas take simple form and are incompatible with functional material like negation and auxiliaries because they are merged directly in the context domain. The double-*is* clause lacks the

T and V-domains.

Double-*is* sentences are also frozen: movement of or out of the two major constituents is impossible.

(128)   a.   A: *What$_i$ do you think t$_i$ is, is we're hungry?

         b.   *What$_i$ do you think the problem is, is t$_i$?

(129)   a.   *[How big]$_i$ do you think [the t$_i$ problem] is, is that he forgot his lunch?

         b.   *What$_i$ do you think the problem is, is that he forgot t$_i$?

The structural frozenness of these sentences follows from the fact that the major constituents occupy information structurally marked positions in the left periphery of the clause. Movement out of these positions is normally impossible, and even if it were permitted by the syntax, it would assign the moved constituent two different interpretations, leading to an incoherent information structure. Double-*is* sentences are also unembeddable in non-bridge contexts, which is expected since they include a rich left periphery.

The Topic-Comment structure of the double-*is* sentence differs from that of the other weight-initial amalgam types, because the first copula in the string occupies a high Top position, while the second one occupies the lower Top: position. This analysis differs from the previous literature in treating neither cop$_1$ nor cop$_2$ as an ordinary verbal copula in T. Crucial for the present proposal is the fact that all of the amalgam sentence types support double-*is*. The counterweight-initial amalgam pseudocleft is the noticeable exception: it is the only amalgam sentence type where the focus rather than the topic occurs in precopular position. Compare (130) and (131).

(130)   That's much the problem with reporting on Pakistan **is is** those short news items...

         (COCA)

(131)   *He forgot his lunch **is, is** the problem.

Because cop_1 is a topic-marker, it is incompatible with the focused counterweight.

5.4.4.2 Evidence for the structural status of the copulas

The left-peripheral position of cop_2 (as the spell-out of either Top: or Top+Fin) is supported by an analysis of 85 million words of spoken English from Corpus of Contemporary English (COCA, Davies 2008–), which yielded 514 double-*is* sentence tokens.[45] Chapter 4 presents the details of the corpus study (see also O'Neill to appear). Of these 514 sentence tokens, I identified only two where cop_2 was non-bare:

(132)   a.   In my mind the measures of progress **are are** you train the Iraqis to take control of

the situation.

b.   And his whole stance **is has been** to just say basically things will get better.

In the plural token, it follows an occurrence of cop_1 as *are*. It is impossible to tell simply by reading the transcript whether this lone example is a disfluency. Informal acceptability surveys with additional functional material on cop_2 showed that it must indeed remain bare. The incompatibility of cop_2 with functional material is consistent with its position in the context domain of the clause.[46]

The status of cop_1 is less clear. The corpus search yielded 31 tokens where cop_1 was in a finite non-bare form, and 29 tokens where it was in non-finite form (*being*).[47] The existence of such

---

[45]Tokens where double-*is* occurred in combination with an amalgam pseudocleft or *that's x is y* were excluded from this count and analyzed separately. These tokens all included simple *is, is* sequences, as in (130).

[46]In this component of the analysis, I am in agreement with Massam (2013).

[47]The use of *being* with proposition-denoting topics has an unclear syntactic status in English. It may be a variant of bare *is* as a simple topic-marker, but it may also derive from a reduced *with*-absolute adverbial modifier. It occurs outside of the double-*is* context, and co-occurs with double-*is*. The following examples are based on what I have observed in the wild.

(i)   My issue being, I don't know how we'll get there on time.

(ii)   The thing being is, is I would rather not leave so late.

forms casts doubt on the structure in (125): how can a plain topic-marker support extra functional material? The answer is in fact that it cannot. These forms show that true double-*is* cannot be the only source for strings of apparently superfluous copulas. Previous accounts of double-*is* (e.g., Massam 1999, 2013) have analyzed cop_1 as an ordinary verbal copula, which is compatible with these problematic forms, but I will now present evidence that true double-*is* is in fact the more basic derivation for double-*is* sentences.

The interaction between the category of precopular constituent and the morphological form of cop_1 provides key insight into this issue. When cop_1 behaves as an ordinary verb (thus admitting extra functional material), the precopular constituent must be consistent with a reduced specificational pseudocleft structure; that is, the precopular constituent must be a licit structural subject of that verb.

(133)   [The only difference **could be**] **is** that you want to give more state control...   (COCA)

An instance of cop_1 with extra functional material, i.e., one that is part of an ordinary verbal projection, cannot occur between the weight and counterweight of an amalgam. Co-occurrence of double-*is* with an amalgam does not make available any ordinary verbal position for cop_1, so the present analysis predicts that when cop_1 occurs in an amalgam, it must take a bare form.

The corpus findings support this prediction. Whenever the precopular weight expression was not an ordinary DP, the copula was bare. This evidence is suggestive, but the relatively low frequency of non-DP-initial double-*is* in the corpus (63 tokens, including *that's x is y*), makes it inconclusive. Is the amalgam + functional material structure merely unattested, or is it actually ungrammatical?

In Chapter 4, I presented an acceptability survey that supports the latter possibility. This survey compared the single copula in an amalgam pseudocleft (independently shown to resist extra

functional material) to cop_1 in the corresponding double-*is* version of the pseudocleft, e.g., (134).

(134)   a.   What he likes is he likes pizza.

            b.   What he likes is, is he likes pizza.

            c.   What he likes could be he likes pizza.

            d.   What he likes could be, is he likes pizza.

The results were conclusive: the presence of extra functional material like the modal significantly reduced the acceptability of both sentence types ($p < 0.01$). In addition, the effect on both sentence types was similar; that is, there was no interaction between the double-*is* condition and the single-*is* condition with respect to the effect of adding material to the copula. These results show that in conjunction with amalgams, only the true double-*is* structure is available, as expected.

5.4.4.3   Faux double-*is*

For the DP-initial double-*is* sentences that permit extra functional material, I propose a structure analogous to Topic-Comment amalgams with single-*is*. A specificational copular clause serves as the weight (just as in *that's x is y*). In this context, cop_2 is the amalgam copula, the head of Top:. Cop_1 is simply an ordinary verbal copula inside the weight clause. The value inside the weight clause—the counterpart of *that*—is a null element. So far, this proposal is very much in line with the structure Massam (2013) develops for *all* double-*is* sentences and amalgams. She treats cop_2 as an Appositive head (Appo; functionally identical to : "Colon"), an applicative element that takes an "anchor" in its specifier (a null category bound by the precopular DP or CP) and the focal postcopular constituent (which specifies the content of the anchor) in its complement.

(135) $[_{IP} [_{DP_i}$ the thing$] [_{I'} [_I] [_{VP} [_{V'} [_V$ is $] [_{AppoP} [e_i] [_{Appo'} [_{Appo}$ is $][_{CP}$ that I like you$]]]]]]]$

As we have seen, this structure cannot account for *true* double-*is*, since if cop$_2$ occupies Appo, there is no structural position for cop$_1$ when double-*is* occurs with an amalgam sentence type. Massam's (2013) applicative appositive structure admits a position for one "intrusive" copula, but not two.

I assume that the appositional head Top: is like a sentential connective, unlike the Appo model above. It must relate propositional elements. Since the anchor in the weight clause (the expression introducing the variable) need not be adjacent to the copula, this analysis can straightforwardly account for examples like (136), where they are non-adjacent.

(136)   Okay, so maybe I should write about [that] in the first paragraph, is [the fact that the biological parent is the only one who has standing].

Given the sentential-connective structure, the faux double-*is* copular clause is the same as *that's x is y*, (137).

(137)
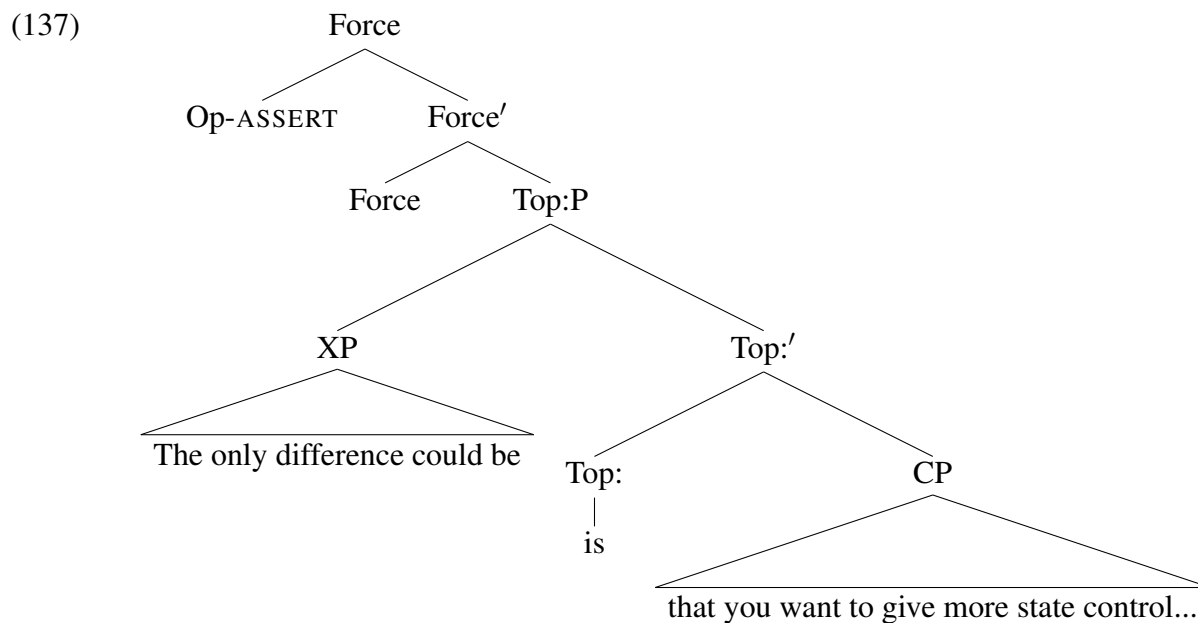
In faux double-*is* sentences with a truly verbal cop$_1$, the associate of the counterweight is a null operator occupying a position above the fronted predicate, e.g., *the thing* in (135).[48] The proposed structure for the weight clause in sentences like (133) is given in (138).

(138)

```
                        CP
                      /    \
                   Op_i     C'
                          /    \
                         C      TP
                              /    \
                          DP_i      T'
                         /    \    /  \
                    One reason  T    VP
                               |     |
                            could    V'
                                    /  \
                                   V    LP
                                   |   /  \
                                  be  t_k   L'
                                         /    \
                                      L+R      RP
                                              /  \
                                           e_i    R'
                                                 /  \
                                               t_R   DP
                                                      |
                                                     t_k
```
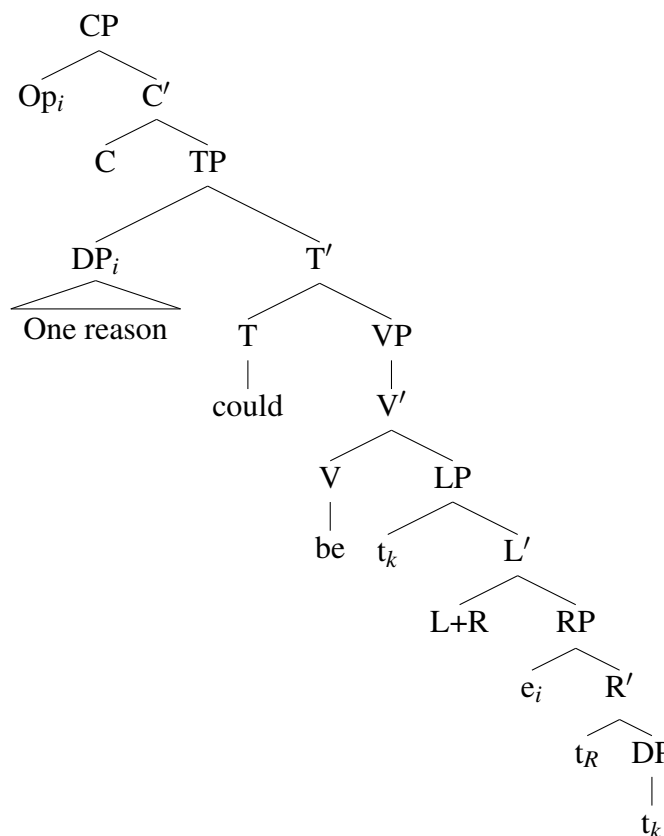
In most contexts, null operator movement over a topic is barred, as it would yield an information structure clash, but it is not necessarily ill-formed. Extraction of an *overt* operator across a fronted predicate brings about a clash, since the operator corresponds to a focus in its base position (as the value of a specificational sentence), but is associated with a topic interpretation in its landing site. A *null* operator receives no local information structural interpretation, so it avoids this problem. The null operator must, however, receive content from somewhere. If the constituent providing its content is a topic, once again, a clash arises, since the null operator is syntactically associated with the focus of specification:

---

[48]Thanks to Marcel den Dikken for suggesting this line of analysis.

(139) a. #The burgers$_i$ Op$_i$ that [what he ate] is e$_i$ were greasy.

  b. #Those burgers$_i$ are hard to believe Op$_i$ that [what he ate] is e$_i$.

In the weight clause of an amalgam, by contrast, the operator is identified only with the content of the postcopular constituent, which is a focus. The operator therefore receives one coherent information structural interpretation in the derivation: it is the focus of its local specificational clause, and its postcopular associate is the focus of the larger specificational structure. This recursive specificational structure is more complex than the *true* double-*is* structure, where the precopular constituent is a base-generated Topic. I speculate that, although all double-*is* sentence strings with simple DPs or free relatives in precopular position are ambiguous between the recursive specificational structure and the more direct Topic-Comment structure, the latter is preferred, and thus extra functional material occurs infrequently.

The more common true double-*is* structure, where cop$_1$ and cop$_2$ occupy the same spine, easily allows intervening modifiers.

(140) The thing is, however, is that he's truthfully the most efficient offensive weapon in the game right now and this is backed up by a pretty damn good measure.[49]

In this respect, it is much like the multiple *que* topic-marking clauses found in Spanish (and Italian, with *che*), where a finite complementizer occurs in multiple left-peripheral structural positions, encoding related information structural functions.

(141) Le   gritó   que qué mala cara que tenía.
       CL.3SG shout.3SG.PAST that what bad   face that have.3SG.IMPF
       'He shouted at her how awful she looked.'   (Demonte and Fernández-Soriano 2009:

       (22))

---

[49]http://www.slipperstillfits.com/2015/1/14/7542791/a-statistical-look-at-kevin-pangos-absurd-senior-season

With an intervening modifier, this doubling is very common.

(142)   Le   dije   que a su madre que le   dara   un beso.
        CL.3SG say.1SG.PAST that to her mother that CL.3SG give.3SG.PST.SUBJ a kiss
        'I told her to give her mother a kiss.'

The fact that these two constructions, double-*is* and double-*que*, are exempt from any haplology

filter suggests the presence of intervening phonological material. In (140), this material is overt. In

examples without the intervening modifier, it is the empty topic that blocks haplology.

   A further prediction of the present proposal deriving double-*is* strings from two different struc-

tures is that the *true* double-*is* structure should be able to contain the other, resulting in a string

of *three* copulas. Indeed, three copular forms can occur in a row (143). Such examples have been

noted in the previous literature, but not incorporated into a unified analysis.

(143)   a.   The fact **is is, Howie, is** that with a quote like that, you press the subject...

             (Zimmer 2011)

        b.   [The problem **remains**] **is**$_{cop1}$, **is**$_{cop2}$ that I haven't eaten.

        c.   ...[One reason **could be**] **is**$_{cop1}$, **is**$_{cop2}$ that I'm as addicted to him as he is to his

             DOC.[50]

5.4.4.4   Summary

To summarize this section, both cop$_1$ and cop$_2$ in true double-*is* are simple functional heads. Sen-

tences in which cop$_1$ occurs with baggage actually have a different structure, where cop$_1$ is a verb.

They share their basic structure with copular amalgams, where cop$_2$ occupies a Top head with

Colon function, and no special topic-marker is projected. By teasing apart these two sources for

double-*is* strings, I have accounted for the puzzling properties of the double-*is* construction, and

---

[50]http://www.soberrecovery.com/forums.html

its co-occurrence with other copular amalgam sentence types.

5.4.5 A cline of spines

This section took as its starting point that observation that the acceptability of the different copular amalgam sentences types varies across speakers. There is no simple implicational hierarchy that predicts which speakers will accept which sentence types. Focusing on the interactions of the sentence types with the main hallmarks of the context domain—finiteness morphology, topic-focus structure, and illocutionary force—I proposed that the counterweight-initial amalgam pseudocleft is the most complete instantiation of the unusual Fin-headed clause type that is the focus of this dissertation. It includes a fully specified left periphery, from Fin to Force. This structure is also the rarest cross-linguistically, occurring only in some Englishes, as far as I am aware.

The weight-initial amalgam pseudocleft showed mixed properties. For some speakers, it shows the properties of both Fin and Force, e.g., agreement with the structural subject and independent illocutionary Force. For these speakers, I proposed that the weight-initial pseudocleft is derived from the application of predicate inversion and domain-extending head-movement in the counterweight-initial pseudocleft. The copula in this structure is obligatory. Other speakers have a weight-initial pseudocleft with more restricted behavior. In their grammar, the weight-initial pseudocleft is a base-generated Topic-Comment structure, where the amalgam copula heads a low Topic projection. These pseudoclefts lack independent illocutionary force, and show a much more restricted copula.[51]

Considering the appositional function of the copula and the bisentential nature of the non-reversible amalgam sentence types (i.e., *that's x is y*, the question-answer amalgam, and the free-*be* amalgam), I extended the model of "specifying coordination" to copular sentences. In particular,

---

[51]Some speakers probably have both structures.

the weight-initial copular amalgams with root-like weight clauses lend themselves to an analogy with coordinate structures. Although the copula in these sentences is morphologically finite, suggesting some association with a finite left periphery, it cannot bear independent illocutionary force. The spine of the copular clause in this case includes Top: and Force. The illocutionary operator in its highest specifier binds the proposition in the weight clause.

Lastly, double-*is* sentences provide indirect support for this structure, since sentences where cop_1 shows ordinary verbal behavior are consistent with cop_2 in the role of Top:—its behavior is the same even in true double-*is* sentences where cop_1 cannot possibly be parsed as an ordinary verb. Double-*is* structures show that the English copula is quite ubiquitous in the functional spine of the clause. Just as in Romance languages, where the finite complementizer (e.g., Italian *che*, Spanish *que*) can multiply in left-peripheral positions, occurrences of the finite copula can be found in Top and Fin. It is uncontroversial that the copula can occur in T. We saw in counterweight-initial amalgams that it can occupy Fin; weight-initial amalgams showed that it can occupy a low Top position; and true double-*is* shows that it can occupy a high Top position as well. Table 5.4 presents this inventory of spine types, from most finite, to least.

Counterweight-initial	[Force			[Fin
True double-*is* 1	[Force	[Top	[Top:	[Fin
Inverse weight-initial	[Force		[Top:	[Fin
True double-*is* 2	[Force	[Top	[Top:	
Base-generated weight-initial	[Force		[Top:	
That's x is y	[Force		[Top:	
Question-answer			[Top:	
Free-*be*			[Top:	

Table 5.4: Gradient finiteness in context-domain spines

5.5 Conclusion

This chapter has offered a syntactic analysis of the copular amalgam sentence family. Given the amalgam copula's inability to combine with functional material from the displacement/T-domain of the clause, and its Force and Finiteness functions, I proposed that the copula in the amalgam pseudocleft is base-merged in Fin, where it relates the propositional weight and counterweight directly. Copular amalgams are thus finite small clauses.

Taking a closer look at differences in the distribution of agreement and illocutionary force in the different copular amalgam sentence types, I proposed that the "spine" of the copular amalgam clause comes in a variety of sizes. The fully finite version, the reversible amalgam pseudocleft, includes both Fin and Force, while the least finite version includes only Top. As the structure decreases in its finiteness characteristics, it increasingly resembles coordinate structures. The next part of the dissertation focuses on the fully finite amalgam pseudocleft, examining the implications of the proposal that a clause in a tensed language can be fully finite in the absence of the displacement domain.

PART 3: Finiteness without Tense

We're right here right now, is my point.

Chapter 6

Finiteness from the top down

6.1   The core of finiteness

How can a clause consisting only of the left-peripheral domain be fully finite? This question guides Part 3 of this dissertation, which shows that the phenomena associated with finiteness are all instantiated in the left periphery, so the unique properties of copular amalgam clauses turn out to be precisely what we expect to find.

In Minimalist clausal architecture, finiteness originates in the left periphery—the context domain. Finiteness is a morphosyntactic category, a feature specification of a head Fin(iteness) (see, e.g., Rizzi 2004; Adger 2007; Nikolaeva 2007), but its presence is associated with a constellation of morphological, syntactic, and semantic phenomena, including:

(1)   Phenomena of finiteness

    i.   tense and agreement marking on the verb

    ii.   instantiation of an opaque syntactic domain (Chomsky 1986, 1995, 2001)

    iii. licensing of an independently referential or nominative case-marked subject (see, among others, Bianchi 2003; Landau 2004; Adger 2007; Sigurðsson 2004; Biswas 2013; McFadden 2013)

    iv. independent anchoring of the proposition to the utterance context (e.g., Bianchi 2003;

McFadden and Sundaresan 2013; Amritavalli 2013; Kissock 2013; Ritter and Wiltschko

2014)

Much research is dedicated to examining how the phenomena of finiteness in (i)–(iv) above can be traced to a syntactic component of clausal architecture. The puzzle is that properties (i)–(iii) are subject to considerable cross-linguistic variation; they interact with language-specific morphological and lexical idiosyncrasies. Property (iv), however, is universal: all languages have a way of anchoring clauses to the utterance context. The anchoring property is what allows us to express propositions about events and individuals that are displaced from our immediate experience—this property of language is fundamental to the linguistic and cognitive systems. Moreover, a clause that is anchored can have illocutionary force. Finiteness at its core consists of these two components: anchoring and force.

Anchoring is universal (Enç 1987; Bianchi 2003; Sigurðsson 2004; Wiltschko 2014), but its specific relationship with verbal inflection, opaque domains, and subject licensing varies within and across languages. Moreover, anchoring does not entail independence. A clause can have independent tense, for example, but still be formally subordinate.

It is uncontroversial in Minimalism that lower domains of structure do not imply higher ones. The clause can be "truncated" at various levels of the verbal projection, with the result that the more truncated the structure is, the less finite it is (Rizzi 1997; Grohmann 2003; Adger 2007; McFadden 2013, etc.). Each layer of structure in the clausal spine contributes a different formal and functional property, so the degree of finiteness exhibited by a given clause is composed of whichever of these properties are present. The structure I proposed for copular amalgam sentences in Chapter 5 departs from the traditional model, in that higher domains of structure do not imply the presence of lower ones.

This section of the dissertation shows that the core finiteness properties can compose from the top-down. This is not unorthodox in the Minimalist framework, particularly the phase-based model of Chomsky (2001, 2008), Richards (2007), and others, which argue that although structure-building proceeds via merge, which is strictly bottom-up (respecting the Extension Condition, and cyclic spell-out), there are also relations that proceed top-down, like Feature Inheritance and Agree. Just as in the bottom-up characterization of clausal finiteness, different levels of truncation from the top-down result in different expressions of properties (i)–(iv). Nothing rules this out in principle, but the current theories of clause structure have not pursued this possibility, I suspect because they have not found empirical evidence for it.

The contribution of this thesis, and Part 3 in particular, is to show how the copular amalgam provides exactly the evidence we need. The challenging properties of the copular amalgam clause follow if finiteness can be constructed from the top down—with the C-domain independent of the T-domain.

6.2 Gradient independence

Let us briefly consider the most basic component of full finiteness: independence. Independence is associated with the presence of both Force and Fin in the context domain of the spine, where formal distinctions allow some clauses, but not others, to serve as independent utterances. In English, (2a) but not (2b) can be independent of superordinate syntactic or semantic structure. Not coincidentally, (2a) includes a tensed, agreeing form of the verb, and (2b) includes a non-tensed, non-agreeing form.

(2)　　a.　He runs.

　　　　b.　*Running.

The traditional assumption is that the independence of a clause is related to the presence of tense and agreement morphology.

If only the case were this simple. First of all, it is unclear whether the independence/dependence distinction is binary. For example, in English, can the following clauses be partitioned cleanly into two groups with respect to their independence?

(3) a. He leaves early.

   b. But he left.

   c. Leave!

   d. It was imperative that he leave.

   e. Why leave early?

   f. Him, leave early?

   g. If only he had left earlier!

   h. That he left

   i. For him to leave early

   j. His leaving early

   k. Him leaving early

   l. Him leave

Most linguists would draw the line either between (d) and (e), between (f) and (g), or between (g) and (h). A prescriptivist, on the other hand, may raise her eyebrows at (3b), because the first conjunct is missing. Even the minimal structure in (l) has the form of a so-called Mad Magazine sentence, which can be used in isolation with interrogative force. One might reasonably quibble with these choices, in the absence of more clearly articulated criteria.

What about the (morphological) finiteness–independence connection? Not surprisingly, natural language does not offer a clear-cut surface morphological distinction on the verb or auxiliary that signals clausal independence. If it did, we would have the following two groups:

(4)   Independent/inflected: (a), (b), (g), (h)

(5)   Dependent/uninflected: (c), (d), (e), (f), (i), (j), (k), (l)

This grouping seems to be on the right track, but it is far from perfect. (5a) is unambiguously an independent clause, and its verb is fully inflected for tense and agreement. The other clauses with inflected verbs, however, are introduced by functional elements that signal dependent status.

(6)   a.   Independent, inflected: (a)

        b.   Dependent, inflected: (b), (g), (h)

The uninflected group, (5), is also heterogeneous with respect to dependence. (6c), for example, is imperative. It is functionally independent, but it fails to express distinctions for tense or person. Although lexical verbs in English use the bare form for second person, the copula inflects; since the imperative form of the copula is the non-agreeing *be*, imperatives lack person agreement. (6d)–(6f) likewise lack tense and agreement inflection, but like (6c), they can be uttered in isolation. Whatever makes them independent, it isn't morphological tense, and it may not be semantic tense either.

(6g) is introduced by a subordinating element, but it may well be uttered in isolation, with perfectly comprehensible propositional content. (6h), likewise, is introduced by a subordinating element, but although its verb form is fully inflected, it is unambiguously dependent. The others are both uninflected and clearly dependent.

(7)   a.   Independent, uninflected: (c), (d), (e), (f)

       b.   Dependent, uninflected: (i), (j), (k), (l)

A binary distinction relying on an implicational relation between morphological finiteness and intuitive independence cannot be maintained, even for English. When other languages added to the mix, the pattern becomes even more murky.

6.3 A brief history of ... time?

6.3.1 From inflection to Infl

Traditional grammatical description conceptualized finiteness purely as a morphosyntactic distinction expressed on the verb: a verb inflected for person and tense is finite, and a verb that is not so inflected is non-finite. This classification correlates robustly with other distinctions in the well-studied languages of Europe. One of the most widely studied of these distinctions in generative grammar is the distribution of case-marked subjects. A clause with a finite verb form licenses a nominative case-marked subject, while a non-finite verb form does not. The following simple examples from (Standard) Spanish illustrate the contrast.

(8)   a.   Ellos   salieron   ayer.
             they.NOM go.out.3SG.PST yesterday
             'They went out yesterday.'

       b.   Es bueno (*ellos)   salir   de vez en cuando.
             it.is good   (*they.NOM) go.out.INF of time in when
             It is good (*they) to go out from time to time.

The second key correlate of morphological finiteness is the availability of an independent temporal interpretation. The tense of a finite verb is interpreted with reference to the utterance time (or with reference to the superordinate event in finite embedded clauses), while the tense of a non-

finite verb receives a dependent interpretation.[1] It is anaphoric to the tense of the superordinate verb.

(9) a. I heard that she left.

        LEAVE-time<utterance-time

    b. I wanted her to leave.

        LEAVE-time=WANT-time

The model of finiteness based on the morphological form of the verb leads to some useful observations, but it turns out to be insufficient in several respects. First of all, it fails to describe languages where verbs do not inflect for tense and/or agreement. Many such languages neverthe-less have oppositions between clauses that license subjects and clauses that do not. In Telugu, for example, the distribution of nominative subjects does not correlate with tense and agreement morphology on the verb (Kissock 2013). The two sentences in (10) contrast with respect to "finite-ness": (10a) is independent and has a nominative case-marked subject, while (10b) does not. The contrast is not expressed by tense or agreement marking, however.

(10) a. Atanu   i n tiki   ve l la li.
        he-NOM house-DAT go-OBLIG
        'He should go home.'

    b. *Atanu   i n tiki   ve l li.
        he-NOM house-DAT go-ABS
        'He go home.' (Kissock 2013: (62))

Languages like Chinese, which lack tense inflection, also contrast clauses that are anchored and clauses that are dependent.

---

[1]In traditional grammatical description, non-finite verbs are assumed to be tenseless. My characterization in terms of anaphora is anachronistic.

Since the 1980s (Chomsky 1981, 1986; Pollock 1989), generative syntax has distinguished between the morphological finiteness of the verb and the structural finiteness of the clause, as realized by a functional head. The theory moved away from an idiosyncratic phrase structure rule specific to the clausal category S toward an endocentric model of the clause, whose head is the functional category Infl.[2] In this model, the source of finiteness is the feature composition of Infl, not the inflection borne by a lexical head V.

Identifying Infl as the head of the clause allowed the theory to make significant strides in explaining the distribution of overt nominative DPs in subject position, and to formalize the theory of raising and control (e.g., Chomsky 1981, et passim; Lasnik and Saito 1991; Hornstein 1999; Culicover and Jackendoff 2001; Landau 2004; among many others). A DP in a feature-checking relationship with an Infl head that is positively specified for strong [Agr] bears case, so it can be overt and referentially independent, bearing nominative Case.

(11)   a. He_i hopes that he_i/*e_i will win.

      b. He_i hopes *he_i/e_i to win.

In a non-finite clause, where Infl lacks strong [Agr], nominative case is unavailable, so the subject depends for its licensing on a superordinate clause. Depending on the selectional properties of the higher verb, the subject of the non-finite clause either undergoes A-movement (raising) into the higher clause, (12), or it is an anaphoric pronoun PRO in a referential dependency with an argument in the higher clause (control), (13).

(12)   a. $John_i$ seems [e_i to be e_i a nice guy].   raising to subject

      b.   They proved $John_i$ [e_i to be e_i guilty].   raising to object

---

[2]In Pollock's split-Infl model, Tense and Agreement project separately.

(13)   a.   John$_i$ wants [PRO$_i$ to e$_i$ leave].   subject control

         b.   They forced John$_i$ [PRO$_i$ to e$_i$ leave].   object control

In (12), the empty category in the non-finite subject position is a trace, and in (13), it is a special

anaphoric pronoun PRO. The Government and Binding-era theories of Case and Control make an

explicit connection between the binary specification of a clausal functional head, the morphological

specification of the verb, and the availability of subject licensing.

6.3.2   [+Agr] $\longrightarrow$ [+nominative]?

Unfortunately, this version of Case Theory does not take long to crumble. Even in languages where

the morphological form of the verb transparently reveals the feature composition of Infl, neither

tense, nor agreement, nor both together can be identified as a stable predictor of the distribution

of nominative vs. non-nominative subjects (Raposo 1987; Bianchi 2003; Landau 2004; Szabolcsi

2009; McFadden 2013, etc.). The verbs in the bracketed clauses below show different combinations

of "finiteness" morphology associated with Infl, and yet all of them allow a nominative case-

marked subject.

(14)   Era importante [eles   sairem].
         was important   [they.NOM leave.INF.3PL]
         'It was important for them to leave.'   European Portuguese (Raposo 1987)

(15)   [Avan   saadatt-æ saappi da-ae], naan   ve lijae poo-n-een.
         [he.NOM rice-ACC eat-INF],   I.NOM outside go-PST-1SG
         'I went out side in order for him to eat rice.'   Tamil (McFadden 2013: (2))

(16)   'That were shame unto the,' seyde sir Launcelot, '[thou   an armed knyght to sle   a
         'That were shame unto you,' said   sir Launcelot, '[you.NOM an armed knight to slay a
         nakyd man by treason].'
         naked man by treson].'
         "'That would be a disgrace on you," said sir Lancelot, "for you, an armed knight, to slay

a naked man by treason"'.'   Middle English (Garrett 2012, cited in McFadden 2013)

(17)   Senki   nem akart   [csak ő   leül-ni].
       nobody not   wanted-3SG [only he/she.NOM sit-INF]
       'Nobody wanted it to be the case that only he/she takes a seat.'   Hungarian (Szabolcsi

       2009:2)

6.3.3   FinP

The next big move in Minimalism with the rise of cartography (Rizzi 1997, 2004; Cinque 1999)

was to pin the finiteness specification not only on Infl/T, but also on a head in the C-domain.

Complementizers in many languages show a sensitivity to the presence of tense and agreement

on T, so the C-domain must participate in marking the finiteness properties of the clause. For

instance, tenseless, non-agreeing Infl in Italian is selected by the complementizer *di* 'of'. Clauses

introduced by *di* seem to be smaller than finite complement clauses, with a truncated left periphery.

For instance, they do not license fronted topics, (18a). Finite Infl on the other hand, is selected by

the complementizer *che*, which allows a more richly specified left periphery, as in (18b) (e.g., Rizzi

1997, 2004).

(18)   a.   *Credo   di   il tuo libro, apprezzar-lo   molto.
            believe.1SG.PRES COMP.INF the your book, appreciate.INF-CL much
            'I believe 'of' your book to appreciate it a lot.'   (Rizzi 1997:288 (10)–(11))

       b.   Credo   che   il tuo libro, loro lo apprezzerebbero   molto.
            believe.1SG.PRES COMP.FIN the your book, they CL appreciate.COND.3PL much
            'I believe that your book, they would appreciate it a lot.'

In other languages (e.g., Irish, West Flemish) the finite complementizer actually inflects for

tense or agreement, indicating the close syntactic relationship between the C and T-domains (Haege-

man 1992; Zwart 1993; McCloskey 1996; Rizzi 1997; Adger 2007; Chomsky 2008).[3]

---

[3]Later work on agreeing complementizers in West Germanic, e.g., Carstens (2003), Koppen (2005), and Haegeman
and van Koppen (2012), has reconsidered the idea that the [φ]-features of C agree with those of T. Although this

(19)   a.   dan-k   (ik) goan
              that-1SG (I)   go
              'that I go'

        b.   da-se   (zie) goat
              that-3SG.F (she) goes
              'that she goes'

        c.   dan-ze   (zunder) goan
              that-3PL (they)   go
              'that they go'   West Flemish (Haegeman 1992)

Alternations like those in (18) and (19) are typically binary. The finiteness distinction in the left periphery is simpler than the one encoded in Infl/T. Thus, the head Fin, at the bottom of the C-domain, is born.

Fin is inherently [±finite]. A [+finite] specification is associated with the presence of a deictic utterance context variable in Spec,FinP, while a [−finite] specification is associated with an anaphoric context variable (Bianchi 2003). Fin generally also bears [T] and [ϕ], which have counterparts on Infl/T in a dependency with the lexical verb and the subject, respectively.[4] This move of finiteness to the left periphery is crucial from the perspective of the present work.

6.3.4   Reversing the direction of the dependency

With the distribution of morphological finiteness features unable to predict the distribution of nominative case-marked and independently referring subjects, Landau (2004) suggests a paradigm shift: why do we assume that the subject is dependent on the features of the verbal projection, rather than the other way around? His observation is based on the subject licensing patterns in subjunctive clauses, and the distribution of different interpretations of controlled PRO. (Note that the licensing

---

complicates the analysis of the relationship between C and T, the important point for the present discussion is simply the connection between finiteness and the features of C.

  [4]Proponents of a split-CP model differ with respect to the division of labor between Force and Fin. In Chapter 5, I proposed that [T] and [ϕ] originate on Force, and are transmitted to a lower C-domain head under binding—to Fin if there is no intervenor.

question is now cast in terms of the subject's phonological and semantic content (overt vs. covert, independently referring vs. anaphoric), rather than its morphological form (nominative vs. null or accusative case).)

Subjunctive clauses in many languages inflect for subject agreement, but form a heterogenous class with respect to tense form, temporal interpretation, and the licensing of referentially independent subjects. It would therefore seem hopeless to set up a perfect implication between any of those features and subject licensing. Even more problematic, in Hebrew and several Balkan languages, including Greek, Albanian, Romanian, and Bulgarian, agreeing subjunctive clauses that include a separate temporal domain, analogous to English *want*-complements (20), permit *either* a referentially independent subject *or* a controlled PRO.

(20) I_i wanted today [PRO_i to t_i leave tomorrow].

The optionality is illustrated in the Hebrew subjunctives below.

(21)   a.   xole ha-amnezya$_i$ cipa   [Se-hu$_i$ yizke   be-medalya].
               sick the-amnesiac$_i$ expected that-he$_i$ will-win.3SG.M in-medal
               'The amnesiac expected that he would win a medal.'

         b.   xole ha-amnezya$_i$ cipa   [Se-e$_i$ yizke   be-medalya].
               sick the-amnesiac$_i$ expected that-e$_i$ will-win.3SG.M in-medal
               'The amnesiac expected that he would win a medal.'   (Landau 2004:825 (15))

The subject of the future-tense subjunctive clause may be either *hu* 'he' or controlled PRO. To argue that the null argument in (21b) is in fact PRO and not referentially independent pro, Landau shows that it must be interpreted *de se*, and that it supports only a sloppy reading under ellipsis.[5] Clearly, it is not morphological [T] that is behind the "tense"-case connection.

---

[5]Landau points out that since Hebrew lacks 3SG pro, it offers a clear approach to identifying environments that allow both referentially independent and controlled PRO subjects. The same environments in the Balkan languages are more difficult to identify, since they have 3SG pro—these temporally independent subjunctives allow both pro and PRO.

In fact, since independent and controlled subjects are not in complementary distribution, any attempt to predict their distribution using morphological or semantic features of the clause is doomed to fail. Landau argues that it is actually the features of the clause that need licensing, rather than the features of the subject.[6] He calls the relevant feature [R], basing the analysis on Reinhart and Reuland's (1993) distinction between [+R] and [−R] noun phrases: [+R] DPs have independent reference and inherent [φ]-features, while [−R] DPs are obligatorily bound and receive [φ]-features under binding. Overt DP and pro are inherently specified as [+R], while controlled PRO is inherently [−R].

Clausal functional heads (C and T) can receive a value for [R] in the course of the derivation, as a reflex of their [T] and [φ]-features. [T] is positively specified on C where the clause has an independent temporal interpretation, while [φ] is positively specified on C in a more stipulative set of environments.[7] This value for [R] must Agree with ("be checked by") a matching value, either on the other clausal head, or on a nominal subject. The distribution of referential vs. controlled subjects is thus only constrained by whether C, T, or both need to check an [R]-feature. When C and T lack a value for [R] (either because they lack [T] and [φ] or because they have checked each other's [R]-values), independent and controlled subjects overlap in distribution.

While this approach is not without problems (e.g., the unclear theoretical status of [R]-checking, and the sketchy evidence for the distribution of [φ] on different C-heads), it represents an important advance. It shifts the burden for the finiteness-case connection off of Infl/Fin and onto the subject DP itself. Because the temporal interpretation of the clause is a key predictor of its subject licensing properties, Landau's approach also brings the anchoring function of the clause, rather than just [T] inflection, to the forefront.

---

[6]This is simplifying slightly, since this dissertation is not concerned with the analysis of control, but rather, with the analysis of the finiteness-case connection.

[7]Landau uses [Agr] where I use [φ].

6.3.5   The external logophoric center

In this advance in our understanding of the finiteness-case connection, the spotlight turns to the semantic and functional counterparts of tense and agreement features. Bianchi (2003) marries syntactic and functional notions of finiteness, proposing that a [+fin] Fin head encodes the "external logophoric center", while a [−fin] Fin head encodes an "internal logophoric center".

(22)   "Every clause is anchored to a Logophoric Centre: a speech or mental event, with its own participants and temporal coordinates, which constitutes the centre of deixis" (Bianchi 2003: (3))

The external logophoric center is the utterance context, while the internal logophoric center is a speech event or mental event internal to the sentence, e.g., associated with a verb of saying or thinking.

If [+fin] encodes the external logophoric center, the clause's tense and event participants can be independently anchored to the utterance context. The logophoric center licenses subject agreement, which in turn licenses nominative case, by housing the anchors for the verb event participants: the speech event participants. A non-finite Fin encodes an anaphoric speech event, so the subject of a non-finite clause is likewise anaphoric.

Bianchi's insight allows us to reimagine the finiteness-case connection as something not coincidental, but deeply tied to the anchoring property of finite clauses, which is syntactically expressed at the bottom of the clausal left periphery, in Fin.

Sigurðsson (2004) takes an approach that is similar in spirit. Like Bianchi, he takes the left periphery of the clause to encode the logophoric center. The grammatical domain of the clause (the Infl domain) encodes the relation between the logophoric elements {speech time, speech location,

agent=speaker, patient=addressee} and the elements of the verb event {event time, event location, agent, patient}. I will show in this final part of the dissertation that these functional-syntactic approaches to the finiteness phenomena provide the right tools to make sense of the finiteness properties of copular amalgam sentences.

6.4 Finiteness in copular amalgams

Although finiteness is not a monolith, the core properties of finiteness can all be traced to the left periphery of the clause. The crucial elements of this domain of structure, in the current split-CP framework, are Fin and Force. The former introduces the binary specification for [±finiteness] and the logophoric context, and the latter introduces [T] and [φ] probes and an illocutionary operator binding the proposition, the context, and the speaker.

The necessary and sufficient ingredients for independent finite clause-hood are the following:

(23)   Participants

      a.   a context (a tuple {speaker, addressee, time, world, location})

      b.   a topic situation (event, state, context, etc.)

      c.   a predicate

(24)   Structure

      a.   a functional head $\text{Force}_{[uT,u\phi]}$

      b.   a functional head $\text{Fin}_{[+Fin]}$

(25)   Relations

      a.   an illocutionary force relation (proposition, context, speaker)

      b.   an anchoring relation (topic situation and context)

Despite the top-down nature of finiteness dependencies, a traditional assumption persists in Minimalist clausal architecture. This assumption is that there is a strict implicational relationship between higher and lower heads in the clausal spine/functional sequence:

(26)   A functional head F^n must combine a head F^{n-1} (Grimshaw 1991).

A more plausible, modified version of this constraint is that a functional head F^n takes a complement with the value $F^{\leq n}$.

The Extended Projections model helps to formalize constraints on selection by functional heads. A functional head C, which has the value F^2, has a limited set of complementation possibilities.

(27)   a.   C - CP

        b.   C - IP

        c.   *C - VP

        d.   *C - DP, etc.

This model allows properties of the lexical head to project upward, addressing a problem for the locality of selection. For example, without extended projections, properties of an NP would be invisible to a V selecting the DP that contains it, leading to mismatches like (28).

(28)   They counted *[the color] / [the colors].

If the Extended Projections model of the clause is right, then a clause cannot be finite unless it is the extended projection of a V.

In Chapter 5, I argued that this is not so. Higher domains are in principle independent of lower ones. Ordinary copular sentences provide evidence against Extended Projections. First of all,

the existence of pronominal copulas is problematic. Either C (V^2) must select D (N^1)—a mixed projection—involving an embedded copular clause, or the pronominal copula must be treated as an I (V^1) without a lower lexical head. Secondly, I presented evidence that in canonical copular sentences in English, T can project in the absence of a lower V. Copular amalgams provide the missing link to support what is logically the next hypothesis: C (Fin) can project in the absence of a lower T.

Not only is Fin in amalgams independent of T—the amalgam pseudocleft and DP-weight amalgams are also fully finite in the absence of the lower domains. These amalgam types exhibit all the core properties of finiteness.

(29) i.   Their verbal element, the copula, is inflected for [T] and [φ].

ii. They instantiate an opaque domain for extraction.

iii. They license an independently referring structural subject.

iv. They are independently anchored to the utterance context.

The absence of the lower domains in copular amalgams demonstrates that finiteness is not constrained by a strictly monotonic functional sequence.

Instead, we must recognize gradient finiteness from both a bottom-up and a top-down perspective. A *maximally* finite independent clause includes Force, Fin, and T. Clauses can decrease in finiteness in *either* direction. A clause may be finite in the traditional sense, but lack independent Force. In this case, it is embedded and lacks its own illocutionary content, like a verb-final embedded clause in German. A clause like an amalgam or an imperative may also be independent, but lack independent temporal reference.

The next two chapters are concerned with finiteness properties (iii)–(iv) in copular amalgams. Chapter 7 addresses the status of the root clause subject in copular amalgams. How is the structural

subject licensed in a copular amalgam, and what is its relationship to Fin? Moreover, why are independent, root-like finite clauses able to serve as structural subjects only in finite clauses that lack functional projections below Fin? The answer will pursue insights about the connection between the referential independence of both the clausal heads and their subjects, showing that the root-like clausal subject has a particular type of referential independence that English DPs lack.

In Chapter 8, I address the question of how copular amalgam clauses are anchored to the utterance context in the absence of Tense. Independent temporal anchoring in English is encoded in the T-domain of the clause, but deictic anchoring obtains in the domain of Fin, without the participation of T. The copular amalgam clause is anchored deictically to the external logophoric center, since its topic situation is the utterance context associated with Fin, rather than a topic time associated with T. Moreover, since the subject of the copular amalgam is a root clause rather than an individual-denoting expression, the utterance context (topic situation) of the amalgam is characterized by the relation between a property of a speech act (the weight) and a speech act (the counterweight). This unusual set of relations sheds light on the intuition that copular amalgams bear some relationship to quotative or free indirect discourse constructions.

Chapter 7

Sentential subjects in the left periphery

7.1 Introduction

One of the most salient properties of copular amalgam sentences setting them apart from other sentence types is the fact that their logical subject is a bare (i.e., *that*-less) finite clause. Not only does a bare finite clause serve as the logical subject of the proposition; it can also occupy what appears to be the structural subject position.

(1) a. [$_S$ She likes coffee] is what she likes.

   b. [$_S$ That's what she likes] is (she likes) coffee.

   c. [$_{DP}$ Coffee] is what she likes.

The puzzle undertaken in this chapter has two elements. First, bare finite clauses function as structural subjects only in amalgams; they cannot serve as subjects in canonical clauses.

(2) *[$_S$ She likes coffee] bothers him.

Second, despite the unusual occurrence of bare sentential subjects, the relation between the clausal functional head and the subject in amalgams patterns with canonical clauses in a number of ways.

7.1.1 Bare vs. headed sentential subjects

It is not at all obvious how to reconcile the existence of bare sentential subjects in amalgams with current theory. Since bare finite clauses cannot serve as structural subjects in ordinary sentences, one of the empirical goals of each new analysis of sentential arguments over the past few decades has always been to rule them out (Stowell 1981a; Kayne 1981; Pesetsky 1982; Webelhuth 1992; Doherty 1997, 2000; Bošković 1997b; Pesetsky and Torrego 2001, 2007; Bošković and Lasnik 2003; Landau 2007; Takahashi 2010). The data from amalgams, where a bare sentence can be the subject of a clause headed by a finite copula, shows that these theories lack empirical adequacy.

Previous analyses of constraints on sentential subjects are of course concerned with canonical clauses, where the displacement domain is projected, not copular amalgams, which consist only of the context domain (CP). Ideally, the extremely limited distribution of bare sentential subjects—in copular amalgams only—will follow from the unconventional structure proposed for copular amalgams in Chapter 5. What is it about the structure of the CP that allows finite C-domain heads but not T-domain heads to take bare sentential subjects? The answer developed in this chapter bears on the second property of finiteness: the ability of a finite clause to license an independently referring (nominative) subject, or the *finiteness-case connection*.

The distribution of bare sentential subjects in amalgams parallels the distribution of overtly nominal DP subjects in canonical clauses, with respect to the finiteness-case connection. The constructions in (3) and (4) demonstrate that like DP subjects, sentential subjects cannot remain in the structural subject position of a non-finite clause.

(3)   a.   *It seems [she likes coffee] to be what she likes.

        b.   *I would be disappointed [she likes coffee] to be what she likes.

(4)   a.   *It seems [coffee] to be what she likes.

        b.   *I would be disappointed [coffee] to be what she likes.

Despite this parallel between bare sentential subjects and nominal subjects, their distribution is distinctly non-overlapping when non-specificational sentence types are considered. Bare finite clauses cannot be ordinary subjects in English (or indeed in any language that I am aware of), except in specificational copular sentences. Sentential subjects in ordinary finite clauses must be introduced by a functional element. Compare the examples in (5):

(5)   a.   *She went to the gym surprised/would surprise me.

        b.   **That** she went to the gym surprised me.

        c.   **For** her to go to the gym would surprise me.

        d.   **To** go to the gym is beneficial.

In copular amalgam sentences, however, the distribution of bare and overly C-headed sentential arguments is reversed:

(6)   a.   She went to the gym is where she went.

        b.   ***That** she went to the gym is where she went.

        c.   ***For** her to go to the gym is where she went.

        d.   ***To** go to the gym is where she went.

Complicating matters, in specificational copular sentences, either the logical subject or the logical predicate may occupy the structural subject position. The inverted predicate of an amalgam specificational sentence may be a DP or an indirect question CP, both of which are perfectly compatible with the canonical structural subject position. Whatever mechanism makes bare CPs able to serve as structural subjects of copular amalgam clauses must allow non-bare CPs and DPs as

well.

(7)   a.   The place she went to is she went to the gym.

       b.   The place she went to is the gym.

(8)   a.   Where she went is she went to the gym.

       b.   Where she went is the gym.

The partially overlapping distribution of subject expression types in canonical and amalgam clauses requires explanation.

7.1.2   The structural subject position

At this point, a clarification of the notion of *structural subject position* is due. I have argued that copular amalgam sentences diverge from canonical verbal sentences with respect to a number of formal properties. Most notably, as the previous chapter showed, copular amalgam clauses lack the lower and middle functional fields of the clause. For example, aspectual auxiliaries and modifiers are not grammatical in amalgams, while they are grammatical in closely related canonical specificational sentences:

(9)   a.   *What he liked (then) had been he liked coffee.

       b.   What he liked then had been coffee.

(10)   a.   *What he likes has long been he likes coffee.

        b.   What he likes has long been coffee.

(11)   a.   *What he likes is still he likes coffee.

        b.   What he likes is still coffee.

Similarly, negation, a projection in the middle functional field, can occur in canonical sentences, but not in amalgams.

(12)   a.   *What he likes isn't he likes coffee.

        b.   What he likes isn't coffee.

On the basis of such asymmetries, I concluded that amalgams, unlike canonical sentences, lack the T and v/V fields of the clause, and are instead projected directly from the highest functional field. Adopting a split-CP model of the clause (e.g., Rizzi 1997), I proposed that copular amalgam clauses are projections of Fin or Top, rather than V. One of the consequences of this model is that a clause lacking the middle functional field lacks the canonical (A-)subject position.

I will continue to refer to the precopular expression in amalgams as the structural subject of the sentence, even though I do not analyze it as the occupant of Spec,TP.[1] The label *structural subject* is generally applied to expressions that occur in preverbal (or pre-finite element) position (in English) in a formal dependency with the finite element. This label applies just as well to the subject of an amalgam clause, since it likewise stands in a formal dependency with a finite element—the difference is simply that this finite element is Fin itself, rather than T/Agr. I use the term *structural subject* instead of the pre-theoretical term *precopular expression* for two reasons: (i) in some cases, the precopular constituent of an amalgam occupies a high topic position, rather than the subject position; (ii) the canonical subject-like patterning of subjects in amalgams is precisely what this chapter seeks to explain. The analytical approach of this chapter uses the unique functional structure of copular amalgams to probe the connection between subjecthood and finiteness.

---

[1] Since temporal arguments are local to T, the actual position of the subject may be an adjacent head Subj. I continue to use the more familiar label TP, unless I specifically distinguish between the *nominal* structural subject position and the *temporal* subject (i.e. Topic Time) position.

7.1.3 Sentential subjects?

Before I compare sentential subjects in ordinary vs. amalgam clauses, there is an ontological question to confront: are there sentential subjects at all? The external distribution of sentential subjects has always been a challenging topic in syntactic theory. While sentential subjects pattern in many respects with overtly nominal subjects, their distribution is non-overlapping in several environments. Restrictions on their external distribution have been taken as evidence that sentences associate with argument positions only indirectly (Koster 1978; Stowell 1981a; Alrenga 2005; Takahashi 2010, etc.). Nevertheless, the judgment data are highly variable, and empirical findings on sentential subjects point in different directions (Davies and Dubinsky 2009; Lohndal 2013). I conclude that English has CP arguments where related languages have only DPs. This line of reasoning naturally raises the question of to what extent English-type sentential subjects are nominal.

7.1.4 Roadmap

This chapter accounts for the occurrence of bare sentential subjects in amalgams by pursuing two lines of analysis. The first, developed in section 7.2, compares the external distribution of sentential subjects with that of overtly nominal subjects. I will show that ordinary sentential subjects can in fact occupy the canonical subject position, where they Agree with clausal functional heads. The second line of analysis, developed in section 7.3, compares the internal composition of the CP and the DP. It examines the asymmetries in the distribution of overtly headed (e.g., *that*-ful) CPs vs. ∅-CPs on the basis of the feature composition of the C-domain. In section 7.4, I present the *that*-omission asymmetry and show how existing analyses of CP and IP arguments are incompatible with the existence of bare sentential subjects in copular amalgam clauses. Next, in section 7.5, I briefly revisit the descriptive profile of the amalgam sentential subject itself, arguing that it is a type of embedded root.

Bringing the two major lines of analysis together in sections 7.6 and 7.7, I explain why bare sentential subjects are restricted to T-less clauses. I show how copular amalgam sentences bear out the finiteness-case connection. While nominal subjects are dependently anchored, in an Agree relation with T/Subj, bare clausal subjects in the C-domain are deictically anchored. The interplay between [T] and [ɸ]-features in the T and C domains derives the patterns of subject licensing in both canonical and amalgam clauses. The analysis in this chapter shows that the reduced clause structure of the copular amalgam in fact predict that it licenses bare sentential subjects.

7.2   CPs can be subjects

It is not a settled matter that expressions of categories other than DP can serve as structural subjects (or arguments) at all. This section once again takes up the analytical approach used in Part 2. To understand the peculiar properties of amalgam sentences, we must first understand their canonical counterparts. In order to establish a baseline for analyzing bare sentential subjects, this section analyzes canonical CP arguments, and subjects in particular.

CPs can serve as arguments much less freely than their nominal counterparts. For example, CPs are generally unacceptable as indirect objects (13).

(13)   a. ??I didn't give [that you left early] another thought.

    b.   I didn't give [that idea] another thought.

On the other hand, CPs with argument roles can occur in (structurally) caseless positions more easily than their DP counterparts. For instance, CPs, unlike DPs, are easily extraposed (14) and can be complements to verbs that do not license accusative DPs (15).

(14)   a.   It disappointed me that you left.

    b.   *It disappointed me*(,) your departure.

(15)   a.   She wondered [what time it was].

           b.   *She wondered [the time].   (Pesetsky 1982: Sect. 2.5)

Because of such asymmetries between DP and non-DP arguments, it is not clear that non-DPs can serve as arguments directly. In the specific case of putative sentential subjects, it has been argued that CPs cannot occupy Spec,TP—they serve as subjects indirectly, by association with some DP (Koster 1978; Stowell 1981b; Webelhuth 1992; Alrenga 2005; Davies and Dubinsky 2009, and others). Another camp argues that the distribution of sentential subjects is constrained by the internal featural composition of CP, which sets them up for different Agree relationships with the clausal functional heads T and C (see, among others, Pesetsky and Torrego 2001, 2007).

I will follow the latter line of analysis. Internal structural differences between different CP types constrain their interactions with clausal functional structure; crucially, some CPs can serve as subjects of T, while others can only serve as subjects of C/Fin. To lay the groundwork for this analysis, the present section compares the distribution of DP and non-DP (primarily CP) arguments (section 7.2.1), as well as the distribution of CPs in A′-positions (section 7.2.2), and concludes that CPs can in fact be canonical subjects. Following Davies and Dubinsky (2009), I argue that the limited distribution of CPs with respect to DPs can be attributed to processing constraints, rather than grammatical ones (section 7.2.3).

7.2.1   External distribution of DP and CP arguments

The robust distributional overlap between CPs and DPs suggests two competing hypotheses. The first is that (non-root) CPs actually are DPs, and the second, which I favor, is that similarities between the two categories arise from the functional and formal similarity of the extended verbal

and nominal domains (e.g, Abney 1987). The present chapter is concerned with the role of CP as structural subject, but since the occupant of the structural subject position is generally base-merged somewhere else, this subsection considers CPs and DPs in a variety of environments.

7.2.1.1 Honorary NPs

While prototypical arguments are straightforwardly nominal, English and many other languages permit expressions that are superficially non-nominal to occur in argument positions. Safir (1983) characterized such expressions as "Honorary NPs". Honorary NPs are headed overtly by non-nominal categories, like A, V, and P. In the following examples, the Honorary NPs occupy the subject position of the clause (copular clauses in particular) (16), but they can also serve as objects of V and P.

(16) a. Honest is the most important thing to be.

  b. Under the bed is a good place to hide.

  c. John fired is what I'd like to see.

  d. Leave early is what we should do.

The question that Honorary NPs raise is the source of their nominality. Are they embedded under a null D-layer or are they linked to a null noun?

(17) a. [DP [D ∅][PP under the bed]]

  b. [[DP/NP ∅][PP under the bed]]

7.2.1.1.1 Type-shifting Honorary NPs

At the semantic level, a predicate expression like *honest* can be type-shifted (via *nom*) into an entity-denoting expression (Partee 1987). In Chierchia and Turner's (1988) framework, the type-shifter "down" maps a property-denoting expression onto its lower-type entity correlate.

(18)   nom: honest$_{e,t}$ → $^{\cap}blue_e$ (Chierchia and Turner 1988)

The output of this type-shift is a referring expression: something semantically indistinguishable from a DP. The fact that the Honorary NP must be interpreted as definite supports the null-headed DP analysis.

7.2.1.1.2 Case and agreement

The external syntactic distribution of Honorary NPs bears some of the hallmarks of nominal syntax. They pattern with DPs with respect to the finiteness-case connection, occurring only in A-positions that also license DPs (Safir 1983; den Dikken 2005b):

(19)   a.   Under the bed seems to be a good place to hide.

         b.   *It seems under the bed to be a good place to hide.

(20)   a.   The closet seems to be a good place to hide.

         b.   *It seems the closet a good place to hide.

In addition, coordinated Honorary NPs in subject position can control plural agreement on the verb and associate with plural quantificational expressions like *both*.

(21)   a.   [[In the woods] and [to the beach]] **are equally** where I'd like to go.

         b.   [[John fired] and [Mary promoted]] **are both** what I'd like to see.

The presence of [ɸ]-features on these expressions suggests a distinctly nominal character, which cannot be attributed to their lexical heads.

7.2.1.2 *Wh*-**CP arguments**

Another class of potential DPs-in-disguise, which is of more direct relevance to the status of (declarative) sentential arguments, consists of CPs introduced by *wh*-elements: free relatives and indirect questions. Unlike Honorary NPs, the source of a *wh*-CP's nominal behavior does not necessarily depend on an element of category D. Free relatives and indirect questions may owe their nominal properties to the occupant of Spec,CP, or C itself.

7.2.1.2.1 Nominal properties of free relatives

Free relatives have a DP-like distribution (Caponigro 2003), occurring as complements of prepositions and accusative case-assigning verbs.[2]

(22)   a.   I went to [where you went].

          b.   I like [what you like].

These free relatives, like DPs, are sensitive to the finiteness-case connection.

(23)   a.   What you like seems to be nutritious.

          b.   *It seems what you like to be nutritious.

On a headless analysis of free relatives, their nominality is due to the properties of the *wh*-expressions themselves. The *wh*-expressions *what, who, when, where, why,* and *how* can all be

---

[2]Caponigro (2003) observes that free relatives introduced by *when, where,* and *how* can also have a PP-like distribution:

(i)   a.   I went [where you told me to].
        b.   I went [*(to) the place that you told me to].

(ii)   a.   I'll meet you [when you are at home].
        b.   I'll meet you [*(at) the time that you are home].

(iii)   a.   I play [how he taught me].
        b.   I play [(in) the way that he taught me].

paraphrased by definite nominals: *the thing, the person, the time, the place, the reason,* and *the way*. Although they occupy a specifier position, they can in a sense "project" their nominality up to the CP. For Caponigro (2003), this is accomplished at the semantic level by means of a type-shifting operation: a functor adjoined to CP takes a predicate and yields a maximal entity.

Taking a syntactic approach, Citko (2008) proposes that the D-label of the *wh*-expression is actually projected when it merges with CP. The free relative is a configuration that allows the goal, rather than the probe, to project (see also Donati and Cecchetto 2011).

A more widely adopted approach to free relatives captures their DP-like distribution by treating them as bona fide DPs, headed either by a null element (e.g., den Dikken 2006) or by the *wh*-word itself, in a position external to the CP (e.g., Larson 1987; Kayne 1994). If either of these approaches is correct, then we are still in a position to be uncertain as to the existence of non-DP arguments. Honorary NPs and free relatives are both amenable to a D-complementation analysis: they are DPs in disguise.

I turn now to the status of interrogative argument CPs, which provide more promising support for the existence of non-DP arguments.

7.2.1.2.2 Interrogative CP arguments

While free relatives are formally DPs, with nominal semantics, indirect questions and exclamative clauses are CPs, with propositional semantics. One piece of evidence in favor of a CP treatment of these *wh*-clauses comes from the behavior of *why*. Unlike the other *wh*-words, *why* does not introduce DP-like free relatives:

(24)   a.   *I did it [why you told me to]

         b.   *She stayed home [why we we did].

Why-interrogatives, on the other hand, can occupy argument environments.

(25)   a.   I don't know why he did it.

            b.   Why she left surprised him.

            c.   You should give why you left some serious thought.

Since *why* does not make available a DP-like interpretation of free relatives, it likewise does not turn interrogative CPs into DPs; however, *why*-CPs can function as arguments.

    Semantic selection constrains the distribution of interrogative and exclamative CPs. Predicates like *ask, wonder, figure out, tell*, and *know* select for question-denoting expressions.

(26)   a.   I asked what his name was.

            b.   She wonders how he manages.

            c.   I can't figure out where he went.

            d.   Tell me what I should do.

            e.   They know why we did it.

The examples in (26) are all CPs, but certain DPs can also have question-denotations, as discussed in Chapter 4. These DPs, known as "concealed questions", pattern semantically with indirect question CPs. Concealed questions thus reverse the situation found in free relatives, where what appears to be a CP patterns with DP semantically. Observe, for example, the interrogative complements in (27) and (28).

(27)   a.   I asked what his name was.

            b.   I asked his name.

(28)   a.   She told me where to go.

    b.   She told me the address.

Similarly, some predicates selecting exclamative CPs also select DPs (Grimshaw 1979):

(29)   a.   I can't believe how tall he is.

        b.   I can't believe his height.

Not all syntactic environments that semantically select for interrogative or exclamative propositions license DPs, however. Many predicates can take CP complements, but cannot take DPs.

(30)   a.   She wondered what time it was.

        b. *She wondered the time.

(31)   a.   He inquired what her name was.

        b. *He inquired her name.

Grimshaw (1991) attributes the asymmetry to the case-assigning properties of predicates like *wonder*. If *wonder* fails to assign accusative case, then it cannot license a DP complement. The structure can be saved, however, if instead of selecting the DP directly, the verb selects a PP: the preposition *about* takes an accusative DP as a complement (or assigns lexical case, in a more current implementation), so (32) is grammatical.

(32)   She wondered about the time.

Because of this asymmetry between semantic and syntactic selection, there is a case to be made for the existence of true CP arguments. *Wh*-CPs may occur where DPs cannot be licensed, so they are not dependent on any DP associate or D-layer.

Interrogative and exclamative CPs do manifest some formally nominal properties, however. The nominal properties of any argument CP stem from the structure of the CP itself. For example,

coordinated interrogative or exclamative CP subjects can value plural agreement on the verb, so they must have [φ]-features.

(33)   a.   [[Where we go] and [how we get there]] **depend** on traffic.

   b.   [[How tall he has gotten] and [what a nice guy he is]] **are** really surprising.

In addition, in predicational contexts, they can associate with plural quantifiers like *both* and *equally*.

(34)   a.   [[What he did] and [how he did it]] are **both** disturbing.

   b.   [[How friendly he is] and [what a generous guy he turned out to be]] are **equally** inspiring.

Although these CPs have nominal [φ]-features, they retain their propositional structure and semantics.

7.2.1.3   Declarative sentential arguments

Declarative CPs, like their interrogative and exclamative counterparts, distribute externally like DPs in many respects. They occur in complement positions and subject position and enter into Agree relations for [φ] and [T]-features. Unlike DP complements, they can occur where accusative case is unavailable; nevertheless, they are constrained by the finiteness-case connection in subject position. I attribute these patterns to the internal structure of CP and not to the presence of a null-D element.

As with indirect questions and exclamative clauses, some of the predicates selecting declarative CPs as internal arguments also select DPs. These predicates include, for instance, verbs of saying and epistemic verbs.

(35) a. She said that we should go.

  b. She said her idea.

(36) a. They know that he is hungry.

  b. They know the truth.

The DPs that alternate with CPs in these environments have propositional denotations. The selecting predicates must therefore semantically require arguments with propositional content. If selection is strictly local, then the most parsimonious analysis of these CPs is that they are selected directly, without the mediation of a DP.

In some argument environments, declarative CPs are more restricted than their DP counterparts. As mentioned above, CPs do not readily serve as indirect objects:

(37) a. ??She gave [that he wanted to leave] careful consideration.

  b. She gave [the idea] careful consideration.

Finite CPs are sometimes, but not always, degraded as complements to prepositions.

(38) a. ??They were horrified at [that she lied].

  b. They were horrified at [the idea].

(39) a. She talked about [that you are quitting].

  b. She talked about [your plan].

CPs likewise have a more restricted distribution than DPs as subjects of clauses embedded under complementizers; compare (40) and (41).

(40) a. ?Mary is unhappy because for her to travel to Tahiti is no longer necessary.

  b. ?Although that the house is empty depresses you, it pleases me.

    c.   ?Jim raised the possibility that for the house to be destroyed would upset you.

    d.   ?That for us to smoke would bother her, I didn't expect.

    e.   ?That for us to smoke bothers her is quite obvious.   (Alrenga 2005:178 (7))

(41)   a.   Mary is unhappy because her trip to Tahiti is no longer necessary.

        b.   Although the house's emptiness depresses you, it pleases me.

        c.   Jim raised the possibility that the house's destruction would upset you.

        d.   That our smoking would bother her, I didn't expect.

        e.   That our smoking bothers her is quite obvious.   (Alrenga 2005:178 (8))

Although CPs are not as acceptable in these environments as DPs, it is important to point out that they are not categorically unacceptable (Delahunty 1983; Authier 1992; Lohndal 2013; Davies and Dubinsky 2009). For example, there are many acceptable instances of CPs in embedded subject positions.

(42)   a.   She hopes that [that he's gone] is true.

        b.   They regretted that [for him to leave] was necessary.

        c.   I expected [that he wanted to leave] to be obvious.

Data like (40) thus constitute a questionable empirical basis for arguments against the existence of true CP arguments. The lower acceptability of these CPs must be accounted for, but the gradience of the judgments raises the possibility of a usage or processing-based explanation (see Davies and Dubinsky 2009 and discussion in section 7.2.3 below).

    Complement CPs do not depend on accusative case for licensing, unlike their DP counterparts. With raising predicates, for instance, CPs can remain in-situ, while DPs must raise into a higher A-position. (DPs in these environments can extrapose, as in (43c), receiving a backgrounded inter-

pretation and deaccented prosody.)

(43) a. It's amazing [that she won].

   b. *It's amazing [her win].

   cf. It's amazing, HER WIN.

The same pattern is observed in passives.

(44) a. It's widely believed [that she won].

   b. *It's widely believed [her win].

   cf. It's widely believed, HER WIN.

CP complements thus are not required to occupy a case-licensed A-position.

On the other hand, CP subjects are constrained by the finiteness-case connection, like DPs. They cannot serve as independently referring (nominative-like) subjects of non-finite clauses.

(45) a. *It appears that he's bored to be obvious.

   b. *It appears his boredom to be obvious.

(46) a. *[[That he likes her] to be obvious], he should just speak up.

   b. *[[His preference] to be obvious], he should just speak up.

The connection to nominative case is underscored by the fact that CP subjects enter Agree relations for [ɸ]-features with clausal functional heads. Coordinated CP subjects (in certain semantic environments), like coordinated DPs, can trigger plural [ɸ]-feature agreement on the verb. (The examples below are from McCloskey 1991:564.)[3]

(47) a. That the president will be reelected and that he will be impeached **are equally** likely

---

[3]See discussion in section 7.5 regarding plural triggered by coordinated CPs in specificational copular sentences.

at this point.

b. That the march should go ahead and that it should be canceled **have** been argued by the same people at different times.

c. That he'll resign and that he'll stay in office **seem** at this point **equally** possible.

A puzzle in the analysis of sentential subjects is therefore why they seem constrained by the availability of Agree with a finite element in subject position, but do not depend on case more generally, for instance, in complement position, in contrast to their DP counterparts. Following Pesetsky and Torrego (2001, 2007) and Landau (2004), section 7.3 shows that the impetus for the finiteness-case connection is not only the licensing needs of the subject expression, but also the needs of the clausal functional heads.

7.2.2 The position of sentential subjects

I will conclude in this section that English sentential subjects can occupy Spec,TP directly, without the mediation of some DP element and without extraposing to an A′-position.

Opposing analyses have been widely defended since the early days of Transformational Grammar (Rosenbaum 1967; Emonds 1970; Koster 1978; Stowell 1981a; Delahunty 1982, etc.). This vein of research argues that some DP occupies the argument position, and its CP associate occupies the right or left periphery. The general motivations for this proposal are the fact that DP and CP arguments do not have entirely overlapping distribution, the fact that CP subjects seem to cause A′-like intervention effects, and the fact that CP arguments are typically pronounced at the periphery of the clause. Taken together, these facts naturally suggest the hypothesis that CP arguments must be in the periphery.[4]

---

[4]Each of the analyses mentioned above proposes that sentential arguments occupy the periphery at some level of representation, assuming an early transformational T-model of grammar, but they differ as to whether that level is D-structure or S-structure.

7.2.2.1 Extraposition and Case resistance

CP arguments are found most typically in peripheral environments:

(48)   a.   That he lied surprised us.

        b.   It surprised us that he lied.

        c.   I'm surprised that he lied.

        d.   That he lied, it surprises me.

        e.   I gave it careful consideration that he lied.

In these cases, the pronominal element *it* stands in for the CP. The question is whether CPs *must* occur in extraposed or peripheral positions.

Stowell (1981a) argues that the answer is affirmative. A Case Resistance Principle is responsible for barring CPs, TPs, VPs, and PPs from case positions. Since these categories assign case (assumed at the time to take place under Spec-Head Agreement or Government), they cannot occur in case positions. CPs cannot be subjects, according to the Case Resistance Principle, because they contain the case-assigning feature [+Tense] themselves. This principle captures the tendency for argument expressions in A-positions to be DPs and APs. CP "subjects", according to Stowell, pass through Spec,TP and raise to Topic position in the left periphery, leaving a nominal trace in subject position.[5]

(49) $[_{TopP}$ $[_{CP}$That he liked beer$]_i$ $[_{TP}$ t_i surprised us$]]$.

Davies and Dubinsky (2009) point out that the Case Resistance Principle is called into question by data from Austronesian (Chung 1991; Levin and Massam 1986). Chung (1991) shows that in

---

[5]Other Honorary NPs predicted not to occur as arguments by the Case Resistance Principle are argued to be embedded under null D-layers, an account that is consistent with what was proposed above, although I do not subscribe to the Case Resistance Principle itself.

Chamorro, *wh*-movement across a verb triggers morphology on the verb, reflecting the grammati-
cal role of the extracted element.[6] (50) exemplifies subject extraction, and (51) exemplifies object
extraction.

(50)   a.   Ha-istótotba   yu' i   buruka.
               INFL-disturb.PROG me the noise
               'The noise is disturbing me.'

        b.   Hafa$_i$ **um**istótotba   hao t$_i$?
               what INFL.disturb.PROG you
               'What is disturbing you?'   (Chung 1991: 82 (15))

(51)   a.   Ha-chiku si Dolores i   neni.
               INFL-kiss   Dolores the baby
               'Dolores kissed the baby.'

        b.   Hay$_i$ ch**in**iku-**ña** si Dolores t$_i$?
               who kiss-AGR   Dolores
               'Who did Dolores kiss?'   (Chung 1991: 83 (17))

Long-distance extraction out of an argument CP results in agreement with the extracted element
on the lower verb, but the higher verb inflects for agreement with the CP argument containing the
extraction site. If the CP is a sentential subject, the higher verb shows subject agreement, and
if it is an object, the higher verb shows object agreement. Both (52) and (53) illustrate object
extraction, with the expected inflection on the lower verb, but (52) shows object extraction out of
a sentential subject, while (53) shows object extraction out of a sentential object. Thus, the higher
verb's inflection in (52) shows subject agreement, and the higher verb's inflection in (53) shows
object agreement.

(52)   Hayi$_i$ **hum**ongngang hämyu [ni **ha**-bisísita   Francisco t$_i$]?
       who INFL.surprise you   [that INFL-visit.PROG Francisco]

---

[6]Whether or not Austronesian voice marking should be analyzed as a type of case is well beyond the scope of this
work. What is important is that the marking indicates the grammatical role of the clause; it does not take oblique form,
for instance, which might be expected if the CP could never be a subject.

'Who does it surprise you that Francisco visits?'

(53)   Hayi$_i$ ma'a'ñao-**ña** si Manuel [pära **u**-lalatdi   t$_i$]?
       who   afraid-AGR   Manuel [will INFL-scold]
       'Who is Manuel afraid he might scold?'   (Chung 1991: 83–84 (19))

Since this agreement pattern in all respects correlates with the case properties of the arguments triggering agreement, these authors conclude that the CP arguments bear case.

The fact that CPs can participate in the Agree relations associated with case is central to the analysis that will be presented for copular amalgam sentences. While CP subjects in canonical sentences can occur in ordinary case positions, bare CPs cannot. The absence of an overt functional element in C prevents CP from distributing like an ordinary subject. If CPs could not serve directly as arguments at all, this asymmetry would be unexpected: there must be something more nominal about the overtly headed CP than the bare CP. This is an old intuition in the literature, and the basis for the hypothesis I pursue in section 7.6.

7.2.2.2 Sentential subjects as topics

Early generative approaches to sentential subjects, like Emonds (1970, 1976) and Koster (1978), take the tendency of CP-extraposition and the asymmetries between CP and ordinary DP subjects seriously. Phrase Structure Rules ensure that the subject position must be occupied by an NP, so CP subjects cannot be generated. Putative CP subjects must therefore occupy some other position, and leave an NP to serve as the real structural subject. Many scholars assume this position is a peripheral topic position.

Emonds (1970, 1976), Koster (1978), and later Alrenga (2005) observe that sentential subjects and topics have similar syntactic distributions, and argue that sentential subjects are topics. The structural subject position is instead occupied by some DP. For example, inversion around a

sentential subject or a topic in questions fares poorly (54), while inversion around a pronominal associate of a dislocated clause is acceptable (55).

(54) a. *Did [that John showed up] please you?   (Koster 1978: (3a))

      b. *Did [yesterday] he leave?

      c. *What does [that he will come] prove? (Koster 1978:53)

      d. *What does [tomorrow] he plan to do?

(55) a. Did [it] please you that John showed up?

      b. Did [John's arrival] please you?

Similarly, negative inversion around sentential subjects and topics is unacceptable, in contrast to negative inversion around subjects.

(56) a. ??Never could [that she stayed out all night] shock me.

      b. *Never could [on Saturdays] she stay out all night.

(57) a. Never could it shock me that she stayed out all night.

      b. Never could she stay out all night on Saturdays.

Alrenga (2005) also observes that sentential subjects and topics do not easily co-occur, attributing the decreased acceptability of the examples in (58) to the same source: fronted topics are not easily stacked.[7]

(58) a. *John, the book, I gave to.   (Alrenga 2005:179 (14))

      b. *John, that the Giants lost the World Series shouldn't have bothered. (Alrenga 2005:177

          (5))

---

[7]Of course, there is no upper limit the number of topic phrases that can occur in the left periphery. This constraint on stacking topics seems to be specific to topicalizing out of an A-position.

A similar example is given by Koster:

(59)   *Such things, that he reads so much doesn't prove. (Koster 1978: (5b))

Davies and Dubinsky (2009) argue, however, that these examples do not strongly support a topicalization analysis of sentential subjects, since the complex NP and extraposed "versions" of the same sentential subject are degraded when a phrase is topicalized above them.

(60)   a. *Such things, the fact that he reads so much doesn't prove.   (Davies and Dubinsky

            2009:122 (33))

        b. *Such things, it doesn't prove that he reads so much.

Since topicalization across the DP subjects in (60), which need not be topics, is just as unacceptable as it is across sentential subjects, it is not necessarily the case that sentential subjects are topics either.

If sentential subjects were topics, then, as Koster (1978) points out, they should be degraded in embedded contexts that do not allow topics. Embedded clause types that are truncated at the level of FinP or below do not allow topics (e.g., Rizzi 1997, and many others).

(61)   a. *I regret [[$_{TopP}$ to John], I said something mean].

        b. *I'd hate [for [$_{TopP}$ to John] her to say something mean].

        c. *I expected [[$_{TopP}$ to John] her to say something mean].

        d. *I want [[$_{TopP}$ to John] to say something mean].

Koster's prediction is partially borne out: sentential subjects do not tend to occur in these environments. Alrenga 2005:178 (9) notes the following examples (judgments are his).

(62)   a. ?*I regret [that [for us to smoke] bothers her so much].

    b. ?*Mary wishes [that [for us to smoke] bothered her more than it did].

The restrictions on embedding are far from categorical, however. First of all, the examples in (62) include a confound: *for-to* infinitives are irrealis (Stowell 1982), so the realis forms of the main verbs cause a clash. (This factor is mitigated in the examples in (63).) Furthermore, sentential subjects are not entirely ruled out, even in embedded environments that do not license topics. While the following are not the most straightforward way to express the intended meaning, they are certainly acceptable for many speakers.

(63)   a.   I regret [that [for her to leave] would cause so much trouble].

           b.   ?I'd hate [for [that it's raining] to stop us from going].

           c.   I expected [[that he was lying] to be obvious].

Since sentential subjects can occur in clauses that do not admit topics, it cannot be maintained that they must occupy a peripheral topic position.

    The ban on inversion around sentential subjects is also not categorical. Delahunty (1982:387) offers the following acceptable examples of inversion around a sentential subject:

(64)   a.   Does that Fred lied to them bother all of the people who bought stock in his company?

           b.   Does that the world is round bother as many people now as it did 500 years ago?

           c.   Does that quarks have wings explain their odd behaviour?

           d.   Does that quarks have wings explain anything at all?

Likewise, Davies and Dubinsky (2009:115) give these examples of inversion around sentential subjects.

(65)   a.   To whom is [that pigs can fly] most surprising?

    b.   Is [that I am done with this homework] really amazing?

I find the negative inversion examples in (66) somewhat acceptable as well.

(66)   a.   Never has that he lied bothered me so much.

           b.   Rarely has that he's extremely short affected my opinion.

Since T-to-C movement cannot target a pre-Topic position, it should not be possible at all in sentences with sentential subjects, if sentential subjects are in fact topics. Given that inversion around sentential subjects is possible, contra Koster (1978), sentential subjects need not be topics.

    Information structure presents another problem with the obligatory topicalization analysis of sentential subjects. If sentential subjects are required to occupy a topic position in the periphery of the clause like other left-dislocated elements, then they should in fact be interpreted as topics, and as a consequence, the predicate should be focused. This is often the case, as (67) illustrates. (Capital letters in indicate the focal pitch accent marking focus which projects up to the CP; anaphoric de-accenting is represented with small caps.)

(67)   a.   Did you know that they lost?

           b.   Yeah, [THAT THEY LOST THE GAME] [really SUCKS]$_F$.

    It is not *necessarily* the case, however. CP subjects can receive a narrow focus interpretation in specificational sentences, including both canonical and amalgam pseudoclefts. Specificational sentences, as the reader will recall, have a fixed information structure, where the value is narrowly focused and the expression containing the variable is a topic. They are also reversible: the value can occur in subject position. Consider specificational sentences where the value is a CP. In the typical specificational order, the focused CP is postcopular.

(68)   a.   WHAT HE THINKS is [that you should LEAVE]$_F$.

        b.   HER CONCERN is [that he LIED]$_F$.

        c.   WHAT HAPPENED is [that they lost their KEYS]$_F$.

If these specificational sentences are reversed, then as expected, the focused value CP is perfectly

acceptable in the precopular position.

(69)   a.   [That you should LEAVE] is WHAT HE THINKS.

        b.   [That he LIED] is HER CONCERN.

        c.   [That they lost their KEYS] is WHAT HAPPENED.

Similarly, the bare CP subjects of specificational copular amalgam sentences are narrowly focused.

(70)   (71)   a.   [You should LEAVE] is WHAT HE THINKS.

                b.   [He LIED] is HER CONCERN.

                c.   [They lost their KEYS] is WHAT HAPPENED.

CP subjects can bear narrow focus, so it cannot be the case that they are obligatorily topics.

7.2.2.3   Sentential subjects in Spec,TP

Although sentential subjects do not overlap completely in distribution with nominal subjects, the

preceding subsections have provided evidence from agreement, case, and syntactic position with

respect to fronted elements that they occupy the canonical structural subject position. Since they

occupy this position, it is thus fruitful to analyze them, like their DP counterparts, in an Agree-

based system. The internal structure of the CP argument constrains its featural makeup: where it

distributes more freely than a DP, it has a valued feature where DP has an unvalued one; where it

distributes less freely than a DP, the reverse is true.

In section 7.3, I will discuss the formal differences between C and D in more detail, drawing heavily on the insights of Pesetsky and Torrego (2001, 2007) and Landau (2004). The features of CP and the features of the clausal functional heads determine the distribution of sentential subjects, and present an analytical approach to the difference between sentential subjects of canonical and amalgam clauses.

Before I make this analytical approach explicit, in the context of the Agree-based approach to CP subjects, I will briefly address the cases of asymmetry between CP and DP arguments that do not fall under the feature-based explanation above.

7.2.3 Restrictions explained

The limited distribution of argument CPs requires explanation if an Agree-based analysis of argument CPs in subject position is to be maintained. An explanation can be found in constraints on processing (Grosu and Thomson 1977; Dryer 1980; Delahunty 1983; Erdmann 1988; Widmann 2005; Davies and Dubinsky 2009).

First of all, as section 7.2.2.2 showed, the syntactic constraints on sentential arguments that are commonly cited as evidence that CPs cannot serve as arguments directly are gradient. While Koster (1978) and Alrenga (2005) state that CPs cannot occur in embedded contexts that do not license topics, others (e.g., Delahunty 1983) report many acceptable examples, like (63) above. Similarly, while subject-auxiliary inversion and negative inversion with sentential subjects can result in lowered acceptability, it is not wholly acceptable, as in (64) and (66) above. Such gradience favors a non-syntactic analysis of constraints on CP subjects, because if CPs really could not occupy Spec,TP, then these judgments should be categorical.

In addition, the tendency for CP arguments to extrapose—from both complement and subject positions—is surprisingly robust from a cross-linguistic perspective (Dryer 1980). While many

languages allow clause-medial sentential arguments, they are dispreferred unless they are overtly nominalized, and in some languages, they are unacceptable. For instance, in Persian and Turkish, which are both SOV, Dryer reports that sentential objects must be dislocated. Given that CPs in other respects (e.g., case and agreement) behave similarly to ordinary arguments, it is implausible that Universal Grammar should be responsible for requiring arguments to dislocate, just in case they are sentential. Why should only CPs show this behavior, and not other categories? Such a universal tendency is perfectly compatible with a processing explanation, especially since the tendency is stronger in verb-final languages, where sentential objects would otherwise occur in a medial position.

Davies and Dubinsky (2009) report on several studies that have pursued a line of explanation in terms of sentence processing; their report is summarized in what follows. Dryer (1980:161) comments in his typological survey of sentential arguments: "If we assume that clauses are the fundamental units of sentence processing, and that material is emptied from short term memory at clause boundaries, clause-internal sentential NPs will interrupt the processing of the main clause." This approach is intuitively appealing, and supported by an eye-tracking study that indicated slower reading times for in-situ sentential subjects than for their extraposed counterparts (Frazier and Rayner 1988), which parallels the slower reading times in response to similar processing disruptions, like center-embedded relative clauses.

Phonological weight is another culprit in the reduced acceptability of certain sentential arguments (Delahunty 1982). Just as in heavy NP shift, where an NP that is relatively heavy with respect to the material that follows its base position tends to dislocate, Erdmann (1988) shows that heavy CP subjects of lighter predicates are extremely likely to dislocate. If the CP is too long, it is taxing to the working memory that has not yet completed processing of the larger constituent

the CP is embedded in. If the CP is too short, then the parser cannot recover from the garden-path triggered by the unexpected embedded clause before encountering the end of the string. Medial CPs of roughly seven words, where the remaining predicate is of similar length, are shown to have the optimal weight (Widmann 2005).

(72)   a.   Does [that the parent wanted to come home] [cause any problem for the older children]?

      b.   Although [that the parent wanted to come home] [caused problems for the older children], it was not a terrible inconvenience.

In the unacceptable examples often cited as evidence that the CP must be a topic, these ameliorating factors are absent.

It should be noted at this point that I do not claim that all languages have CP structural subjects. Given the constraints on processing CP subjects, it is plausible that some languages have grammaticalized a mechanism that ensures that sentential arguments dislocate—this is a different claim than that this mechanism is part of Universal Grammar. English can be contrasted with its V2 Germanic neighbors to show that its sentential subjects pattern more clearly with DP subjects than theirs.

Lohndal (2013) shows, for instance, that in Norwegian Verb-Second clauses, CP subjects can occupy the preverbal position, traditionally treated as a Spec,TopP or Spec,CP position.

(73)   At han kom   så sent, ødela festen.
      that he   arrived so late, ruined the.party
      'That he arrived so late, ruined the party.'   (Lohndal 2013:8 (61))

When a non-subject occupies the preverbal position, by contrast, DP subjects (74), but not CP subjects (75), remain in Spec,TP.

(74)   Bøkene   leser John hver dag.
        the.books reads John every day
        'The books, John reads every day.'   (Lohndal 2013:9 (64))

(75)   *I fjor   overrasket at   John vant prisen,   alle deltakerne.
         in last.year surprised that John won the.prize, all   the.participants
        'Last year it surprised all participants that John won the prize.'   (Lohndal 2013:9 (65d))

The CP, however, can remain in Spec,TP if it is associated with a pronominal element *det* 'it'.[8] An

CP can therefore remain in subject position—even a medial one—if it is overtly part of a complex

NP.[9]

    English certainly has the option of using *it* in such contexts, but it is not obligatory. Compare,

for example, the Norwegian in (75) with the negative inversion example in English (repeated from

(66), above), which resembles Verb-Second, and permits a post-verbal sentential subject.

(76)   Never has (the fact) that he lied bothered me so much.

Examples like (76) show that English is simply more liberal in its treatment of sentential arguments

than its relatives.

7.2.4   Interim conclusion

As Delahunty (1982), Davies and Dubinsky (2009), and others point out, acceptability judgments

concerning sentential subjects tend to be gradient. Processing constraints play a role in limiting

the distribution of sentential arguments. In view of these confounding factors, which cast doubt on

evidence against the existence of true sentential arguments, I maintain that there are CP subjects in

English.

    This conclusion has two analytical consequences for the sentential subjects of amalgams. First,

---

[8]Koster (1978) reports similar facts for Dutch.

[9]For this argument to go through, it must be shown that the pronominal and its CP associate normally form a
constituent in the same syntactic position.

since English differs from other Germanic languages in allowing the CP to associate directly with argument positions, without the mediation of a DP associate, the limited cross-linguistic—and cross-Germanic—distribution of counterweight-initial amalgams can be tied to the limited availability of CP subjects. The subject of the FinP small clause of an amalgam specificational copular sentence is always a bare CP, a type of sentential subject whose distribution is even more restricted than that of the overtly headed CP. It is likely that only languages that allow CP subjects to begin with would allow bare CP subjects. Secondly, the featural makeup of the CP itself, and not some DP associate, determines what kinds of clausal heads a sentential subject can Agree with in English. To approach the question of why the canonical subject position (Spec,TP) cannot house a bare CP subject, while the non-canonical amalgam subject position (Spec,FinP) can, it is fruitful to examine the featural makeup of Fin and T, as it parallels that of D. This task is undertaken in the following section, which takes a closer look at the internal structure of the C-domain.

7.3   Internal parallels between CP and DP

As illustrated in the last section, clauses and noun phrases are similar in their external distributions. They also exhibit formal parallels with respect to their internal structures (e.g., Chomsky 1970; Abney 1987). Abney (1987) fleshed out the similarities in detail, observing that both contain a lexical element at their core, with complement and specifier positions marked by Case and a functional element heading the structure:

(77)
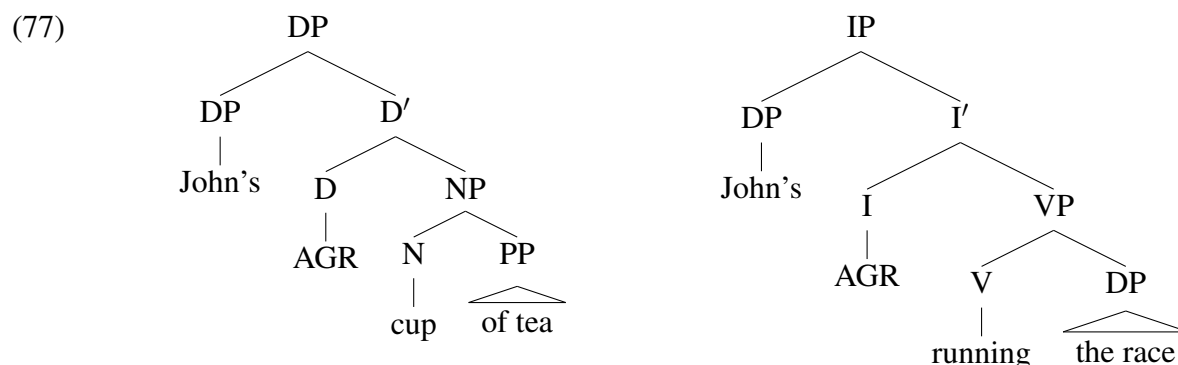
Formal parallels like those above support the conclusion drawn in the previous section, namely, that CPs can serve as arguments, just as DPs can. The similarity in their external distribution is no accident, but rather, due to the architecture of grammar.

It is telling that in the trees above, taken from Abney (1987), D is analogized explicitly to I, rather than C. Similarly, the structure I proposed for amalgam pseudoclefts in Chapter 5 locates the copula in Fin, a position at the boundary between the C and Infl/T domains. The literature on the parallels between verbal and nominal projections is ambivalent about whether the extended noun phrase consists of a single inflectional layer above N (DP) or two (KP/PP). It nevertheless treats the extended verb phrase as unambiguously consisting of the distinct layers IP/TP and CP. This ambivalence leads to confusion in the details of the similarities between the clause and the noun phrase. Is D the counterpart of T (or Infl), or the counterpart of C? Additional confusion stems from the fact that some approaches to the extended noun phrase do not treat determiner elements like *the* as the heads of D, just as some approaches to the extended verb phrase do not treat complementizer elements like *that* as the heads of C.

It is not a coincidence that the boundary between the middle and peripheral domains is fuzzy in both nominal and verbal extended projections. Formal properties of both the T and C domains occur at the structural position of Fin. Copular amalgam sentences provide novel evidence that in a clause where the predicate is non-verbal and the subject is non-nominal, the functions of the T domain are obviated. The abstract function of finiteness (independent anchoring, independent force) and its formal reflexes (agreement and tense inflection, subject licensing) originate in the C-domain, and remain there, if Fin fails to take T as a complement. Because it does not select T and facilitate subject movement to Spec,TP, its specifier position shows both A and A′ properties. Since I treat Fin as straddling the functional boundary between the two domains, the vagueness of

the DP-KP distinction, and the fact that D parallels both T and C in fact reinforces the similarity between the verbal and nominal projections, and the conclusion of the present chapter: that verbal and nominal elements serve as subjects in similar ways, even at the A/A′ boundary.

Similarities between the extended nominal and verbal domains are lexical, featural, and functional. The latter two types of similarity are exploited to account for the alternation of root-like CPs, subordinated CPs, and DPs in the amalgam structural subject position, although I defer discussion of the abstract functions of the parallel domains until section 7.7. The first, which plays only a supporting role, is discussed below.

7.3.1   Elements that occupy C and D

An obvious parallel between the verbal and nominal domains can be found in the form of the function words that head them.

7.3.1.1   Determiners in C

First of all, definite noun phrases and finite clauses in many Indo-European languages can be introduced by formally identical functional elements. Complementizers in Indo-European, for instance, are often morpho-phonologically related to determiners, like demonstratives or *wh*-words.

(78)   a.   Germanic demonstratives: dass, dat, that

      b.   Slavic *wh*-words: čto (Russian), co (Polish) (Citko 2008)

      c.   Romance *wh*-words: que, qui (Rooryck 2000)

      d.   Indo-Aryan *wh*-words: ki (Marathi) (Bayer 1999)

A plausible analysis of the cases where a *wh*-element can be a nominal *wh*-expression or a finite complementizer is that this element is a complex head with properties of both C and D (Citko 2008). More simple, perhaps, is a late-insertion analysis (Halle and Marantz 1993): these elements

are specified for nominal ([φ]) features and little else, and thus meet the criteria for insertion in both verbal and nominal environments. Functional elements in many languages can be found in multiple functional positions, suggesting that they are not lexically associated with a particular category. For example, Abney (1987:28) observes that in Yupik the same agreement suffix marks ergative case and subject agreement (in the verbal domain) and possessive agreement (in the nominal domain):

(79)   a.   Angute-t   kiputa-a-t.
              man-ERG.PL buy-O-S
              'The man bought it.'

        b.   angute-t kuiga-t.
              the   man-ERG.PL river-S
              'the men's river'

From a functional perspective, since the complementizer occurs only with subordinated clauses, both C and D could be analyzed as the element enabling their complement to serve as an argument. This idea is explored in, e.g., Szabolcsi (1987) and Wiltschko (2011). I explore a version of this idea in section 7.6, relating the subordinate status of CPs and DPs introduced by these expressions to the property of dependent (vs. deictic/contextual) anchoring.

Indeed, the category of the complement they select has been proposed to be the only difference between C and D: C selects TP and D selects NP, but their feature content and relationship to argument-marking is the same (Pesetsky and Torrego 2001). If Haeberli (1999), Pesetsky and Torrego (2001), Lecarme (2004), Ritter and Wiltschko (2009) and others are correct that DPs bear the feature [T], then C and D in fact have identical features. In English-type languages, [uT] on both C and D is valued by Agree with the valued occurrence of [T] on T.[10]

---

[10]These authors also explore the possibility that some languages, e.g., Amharic and Halkomelem, have interpretable instances [T] on D. In these languages, DPs can bear grammatical tense, and their distribution is not constrained by the traditional Case Filter. The facts are complex however, as Ritter and Wiltschko (2014) show, and even languages with morphologically tensed DPs manifest some version of the finiteness-case connection. It appears that even if DPs can have inborn valued [T]-features, they are still in an anchoring dependency with clausal functional heads.

7.3.1.2 Prepositions in C

Nevertheless, the existence of non-finite clauses introduced by prepositional elements casts doubt

on a model deriving the nominality of CP from the lexical properties of the head of C. The non-

finite clauses in (80a and b), headed by *to* and *for*, occur in the nominal subject position.

(80) a. **To** succeed in this class requires a great deal of discipline.

b. **For** you to pass would surprise me.

c. **Because** he left, she is sad.

d. **Without** us reading it carefully, we won't understand it.

e. **By** taking her studies seriously, she will improve her grades.

f. I wish he hadn't **of** done that.   (Kayne 2005)

Similarly, the non-finite complementizers in Romance languages are prepositional:

(81) a. Elle a   essayé   **de** partir.
          she have.3SG.PRES try-PART of depart-INF
          'She tried to leave.'   French

b. Mi aiuti   **a** capire   questo   libro.
    me help.2SG.PRES at understand-INF this.M.SG book
    'You help me to understand this book.   Italian

The prepositional nature of these complementizer elements makes it implausible that the nominal

distribution of CPs like (80a and b) is due to the features of the element occupying C at spell-

out; rather, it is due to the [ɸ]-features of C itself. The parallel between the nominal and ver-

bal domains can be maintained if we analyze prepositional elements, particularly case-assigners

like *to, for*, and *of*, as originating in the topmost layer of a nominal or verbal projection that is

anaphorically anchored—one that anchored to an anaphoric deictic context. The formally nomi-

nal determiner/complementizer elements originate in the displacement domain (D/T), and occur

in structures that are dependently anchored to a deictic context, e.g., to the utterance time or to the speech event participants (Bianchi 2003; Sigurðsson 2004; Wiltschko 2014; see section 7.7 for details).

7.3.1.3 Determiners in T

D-elements can also be found in T. Many languages have a copula in T (in certain types of copular sentences) that occurs elsewhere as a pronoun or demonstrative. For example, consider the Russian, Hebrew, and Haitian Creole examples below (see also Benmamoun 2008 on Arabic).

(82)   Russian

     a.   Mark Twain, eto   Samuel Clemens.
         Mark Twain COP Samuel Clemens
         'Mark Twain is Samuel Clemens.'   (Geist 2007:89 (31))

     b.   Eto   chelovek chital   kniga
         DEMON man   read-IMFV kniga
         'That man was reading a book.'

(83)   Hebrew

     a.   Dani hu   hamore
         Dan COP theteacher
         'Dan is the teacher.'   (Doron 1986:313 (1))

     b.   hu   kara   et   hasefer
         PRON read.PST DOM thebook
         'He read the book.'

(84)   Haitian Creole

     a.   Jan se   zanmi mwen.
         Jan COP friend my
         'Jean is my friend.'   (Déprez 2003: (29))

     b.   Se   pou mwen.
         PRON for me
         'It is for me.'   (Déprez 2003: (10))

In fact, there are a number of formal parallels between tenses and D-elements. Tenses and pronouns show similar semantic behavior. Tenses, like pronouns, can be identified by a contextually given assignment function, or they can be bound. Both tenses and pronouns participate in feature transmission under binding (Sequence-of Tenses) (Partee 1973; von Stechow 1995; Kratzer 1998; Ogihara and Sharvit 2012).

Tenses and D elements also show the same proximate vs. distal interpretation. A present tense, 1st/2nd person, or proximate demonstrative indicates overlap with some reference element, while a past tense, 3rd person, or distal demonstrative indicates non-overlap.

The parallel also reveals itself in the morphosyntax of tense. Zagona (1990) and others treat tenses as expressions of category ZP, a DP-like category including a null *time/event*-denoting expression and its definite restrictor. In Germanic, the analogy between the Z restrictor and D manifests itself in the phonological form of past tense affixes. In Dutch, for example, both include the form [de] (Marcel den Dikken, p.c.; see also discussion in Kayne 2015).

(85) a. faal-**de**

         'fail-ed'

    b. **de** man

         'the man'

The past tense ZP means: 'THE time that is before some reference time.'

The fact that these determiner-like elements map to the same domain of structure in verbal and nominal contexts suggests a deep similarity in the way they associate with that domain's abstract function (Wiltschko 2014). These elements map to the displacement domain of the clause or nominal projection, where they encode different dependent anchoring relationships. They relate the tensed clause or definite DP to the larger structure, including a reference context encoded in the

higher CP/KP layer. When the reference context is *also* dependent—in a subordinate clause or an ordinary nominal expression in English—I will claim, elaborating the proposals of Pesetsky and Torrego (2001, 2007) and Wiltschko (2014), that anchoring requires the head of the displacement domain (T or D) remerges in the context domain (C/Fin or K).

7.3.2 Featural similarities between DP and CP: [ɸ] and [T]

It uncontroversial that C can bear [ɸ]-features, which are prototypically associated with DP.[11] In West Flemish, for example, finite complementizers show subject agreement inflection.

(86) a. dan-k   (ik) goan
          that-1SG (I)   go
          'that I go'

      b. da-se   (zie) goat
          that-3SG.F (she) goes
          'that she goes'

      c. dan-ze   (zunder) goan
          that-3PL (they)   go
          'that they go'   West Flemish (Haegeman 1992)

This pattern is far from exotic—complementizer agreement can be found in a range of language families. In the split-CP model I assume in the present work, such inflection on the complementizer reflects the dependency between Fin and T. In the Feature Inheritance model, for instance, all probing features originate on phase heads, to ensure that value and transfer happen simultaneously, and within the same spell-out domain—otherwise, as this model argues, uninterpretable features would creep into subsequent cycles and cause a crash at the interface (Chomsky 2001, 2008; Richards 2007). The [ɸ] probe on T, therefore, originates on Fin. Other models of clause structure attribute the presence of [ɸ]-features on Fin to an Agree or selection relation between Fin

---

[11]It is important here to note that I assume that C *can* bear [ɸ]-features. In root clauses, I will argue that C does not bear [ɸ]-features at the end of its cycle.

and its complement T (Rizzi 1997; Landau 2004; Adger 2007).

In a recent line of research examining the role of both C and D with respect to the finiteness-case connection, Pesetsky and Torrego (2001, 2007) propose that the syntactic dependencies CP and DP both participate in stem from the presence of the same formal features in the C and D domains. They show how an interplay between the formal features [φ] and [T] on *both* clausal and nominal heads can derive the distribution of nominative case effects.

Their main thesis is that nominative case is unvalued [T] on D. Both C and D are dependent on T for the valuation of their [T]-features. This dependency is essential for deriving not only nominative case / subject licensing patterns, but also the subject-object asymmetry between overtly-headed and null-headed CPs: the so-called "*that*-omission asymmetry".

I will briefly describe their approach in the present section, and show in section 7.4 how it cannot be reconciled wholesale with the existence of bare CP subjects in copular amalgam sentences. In section 7.6, I show how some of their main insights can be preserved, to allow bare CP subjects, subordinate CP subjects, and DP subjects to overlap in distribution only in amalgam copular sentences.

In Pesetsky and Torrego's (2001, 2007) framework, D and C-heads in English enter the derivation with unvalued, uninterpretable [T]-features, which must be valued and deleted by the end of the cycle. D-heads also bear valued, interpretable [φ]-features; and C bears an unvalued, but interpretable counterpart. In the course of the derivation, both C and D enter an Agree relation with T, which bears the interpretable occurrence of [T].[12]

Just as complementizer elements can inflect for [φ], they can also inflect for [T] in several languages. Consider, for example, the Irish data below.

---

[12]In Pesetsky and Torrego (2001), T is treated as bearing a valued occurrence of [T]; the formalization of Agree is different in Pesetsky and Torrego (2007), which more explicitly argues against a biconditional relating *valued/unvalued* to *interpretable/uninterpretable*.

(87)   a.   Deir   sé go dtógfaidh sé an peann.
              say.PRES he that take.FUT he the pen
              'He says that he will take the pen.'

        b.   Deir   sé gur   thóg   sé an peann.
              say.PRES he that.PST take.PST he the pen
              'He says that he took the pen.'   Irish (Cottell 1995, cited in Adger 2007)

Additional evidence for the presence of [T] on C comes from selection. Since the [±T] value of a clause is visible to selecting predicates, it must be marked on C (or Fin), assuming that selection is strictly local.

(88)   a.   She thinks [that he left].

        b. *She thinks [for him to leave].

(89)   a. *She is eager [that he left].

        b.   She is eager [for him to leave].

The occurrence of a formal feature [T] on D is supported not just by the distribution of nominative case, but also by the existence of overtly marked and "interpretable" [T] on the DP in some languages (Lecarme 2004; Wiltschko 2003, etc.). I will ultimately reject an analysis that treats interpretability as the key to the distribution of [T], but the occurrence of a valued morphological feature [T] on D is not in dispute.

Because of their similar feature composition, CPs and DPs overlap in distribution with respect to many environments, but since their [T]-features can be valued in different configurations (D's by an external T-head, and C's either by a phrase in its specifier or by the head of its complement), CPs are not dependent for licensing on "case" in the same way as their DP argument counterparts.

To take one example, for the derivation of the root CP in (90) to converge, C's [T]-feature must be valued:

(90)   You are tired.

Since C's feature is valued in the root context, where no external source of the valued [T]-feature

is present, then it is also valued in clause-internal contexts lacking [T], e.g., complement to verbs,

regardless of their ability to assign "case".

(91)   a.   It seems you are tired.

         b.   I think you are tired.

         c.   I doubt you are tired.

         d.   It's too bad you are tired.

The distribution of [T]-features on D, T, and C in Pesetsky and Torrego's system, along with some

ancillary assumptions about Agree, movement, and feature interpretability, predicts a complex set

of facts concerning root vs. non-root and subject vs. object asymmetries in Standard English and

Belfast English. Although this system offers no clear route to an analysis of bare CP subjects of

C (or Fin, in the split-CP model I adopt here), its empirical coverage is broad, uniting very old

problems like the *that*-trace effect under one analysis. I therefore follow Pesetsky and Torrego in

treating CP subjects as goals for Agree with clausal functional heads.

7.4   The *that*-omission asymmetry

No analysis of CP arguments is complete without careful examination of the "*that*-omission" asym-

metry.[13] As I showed in section 7.2, English is much more liberal than other Germanic languages

in its ability to license CP arguments. Moreover, English permits bare CPs in a wider variety of

environments than its Germanic relatives (Thráinsson 2007; Haider 2010).

---

[13]The same asymmetry applies to non-finite CPs headed by *for* or *to*; nevertheless, I use "*that*-omission" as a shorthand to represent all of these patterns.

Bare CPs in English occur in a variety of complement environments, as well as some adjoined positions, but they are categorically impossible in canonical subject position. It has also been observed that an otherwise optional complementizer becomes obligatory in certain cases of extraction and ellipsis. This section will briefly review existing analyses of the *that*-omission asymmetry, and demonstrate how the existence of bare sentential subjects in amalgam sentences requires a new explanation.

Analyses of *that*-less argument clauses in English fall into two main camps: the IP camp and the CP camp. The IP camp argues that there is no C projection in *that*-less clauses, while the CP camp argues that *that*-less clausal arguments are CPs, but that the complementizer head position is (phonetically) empty.

7.4.1 The asymmetries

7.4.1.1 Complement positions

It is well known that bare CPs can occur in complement positions to some verbs (92), but not, at least for many speakers, in complement to some non-bridge verbs, complement to N, or complement to P ((93)–(95)). Variable judgments are indicated with the percentage sign. Examples marked with the percentage sign are typically reported as ungrammatical without the complementizer, but I do not find them severely degraded, and have observed these structures in casual speech.

(92)   a.   I think (that) she's great.

          b.   I figured (that) he had left.

          c.   They believed (that) you were a vegetarian.

          d.   They convinced John (that) she was from Mars.

          e.   She reminded them (that) I needed to leave at noon.

(93)   a.   I regret %(that) you came.

          b.   She whispered %(that) she was hungry.

(94)   a.   They disputed the claim %(that) he had lied.

          b.   I don't know where he came up with the notion %(that) I would agree.

(95)   a.   They're talking about %(that) he will win.

          b.   There's still the problem of %(that) I ran out of money.

While I will not discuss nominal specificational copular clauses like (95b) in this thesis, the acceptability of a bare sentential subject in an inverse nominal predication structure is telling. It is strongly reminiscent of its counterpart in the verbal environment, thus providing indirect support for my claim that bare sentential logical subjects can be base-generated as small clause subjects.

Even complement positions that do admit bare CPs require the complementizer under certain conditions. For example, if a phrase within the CP is topicalized, the finite complementizer must occur:

(96)   a.   I think *(that) to Mary, you should give a book.

          b.   They heard *(that) tomorrow, you are leaving.

That is also obligatory with a complement clause selected by a gapped verb (Bošković and Lasnik 2003), (97).

(97)   Mary thinks that John is nice, and Sally __ *(that) Bill is.

7.4.1.2 Adjoined/specifier positions

Bare sentential arguments are illicit in most specifier positions, where they do not appear to be the complement of any functional head. In topic position and subject position of both finite and non-

finite clauses, CPs must categorically be introduced by an overt functional element, as in (98) and (99) (Ross 1967; Stowell 1981a; Pesetsky 1982; Webelhuth 1992; Doherty 2000; Bošković 1997b; Pesetsky and Torrego 2001).[14] Specifier positions in the vP domain, e.g., the position targeted by raising to object, likewise reject bare sentential arguments (100).

(98) a. *He left, even Mary doubted.

      b. *She needed help, only he believed.

(99) a. *It's raining ruined our party.

      b. *He should leave without telling anyone would be very upsetting.

      c. *I consider he left without telling anyone to be very upsetting.

(100) a. *I want he left without telling anyone not to be true, but I'm afraid it is.

       b. *I expected they must buy their books online to be convenient.

In specifier positions, CPs must be introduced by an overt functional element.

For many speakers, bare sentential arguments are also unacceptable in the right dislocated position that is generally compatible with sentential arguments (101).

(101) a. It really surprised me %(that) he left.

       b. It seemed unlikely to everyone %(that) he would leave.

---

[14]Note, with respect to (98) that the bare CP is topicalized to the left of a focused matrix predicate. The focus particles *even* and *only* ensure that the remnant clauses are not interpreted as slifting remnants, which are backgrounded and provide evidential information. Slifted clauses are root-like, similar to amalgam subjects (Ross 1967; Haddican et al. 2014).

(i) a. She needed help, HE BELIEVED.
     b. She's great, I THINK.

Bare sentences can occur as object relatives (102), but for many English speakers, subject relatives require an overt C element (103).[15]

(102)   a.   I like the boy (that) you met.

        b.   They tried some of the cake (that) she baked.

(103)   a.   I saw the lady %(that/who) lives next door to you.

        b.   There's a man here %(that/who) can help us.

The environments permitting bare sentences form a natural class with respect to requiring the complementizer once a topic or other element intervenes between TP and the embedding context.[16]

(104)   a.   She hopes *(that) this semester, she will earn a better grade.

        b.   She likes the boy *(that) unfortunately, they disparaged.

The bare CP amalgam subject does not belong to this natural class. Even with a fronted topic or a speaker-oriented adverb, no complementizer element occurs.

(105)   a.   (*That) this afternoon, it will rain is why our party is canceled.

        b.   (*That) unfortunately he left is the problem.

In the present analysis, this is because the amalgam sentential subject is a root—an independent clause. The mechanisms requiring the complementizer to be overt in the corresponding environ-

---

[15]Bare subject relative clauses like those in (103), also known as "subject-contact relatives" occur in some varieties of English, like Belfast English and Appalachian English, and have been analyzed by Doherty (2000) and den Dikken (2005a), among others.

[16]In (104b), I use a base-generated speaker-oriented adverb rather than a fronted topic, since topicalization and object extraction cannot co-occur:

(i)   *I like the boy that to Mary they disparaged.

I do not know whether subject-contact relatives with sentence-initial material force the C-element to occur, but I suspect this is so.

ment in canonical clauses are not at work in amalgams; in fact, the dependent anchoring status signaled by *that* is incompatible with the deictically anchored interpretation of the counterweight clause of an amalgam. I now turn to existing explanations for the *that*-omission asymmetry.

7.4.2 Explanations: The IP camp

One simple approach to the *that*-omission asymmetry is the IP analysis. Its simplicity makes it appealing for canonical embedded clauses, but it ultimately fails to offer a route to understanding sentential subjects in amalgams. This line of explanation treats bare sentential arguments as IPs, rather than empty-headed CPs (Webelhuth 1992; Doherty 2000; Bošković 1997b).

Webelhuth (1992) derives constraints on the distribution of bare IP arguments through restrictions on category features. IPs are [+V], while CPs are [−V]. Assuming that structural subjects must be [−V] (recall Stowell's (1981b) Case Resistance Principle), IPs cannot be subjects. An additional component of the proposal prevents IPs from occurring in other illicit positions, like dislocated environments and topic positions. First, sentences must bind [+N] traces (sentence anaphors are also plainly nominal), and all members of a movement chain must agree in categorial features. Given these assumptions, IPs can never move. They leave [+N] traces, but they are [+V]. CP sentential arguments, however, are [+N]; hence, they can occur in and move out of environments where [+N] expressions are selected and licensed.

Doherty (2000) gives a different explanation, which is strikingly simple and does not suffer from the same under-generation problem as many other approaches:

(106)   Selected IP must be a complement of X^0 (C^0, V^0 ...) at all levels of representation.

          (Doherty 2000:38 (83))

He argues that Webelhuth's (1992) analysis is too restrictive, failing to account for Honorary NP subjects, licit IP complements to nouns (including subject-contact relatives), and the fact that CPs and IPs can be coordinated, which is unexpected if they differ in their basic categorial features. Environments that meet the description "complement of X^0" but nevertheless fail to license bare IPs are explained in terms of lexical restrictions.

In principle, this analysis raises no problems for bare sentential subjects in amalgams, since they are not *selected*. They are small clause subjects. Then again, clausal subjects of predicational small clauses are likewise not selected, but they must be introduced by an overt complementizer.

(107)   *He never calls me back depresses me.

Bošković (1997b) and Bošković and Lasnik (2003) assume in a more contemporary Minimalist framework that direct IP complementation is possible when economy allows, but it is heavily restricted. For Bošković and Lasnik (2003), it occurs only when the clausal argument is the complement of a functional projection, precisely in such constructions as (108), under the assumption that the complement of the lexical verb is a small clause of sorts housing the two objects.

(108)   She persuaded John [his pants were on fire].

Since the amalgam subject is not a complement, this approach does not account for its distribution.

Finally, as I argued in Chapter 2, there is abundant evidence that the bare sentential subjects of amalgams are not IPs, but rather, root CPs. They can include subject-auxiliary inversion when interrogative, and they can support fronted topics. The IP-approach to bare sentential arguments therefore does not solve the puzzle of the bare sentential subject of amalgams. I will pursue a feature-based analysis instead.

7.4.3   Explanations: The CP camp

7.4.3.1   Empty Category Principle explanations

Early Government and Binding analyses of the asymmetry in the CP camp attribute it to the Empty Category Principle (ECP) (Stowell 1981a; Kayne 1981). Bare CPs are only permitted where the empty C head can be properly governed. Head-government, in the complement position of a lexical verb, ensures that the emptiness of C is licit. Where this government is blocked, e.g., by an adjoined topic or by gapping of the governing verb, the empty category in C gives rise to an ECP violation. Since N cannot head-govern the head of its complement, sentential complements to N must be introduced by overt C-heads: examples like (94) are generally unacceptable.

Pesetsky (1982) gives an explanation that relies on affix-hopping: null C is affixal, and must attach to a sufficiently local host, e.g., the immediately superordinate V, in order to be licensed. Locality restrictions then constrain the environments where affixal null C can occur. Affixal C cannot occur in a subject clause, for example, because if it moves out of the subject island in order to be licensed, its trace cannot be properly governed and an ECP violation results. This approach runs into problems when object relative clause islands are considered, where null C is perfectly acceptable. (See Bošković and Lasnik (2003) for discussion and a more contemporary implementation of Pesetsky's 1982 approach.)

Another empty-category-based proposal is offered by Landau (2007), who argues that the EPP reduces to a selectional requirement holding at PF: the head of a phrase in an Agree relation and attached to the specifier some other head is p-selected, meaning that it must have phonetic content. A phonetically empty C head violates this condition on Agree and the EPP, so a *that*-less CP cannot serve as a subject.

Clearly, none of these approaches can handle the pattern in copular amalgam sentences, where

a null-headed CP serves as a subject. These approaches also tend to undergenerate—they do not make much room for parametric variation that could allow, for instance, bare CP complements to nouns and prepositions, and bare subject relative clauses.

7.4.3.2 Pesetsky and Torrego's model

Let us now revisit Pesetsky and Torrego's (2001, 2007) analysis of the C-T relationship, and their contemporary analysis of the *that*-omission asymmetry. While I do not ultimately adopt several of the details of their proposal, the key insights—that T can remerge in C, and that clause types can vary with respect to the EPP specifications of their probes—prove integral to the analysis I develop for sentential subject licensing in copular amalgams.

7.4.3.2.1 Derivation of nominative case

Consider the following derivation of a clause with an ordinary DP subject, presented in Pesetsky and Torrego's (2007) framework, where D and T's unvalued [T]-features Agree, and then are co-valued in a second Agree relation between T and the finite verb. C's unvalued [T]-feature can then be valued by D's newly valued occurrence of [T]. Since both T's [T] and C's [T]-features are assumed to bear the EPP property, valuation of these two features is associated with remerge of the goal. In the derivation below, the [T]-feature is prefixed by *u* or *i*, meaning 'uninterpretable' or 'interpretable', respectively. The square brackets following the feature indicate its value—empty brackets indicate that the feature is unvalued.

(109)   a.   [$_{TP}$ [$_T$ *i*T[ ]] [$_{vP}$ [$_{DP}$ *u*T[ ] Sue] [$_v$ *u*T[pres] buys the book]]]]

   T probes the unvalued [T]-feature on D and Agrees; the features are covalued.

   b.   [$_{TP}$ [$_T$ *i*T[x] [$_{vP}$ [$_{DP}$ *u*T[x] Sue] [$_v$ *u*T[pres] buys the book]]]]

   The EPP property of [T] on T requires the DP goal to remerge in its specifier.

c. [$_{TP}$ [$_{DP}$ uT[x] Sue]$_i$ [$_T$ iT[x] [$_{vP}$ e_i [$_v$ uT[pres] buys the book]]]]]

T probes again and finds the valued features on *v*. Agree results in covaluation of the features on T and D.

d. [$_{TP}$ [$_{DP}$ uT[pres] Sue]$_i$ [$_T$ iT[pres] [$_{vP}$ e_i [$_v$ uT[pres] buys the book]]]]]

Merge C, which has unvalued [T], also with the EPP property. C probes the subject DP and attracts it to Spec,CP.

e. [$_{CP}$ [$_{DP}$ uT[pres] Sue]$_i$ [$_C$ uT[pres] [$_{TP}$ e_i [$_T$ iT[pres] [$_{vP}$ e_i [$_v$ uT[pres] buys the book]]]]]]

Uninterpretable, valued occurrences of [T] delete.

7.4.3.2.2   Derivation of T-to-C movement

C also has the option of Agreeing with the head of its complement T. In certain clause types, according to Pesetsky and Torrego, the EPP property of C is satisfied by remerge of the head of T. For example, in matrix interrogatives and embedded object questions, T-to-C movement serves the same purpose as subject movement in matrix declaratives:

(110)   [$_{CP}$ [$_{DP}$ What]$_j$ [$_{T+C}$ iT[fut] will [$_{TP}$ [$_{DP}$ uT[fut] Sue]$_i$ [$_T$ [$_{vP}$ e_i [$_v$ uT[fut] buys e_j]]]]]]]

In (110), T bearing the interpretable feature [fut] remerges in C following Agree between C and T. C now bears an interpretable occurrence of [T], which is not deleted, but remains active as a potential goal for Agree operations in the next cycle. By treating the movement of the subject DP and the head of T to C as triggered by the same feature (the EPP property of [T] on C), Pesetsky and Torrego offer a straightforward account of the unavailability of T-to-C movement in *wh*-subject questions. Since the subject expression bears both valued [wh] and [T]-features, C can value both of its features against the same goal; by economy, this option is preferred over

valuing [wh] by Agreeing with the subject and [T] by agreeing with T, and triggering two separate

EPP-movements. The derivation of the subject question in (111) is therefore quite similar to the

derivation of the declarative in (109), apart from the presence of the additional [wh]-feature.

(111) $[_{CP}$ $[_{DP}$ uT[pres], $i$wh[wh] Who $]_i$ $[_C$ uT[pres], $u$wh[wh] $[_{TP}$ e_i $[_T$ iT[pres] $[_{vP}$ e_i $[_v$

    uT[pres] buys the book?]]]]]]

The subject question in (112), by contrast, with both subject movement and T-to-C movement, is

ruled out by economy.

(112)   *Who does buy the book?

    In object questions, there is no complementarity between *wh*-movement and T-to-C movement,

because the *wh*-element in Spec,CP does not bear nominative case; that is, it does not bear an

occurrence of [T] in an Agree relation with T (or indirectly, with v, in the 2007 version). T-to-C

movement in this case values [T] on C.[17]

(113)   What did he buy?

The same logic applies to all non-subject questions in Standard English, which require T-to-C

movement in addition to *wh*-movement.

---

[17]An obvious question at this point is why movement of the nominative DP subject to Spec,CP (assuming multiple specifiers) cannot value [T] on C in object questions, just as it does in declaratives and subject questions. In other words: why should the availability of T-to-C movement render nominative DP movement impossible? Pesetsky and Torrego (2001: (43)) point out that in fact, under the assumptions of their framework, nominative DP movement and *wh*-object movement to Spec,CP do co-occur in exclamatives, which are incompatible with T-to-C movement:

(i)   a.   What a silly book he bought!
       b.   *What a silly book did he buy!

They propose the following generalization to account for this pattern:

(ii)   "A matrix CP whose head bears uWh is interpreted as an exclamative if a non-*wh*-phrase appears as one of its specifiers. Otherwise, it is interpreted as a question."

This generalization is empirically adequate, but warrants further investigation.

(114)   a.   Why did he buy it?

        b.   Where did he buy it?

        c.   How did he buy it?

7.4.3.2.3   Parametrizing EPP

By endowing C and D heads with unvalued [T]-features and parametrizing the EPP properties for
different clause types and different dialects, the Pesetsky and Torrego system represents several
interrelated sets of facts. It addresses why T-to-C movement occurs in matrix, but not embedded
questions, and why the distribution of T-to-C is different in Belfast English. In embedded interroga-
tives in Standard English, C's [T]-feature lacks the EPP property, so no displacement of nominative
DP or the head of TP takes place. Compare (115) and (116).

(115)   a.   I wonder what Sue bought.

        b.   I wonder why Sue bought it.

(116)   a.   *I wonder what did Sue buy.

        b.   *I wonder why did Sue buy it.

The absence of the EPP property on [T] of embedded C allows C to probe and value its [T]-feature
without inducing movement. In Belfast English, Pesetsky and Torrego propose that EPP is present
on embedded C's [T], reversing the pattern:

(117)   a.   I wonder what did Sue buy.   *Belfast English*

        b.   I wonder why did Sue buy it.   *Belfast English*

7.4.3.2.4 Blocking bare sentential subjects

Now, let us turn to the application of this system to the distribution of subject CPs, and to the question of why bare CPs cannot generally serve as subjects, which this section has yet to address directly. For Pesetsky and Torrego, the presence of [T] on C is what allows some CPs, but not others, to serve as subjects. Clauses in which C bears an undeleted occurrence of [T] are eligible to serve as subjects, since they can Agree with higher T and C heads. If all visible occurrences of [T]—those on the edge and head of the CP phase—are deleted at the end of the CP cycle, then the CP cannot be a goal for higher Agree operations. This is why CPs in which the subject in Spec,CP values C's [T]-feature cannot be subjects: since [T] is uninterpretable on D and C, both of these occurrences of [T] delete at the end of the CP cycle once valued, so they cannot in turn value [T] on the superordinate C. The abbreviated structure in (118) illustrates the problem.

(118)   *[$_{CP}$ [$_C$ **uT[ ]** [$_{TP}$ [$_{CP}$ [$_{DP}$ uT̶[̶p̶a̶s̶t̶]̶ Sue]$_i$ [$_C$ uT̶[̶p̶a̶s̶t̶]̶ [$_{TP}$ e_i [$_T$ iT[past] [$_{vP}$ e_i [$_v$ uT̶[̶p̶a̶s̶t̶]̶

         bought the book]]]]]]$_j$ [$_T$ uT[past] [$_{vP}$ ... upset him]]]]]

In this sentence, [uT] on matrix C cannot satisfy its EPP feature, because the structural subject lacks [T].[18]

     Grammatical CP subjects are those that bear an occurrence of [T] on their edge. This means that either the phase head or the phase edge must have an undeleted [T]-feature. Examples of

---

[18]It must of course be stipulated that the head-movement strategy for satisfying EPP is unavailable in matrix declaratives, in contrast to matrix interrogatives. Otherwise, the following ungrammatical version of the sentence above would be generated:

(i)   *Did Sue buy the book upset him.

Pesetsky and Torrego must also account for the fact that the counterpart of (118) with an interrogative interpretation of the structural subject is ungrammatical, despite the fact that T-to-C movement is licit in interrogatives, and leaves an occurrence of T in C. The other piece of this problem is that the structural counterpart of (i) with a subject question expression in the edge of the subject clause is grammatical.

(ii)   What happened upset him.

grammatical CP subjects are given in (119).

(119)   a.   That Sue bought the book upset him.

           b.   For Sue to buy the book would upset him.

           c.   To buy books is a good investment.

           d.   What he bought surprised her.

First, consider the *that*-ful CP subject in (119a). Pesetsky and Torrego observe that fronted auxiliaries and *that* are in complementary distribution with nominative *wh*-movement in embedded clause environments in Standard English, but not Belfast English.

(120)   a. *I don't know who did read it.   *Standard English*

           b. *I don't know who that read it.   *Standard English*

(121)   a. *I don't know who did read it.   *Belfast English*

           b. *I don't know who that read it.   *Belfast English*

This pattern strongly suggests that the mechanism of T-to-C movement and the distribution of *that* are related. The analysis that Pesetsky and Torrego give is that they are not just related: they are one and the same. Just like T-to-C movement pronounced as a fronted auxiliary, the finite "complementizer" *that* is an instance of T moved to C. C is always null in English. Clauses introduced by *that* are clauses with T-to-C movement, the end result of which is an interpretable occurrence of [T] in C. Since [T] on T is interpretable, head-movement of T to C yields a CP that is eligible to serve as a goal for Agree with higher T and C heads. This is Pesetsky and Torrego's explanation of the so-called "*that*-omission" asymmetry: the reason that *that*-ful CPs can serve as subjects and objects, while *that*-less CPs can serve only as objects:

(122) a. That Sue bought the book upset him.

       b. He believed that Sue bought the book.

(123) a. *Sue bought the book upset him.

       b. He believed Sue bought the book.

A similar account is provided for the non-finite clausal subjects in (119b)–(119c), but I will not spell out the details here.

Now consider the *wh*-CP clausal subject in (119d), repeated here:

(124) [What he bought] surprised her.

In this example, no T-to-C movement has taken place in the subject CP, and yet it is able to value [T] on matrix C. T-to-C movement of auxiliaries only takes place on matrix interrogatives, whose [T]-feature bears the EPP property. The [T]-feature of C in embedded interrogatives lacks EPP in Standard English, so we do not find sentences like:

(125) *[What did he buy] surprised her.

If uninterpretable features must delete at the end of the cycle, and C only contains an instance of interpretable [T] if a T-head remerges in it, then embedded interrogatives should not be possible structural subjects. An ancillary assumption about the lifespan of features is proposed to account for the fact that embedded interrogatives pattern with *that*-ful declaratives: only uninterpretable features with the EPP property are required to delete at the end of the cycle:

(126) A feature marked for deletion as a consequence of an operation (Agree, Move) must disappear at the end of the CP cycle if it has the EPP property. Otherwise it may delete at the end of the derivation.   (Pesetsky and Torrego 2001: (64))

This proposal allows embedded interrogatives to pattern with *that*-ful declaratives as potential structural subjects, because both retain an instance of [T] on their edges after the spell-out cycle. In Belfast English, by contrast, where embedded interrogative C bears an EPP-ful [T]-feature, it is predicted that only interrogative CPs with subject-auxiliary inversion should be possible subjects. Interrogatives where a nominative DP values C's [T]-feature have no interpretable instance of [T] on their edge, just like bare declarative CPs in Standard English, by the condition in (126), so they should not be possible subjects:

(127)   a.   *[What he bought] surprised her.   predicted, *Belfast English*

            b.   [What did he buy] surprised her.   predicted, *Belfast English*

This prediction remains to be verified.

The Agree-based system developed in Pesetsky and Torrego's work relates a range of puzzling facts, and relies on Minimalist principles, so it provides a useful analytical framework for sentential arguments. Since the impetus for the finiteness-case connection in this model lies with unvalued features of both clausal C and argument D, there is an avenue toward explaining why non-DP arguments show some, but not all of the "case"-dependence of DPs: some of the Agree and displacement operations between clausal functional heads and arguments serve to value D's features, but others serve to value those of T and C, so the latter obtain even when the argument is not a DP.

7.4.3.2.5 Some criticisms of the Pesetsky and Torrego system

The formal details of Pesetsky and Torrego's (2001, 2007) system raise some issues, however, which the revised model I put forward in the present work puts to rest. First, by downplaying the role of [φ] in the 2007 proposal, and leaving all the work of deriving nominative case and subject-

object asymmetries to [T], some empirical coverage is lost, as are some conceptually appealing aspects of the 2001 proposal. If only [T] is involved in deriving the distribution of case-licensed DPs, then it remains to be explained how nominative DPs are licensed in infinitival constructions, where they agree with the infinitival for [ϕ]-features, as in Portuguese.

(128) Era importante eles   sairem.
       was important they.nom leave-inf-3pl
       'It was important for them to leave.'   (Raposo 1987)

It is more conceptually appealing to treat the finiteness-case connection as a cross-referencing relationship between prototypically nominal features ([ϕ]) and prototypically verbal features ([T]). Indeed, both features are exploited in a parsimonious way in Landau's (2004) analysis of control and nominative case.[19] The peculiarity of copular amalgams is that they cross-reference verbal features with verbal features.

Secondly, the Pesetsky and Torrego (2001, 2007) formalization of Agree relies on assumptions that may not be warranted. They propose a version of Extreme Functionalism, in which all features must have some interpretable occurrence; all uninterpretable occurrences must delete, or else they cause a crash at LF. As den Dikken (2014b) argues, however, there is no compelling reason to require uninterpretable features to delete. Moreover, valued uninterpretable features are necessary at PF to ensure that the correct vocabulary items are inserted, under a Distributed Morphology implementation of spell-out (not assumed in Pesetsky and Torrego's 2001 model). Feature valuation is a prerequisite for spell-out, but a deletion mechanism that is sensitive to semantics and obligatory is harder to defend, especially if its putative effects find another explanation.

---

[19]He argues convincingly that it is the interplay of [ϕ] and [T] that predicts that in some environments, PRO alternates with independently referring (case-marked) DPs.

I propose an alternative analysis of the effects of the putative deletion of uninterpretable [T] on C/Fin. This analysis resurrects [φ]. Not only is such an analysis *adequate*, obviating the feature life-span proposal in Pesetsky and Torrego (2001); it is *necessary* in order to capture the fact that bare sentential subjects are licensed in amalgam clauses: precisely where they serve to value an instance of unvalued [T] on Fin.

7.5   The form of the sentential subject in amalgams

The amalgam specificational copular sentence is the only environment, to my knowledge, where bare sentential subjects occur with impunity.

(129)   a.   [He left] is the problem.

b.   [It's going to rain] is why our party is canceled.

c.   [She needed help] is what she needed.

d.   [That's what has to happen] is we have to try again.

(130)   a.   *[He left] upset me.

b.   *[It's going to rain] has disrupted our party plans.

c.   *[She needed help] occurred to him.

d.   *[That's what has to happen] is generally understood.

The subject is a root-like CP, which nevertheless occurs as the structural subject of the finite functional head spelled out by Fin. It bears valued [T], but is not specified for [φ]. These features are behind its limited distribution as a structural subject.

7.5.1 The amalgam subject: embedded root

In Chapter 2, I presented evidence that the amalgam subject is an embedded root clause. It shows a range of root properties. For instance, it cannot be introduced by the finite complementizer *that*, and interrogative amalgam subjects show root, rather than embedded word order (Hooper and Thompson 1973; Heycock 2006, etc.).

(131)   a.   She fetched a pail of water is what she did.

        b.   *That she fetched a pail of water is what she did.

        c.   Why did she fetch a pail of water is what is at issue.

        d.   Why she fetched a pail of water is what is at issue. (canonical structure only)

Likewise, in *that's x is y*, where the structural subject is the weight, rather than the counterweight, it must have root syntax.

(132)   a.   That's what she did, is she fetched a pail of water.

        b.   *That that's what she did, is she fetched a pail of water.

        c.   Is that what she did? is she fetched a pail of water?

        d.   *Whether that's what she did? is she fetched a pail of water?

I have also argued that the counterweight of the amalgam pseudocleft is embedded, despite its root-like properties. The simplest piece of evidence is that the amalgam pseudocleft construction allows recursion: a counterweight can be found within a counterweight.

(133)   a.   And [[what they've done is [they've bought... Mercedes, after Mercedes, after Mercedes]], is what they've done].   (Lambrecht and Ross-Hagebaum 2006:20)

   b.   [[That's my feeling], is [[that's the most rewarding way to harvest deer], is [to track it in the snow, and get it]]].[20]

One familiar context where formally root expressions do not function as roots is in quotations, but I argued in Chapter 2 that the sentential subject of the amalgam pseudocleft is not a quotation. If all counterweights were direct quotations, then they should freely allow indexicals to refer to the quoted utterance's context, rather than the matrix (direct report) context, (134). In amalgams, however, even where the whole counterweight serves as the value, indexicals refer to the matrix context, (135).

(134)   a.   "What will $I_{k/*I}$ do tomorrow?" she_k wondered.

   b.   *"What will she_k do tomorrow?" she_k wondered.

   c.   "My$_{j/*I}$ car has been stolen!" exclaimed John$_j$.

   d.   *"His$_j$ car has been stolen!" exclaimed John$_j$.

(135)   a.   What will $I_{*k/I}$ do tomorrow is what she_k wondered.

   b.   What will she_k do tomorrow is what she_k wondered.

   c.   My$_{*j/I}$ car has been stolen is what John$_j$ exclaimed.

   d.   His$_j$ car ??has/had been stolen is what John$_j$ exclaimed.

   e.   My$_{*j/I}$ car has been stolen is what John$_j$ said was stolen.

These facts support an embedded-root analysis of the counterweight clause. What is particularly unique in amalgams is that an embedded root can occupy the subject position.

---

[20]`http://www.howcast.com/videos/460894-Top-9-Tips-for-Deer-Hunting-Hunting`

7.5.2 Amalgam sentential subjects lack [ɸ]-features

The crucial asymmetry between bare and overtly headed CPs that I exploit in the analysis of amalgam subjects is that bare CPs lack [ɸ]-features. Evidence for the absence of [ɸ]-features on the amalgam sentential subject comes from the distribution of number agreement. While the copula in amalgam sentences is typically singular, the copula can take plural form when the weight is a plural DP.

(136)   a.   **The main problems are** he missed the shot and he argued with the referee.

b.   He missed the shot and he argued with the referee **are the main problems**.

Recall from the discussion in Chapter 4 that the sentential subject cannot trigger plural agreement alone. An acceptability experiment showed that a plural copula in combination with a conjunction of *that*-CPs was more acceptable than it was with a conjunction of bare sentential subjects. The experimental results were corroborated by small-scale judgment surveys, which allow for more pragmatic support. I showed that when prosody and pragmatics facilitate the construal of the coordinated propositions as two separate objects, rather than a notionally singular object, the asymmetry is clear.

(137)   A: Is the reason he's upset just that he missed the shot?

B: No, that he is in trouble with the COACH AND that he has a KNEE injury **is/are** why he's upset.

(138)   A: Is the reason he's upset just that he missed the shot?

B: No, he is in trouble with the COACH AND he has a KNEE injury **is/*are** why he's upset.

Further support for this observation comes from *that's x is y*, where the structural sentential subject is not the logical subject. In this sentence type, there is no confound from an underspecified weight expression. There is no potential source for plural other than the coordinated *that's x* weights. Nevertheless, a conjunction of *that's x* weight clauses cannot value plural. (139), with *are*, is strikingly bad.

(139) That's the problem and this is what I was telling you about, **is / *are** we can barely afford our rent.

A bare sentential subject lacks [φ].

7.5.3 A natural class of amalgam structural subjects?

One aspect of the sentential subject licensing puzzle in amalgams that needs to be taken into account is that there are three categories that can occupy the subject position, which do not fall into any salient natural class. Although only bare CPs are base-generated as the subject of Fin, embedded CPs and DPs can come to occupy the subject position of Fin when it is remerged in Top. Bare CP and ordinary subjects are not in complementary distribution (Subsection 7.5.3).

	Bare-CP	*that*-CP	Indirect question	DP
Canonical	✗	✓	✓	✓
Amalgam	✓	✗	✓	✓

Table 7.1: Subject types in canonical vs. amalgam sentences

Since Top does not ordinarily bear its own [φ] and [T]-features, while Fin does, I assume that its features in inverse copular amalgam sentences are contributed by the remerged Fin head (which receives its features in a binding relationship with Force). Given this analysis of the source of Top's features, the head spelled out by the copula has the same probing features whether it takes the bare

sentential subject or the inverted predicate as its specifier. As such, whatever Agree relation the bare sentential subject participates in must be compatible with the licensing of embedded CP and DP as well.

This is a puzzling pattern. Under Pesetsky and Torrego's (2001) approach, for instance, as I illustrated above, there should be no way to value the [T]-feature of Fin in a sentence like (140).

(140)   I'm just too tired is the problem.

Fin does not take TP as a complement, and hence, no instance of T is available to be remerged in Fin. *The problem* is a predicate in this sentence, so it does not need to be formally licensed, but even if it did, its [T]-feature is unvalued. The bare sentential subject *I'm just too tired* likewise has no instance of T remerged in its phase head; it has only a nominative DP in an outer specifier position. This nominative DP has valued [T] due to its Agree relation with the T head in its clause, but under Pesetsky and Torrego's (2001) assumptions about feature lifespan, the [T]-features on D and the C head of the sentential subject are deleted, since they are uninterpretable, so they cannot value the [T]-feature of matrix Fin. The derivation cannot be salvaged under their approach.

I will now turn to my analysis of bare sentential subject licensing, which does not rely on feature interpretability and deletion. Section 7.6 proposes a theory that allows both root clause subjects and ordinary subjects in amalgam sentences, and section 7.7.3 discusses why either the logical or structural subject in amalgams in fact *must* be a root.

7.6   How to allow bare sentential subjects in amalgams

According to the analysis developed in the present thesis, the crucial difference between amalgam and canonical sentences is that in amalgams, the T and V-domains are absent: they are projected directly from the C-domain of the clause. This chapter asks: why should a bare CP subject be

compatible only with a C-domain head? Moreover, what is the crucial contrast between bare and overtly headed CPs that distinguishes their distribution? The task of this section is to explain the unexpected grammaticality of bare CPs in amalgam subject position, while maintaining the ungrammaticality of bare CPs in canonical subject position.

To develop an explanation, I must address two questions regarding the featural makeup of C/Fin, since the two key players in amalgams are two Fin heads.

(i) What is the featural makeup of a canonical Fin at the end of its cycle when it is a root vs. a non-root?

(ii) What is the featural makeup of a root Fin that does not select T?

I derive the complementary distribution of root vs. non-root clauses in canonical vs. amalgam subject position by appealing to the distribution of the EPP property, which can be satisfied in precisely three different minimally local configurations:

(141)   a. $[\text{Fin}_{[uF(EPP)]} [\text{TP} [\text{XP}_{[F]}][\text{T} [... t_{XP}]]]]$ Spec of complement

         b. $[\text{Fin}_{[uF(EPP)]}+\text{T}_F [\text{TP} [\text{T} ...]]]$ Selected head movement

         c. $[\text{Fin}_{[uF(EPP)]} [\text{XP} [\text{X}_F ...]]]$ Unselected complement in-situ

7.6.1   EPP and feature inheritance

The Minimalist enterprise holds that the formal features of C and T are intimately related, but there is no consensus regarding the precise nature and directionality of the relationship. In some approaches, the features of clausal functional heads enter the derivation valued, and their value is constrained by the selectional properties of higher heads (Landau 2004); in others, features are inborn but unvalued and supplied with values via Agree (Pesetsky and Torrego 2001, 2007; Adger 2007). Approaches to the valuation strategy also differ: features may be transmitted (up

the tree or down the tree) through mechanisms like projection/remerge (e.g., Grimshaw 1997, 2000; Broekhuis 2013), inheritance (Chomsky 2001, 2008; Richards 2007; den Dikken 2014b), or spreading (Tortora 2014).

7.6.1.1 Motivation for feature inheritance

Feature inheritance, a mechanism introduced in Chomsky (2001), is motivated by the fact that nominative subjects obligatorily displace (in English-type languages) to Spec,TP, yet they do not interfere with the strictly local relation between C and T. Assuming that only phase heads are inherent probes, Chomsky (2001) proposes that the feature triggering the displacement of subjects to Spec,TP originates on C. He argues that the fact that displacement to Spec,TP, indicating a probing feature on T under standard implementation of the EPP, is obligatory follows from the Principle of Full Interpretation. Full interpretation requires that the object transferred to the Conceptual-Intentional Interface contain no illicit objects: unvalued or uninterpretable features. Valued interpretable and valued uninterpretable features are formally indistinguishable from each other—they can only be distinguished by information found in their derivational cycle. Because a cycle is impenetrable to further operations, the key information distinguishing interpretable from uninterpretable features is unavailable once the cycle is spelled out. To ensure that no valued but uninterpretable instances of a feature sneak into a higher phase, for example, by way of having EPP-moved to the specifier of the phase head, Chomsky (2001, 2008) and Richards (2007) argue that value and transfer must happen simultaneously: in the same phase. This requirement renders feature inheritance obligatory when a phase head bears an EPP probe.

Den Dikken (2014b) argues that the conceptual motivation behind requiring value and transfer to happen together is weak. It is neither empirically nor conceptually clear why valued but uninterpretable features at the interface should cause the derivation to crash. Moreover, valued but

uninterpretable features must be preserved at the PF interface under a late-insertion approach to morphology (Halle and Marantz 1993). If uninterpretable, but valued features are not catastrophic for the CI interface, then there is no reason to require valuation, deletion, and transfer to happen simultaneously.

7.6.1.2   EPP configurations

According to den Dikken (2014b), the distribution of EPP offers a stronger motivation than Full Interpretation for some version of feature inheritance. If valuation must happen as early as possible (in keeping with what Pesetsky and Torrego 2001 propose), then a phase head should value its unvalued features in the minimal structure created by merge with its complement. This means that if a probing feature like [φ], responsible for subject displacement, has the EPP property, it cannot be satisfied in the minimal structure where Fin merges with TP, because its goal, the subject DP, is more deeply embedded in the TP. Feature inheritance presents a solution: the probing feature can be passed to the phase head's complement, allowing Agree and EPP movement to take place to the specifier of its complement. This is a strictly local configuration for Agree, configuration (a) in (141), above.

Another configuration that is compatible with the requirement that EPP be satisfied in the minimal structure created by merge of a phase head with its complement is head movement, configuration (b), above. Head movement and traditional EPP movement can thus be triggered by the same requirement (Pesetsky and Torrego 2001). The head-movement derivation is not implicated in subject licensing in the discussion given in Chomsky (2001), however. Fin cannot satisfy the EPP property of its [φ]-feature through head movement, because its complement [T] does not bear a valued instance of [φ]. This will be relevant for the derivation of different clausal subject types shortly.

Assuming that Fin bears a probing [φ]-feature with the EPP property, the subject in a canonical clause must displace to Spec,TP, the phase head's complement. Crucially, I assume that in root clauses, this leaves the Fin without [φ] at the end of its cycle. Feature inheritance does not leave a copy of the feature on the phase head (it differs from feature spreading; Tortora 2014). Fin also bears a probing [T]-feature, which is valued by Agree with its complement in ordinary clauses: T bears inherently valued [T].[21]

The configuration in (c) above is reserved for instances in which a phase head takes a complement without selecting some property of its head. For example, in a coordinate structure, the conjunction may inflect, in some languages, and behaves for the purposes of locality like a phase head, but it does not trigger displacement.[22] What distinguishes it from other phase heads is that it does not select for any particular category. I propose that Fin in copular amalgams, which is inflected, but does not s-select its complement, can value its EPP properties in the strictly local head-complement configuration.

I follow Pesetsky and Torrego in exploiting the distribution of the EPP property in different clause types to derive different patterns with respect to the distribution of FinP. I depart from their proposal in two main regards: first, an interpretability/uninterpretability contrast does not determine the distribution of valued features at spell-out (features are not deleted upon valuation); secondly, feature inheritance is exploited to clarify the configurations that satisfy EPP.

7.6.2 Derivation of clauses with sentential subjects

I illustrate the proposal by comparing the derivation of a non-root declarative and a root declarative, using Fin instead of C. Embedded declarative Fin bears unvalued [φ] and unvalued [T], both with

---

[21]On an implementation that assumes that V is the source of valued [T]-features, a different derivation would be necessary, for instance, one that uses Pesetsky and Torrego's (2007) model of covaluation. Either way, T bears a valued instance of [T] by the time its projection is selected by Fin.

[22]I am not aware of any independent evidence that the features of the coordinator are EPP-specified, however.

the EPP property. Fin merges with T, which is inherently valued for [T], but does not bear [φ]. [φ] continues to probe, and finds the subject DP in Spec,vP.

Since the EPP property of [φ] cannot be satisfied under head movement of T to Fin, feature inheritance takes place, which endows T with the probing [φ]-feature that triggers displacement of its DP goal to Spec,TP. Fin's [T]-feature, on the other hand, can satisfy its EPP property without feature inheritance: local head to head movement takes place within the minimal structure created by merge of Fin and its complement, hence it respects the locality constraints on EPP. Since T-to-Fin movement takes place, Fin contains an instance of T, with the result that Fin now bears valued instances of both [φ] and [T].

This derivation predicts that embedded finite clauses can themselves be structural subjects, a prediction that is borne out, as we saw in section 7.2. It is also compatible with the existence of tensed complementizers—if the counterparts of English *that* in other languages are similarly instances of T remerged in a C-domain head, then it is wholly unsurprising that the morphology should reflect their [T] and/or [φ]-features.[23]

Root declaratives differ from embedded declaratives in one way: the [T]-feature of Fin does not bear EPP. This small difference predicts their distribution in canonical clauses and amalgams. Just like their embedded counterparts, root declaratives require feature inheritance to satisfy the EPP

---

[23]The form of the [φ]-features on the T+Fin complex does not follow so clearly, if the cross-linguistic facts are considered closely. More recent studies of complementizer agreement in West Germanic show that Fin and T can bear different [φ]-feature values, for instance, in the environment of a coordinated subject (Carstens 2003; Koppen 2005; Haegeman and van Koppen 2012). T must agree with the whole coordinate structure, while Fin can agree with the first conjunct. In light of these facts, the [φ]-features of Fin and T appear to probe separately. Further research is needed to determine whether these facts can be reconciled with feature inheritance.

One avenue for future research is that since these other West Germanic languages, despite their genetic similarity to English, do not allow Fin-headed amalgams, a component of the feature inheritance analysis may be parametrized. For instance, in some languages, perhaps feature inheritance only bleeds the higher probe's EPP specification, but leaves an otherwise intact copy of the probe on the head. If this is the case for West Flemish, then T's inherited EPP-bearing [φ]-probe would have to attract the whole coordinate subject into its specifier (to avoid violating the coordinate structure constraint; Haegeman and van Koppen 2012), but the Fin-head's [φ]-probe, now lacking EPP, would then be sufficiently local to Agree with just the first conjunct of the subject, without inducing further movement.

property of their [φ]-probe. This results in subject displacement to Spec,TP, and an absence of [φ] on Fin. Fin's [T] probe Agrees with its complement T, but no head movement takes place, since it has no EPP property. Consequently, [φ] remains on T, and does not end up on Fin as a result of remerge. The resulting object bears [T]-features, but no [φ].

The proposal that root Fin's [T] probe lacks EPP has the consequence that root clauses cannot serve as goals for [φ] probes in other cycles. The [φ]-features inherited by T in this derivation remain on T—they are not carried back to Fin via remerge of T in Fin. The distribution of the EPP property thus derives the fact that a root clause cannot be a canonical structural subject: it lacks [φ], so it is not eligible to serve as a goal for a T head that has inherited [φ] from a higher Fin.

Recasting this purely technical analysis in terms of the conceptual status of the different functional domains, the canonical clause derivation establishes a relation between two displacement domain elements associated with the dependent anchoring function. A DP and a TP are dependent on superordinate structure providing a deictic context for anchoring. When the clausal subject is a bare CP, with a purely independent context domain (no remerger of T in the left periphery), it is incompatible with dependent anchoring—it cannot occur in the nominative structural subject position.

This proposal predicts that a root clause can be the structural subject of another root clause, crucially restricting this possibility to the amalgam-type derivation. To avoid confusion, I will call the matrix Fin, ultimately spelled out by the copula, Fin_1 and the embedded root Fin in the sentential subject, Fin_2. Fin_1 enters the derivation with unvalued [T], which lacks the EPP property, and valued [φ], which has it. Instead of TP, its complement is the amalgam predicate: an indirect question CP or a DP (predicate of the propositional subject in the amalgam small clause). Both of these expressions bear valued [φ]-features. DP has inherently valued [φ]-features (just as T has

inherently valued [T], see section 7.7 for brief discussion), and Fin in an indirect question has its [φ]-features valued by the fronted *wh*-expression.

The structural definition of EPP is satisfied: Fin_1 Agrees with its unselected complement, satisfying the EPP property within its minimal merge domain. Unlike when Fin merges with T, head movement is not triggered in this case. The difference between these two derivations is that canonical Fin stands in a (c-)selectional relationship with T, while in amalgams, Fin_1 does not select its complement. I must therefore assume, given the pattern of head movement, that head movement satisfies EPP only under selection. Given ordinary constraints on head movement, a strictly local process that typically applies within the limits of the traditional Extended Projection, this assumption is plausible. It would be surprising to find the head of an unselected maximal projection remerging in the domain of a different clause. Future research will determine whether independent evidence for this proposal is available. Hypothetical evidence would come from an agreeing Relator element that can take either a selected or an unselected complement.

Fin_1 must also value its [T]-feature, which lacks EPP. Since it lacks the EPP property, this [T]-probe can be valued in a wider range of Agree configurations than the [φ]-probe: it need not be a complement of Fin_1, the head of Fin_1's complement, or the specifier of Fin_1's complement. In fact, the configuration that values Fin_1's [T]-feature in this case is spec-head agreement: the sentential subject of the amalgam is base-merged in Spec,FinP.

7.6.3   Predicate inversion around Fin

As I discussed in Chapter 5, the Relator phrase (RP) of the amalgam specificational copular sentence functions quite similarly to its canonical counterpart with respect to the availability of predicate inversion. I will now briefly revisit the derivation of weight-initial amalgam pseudoclefts in light of the feature-based analysis of Fin.

Since the RP is a phase, predicate inversion is facilitated by domain-extending head movement. This is clearly necessary in amalgams, where the Relator head is Fin. When Fin remerges, an instance of domain-extending head movement, domain extension renders the predicate eligible to raise to the edge of the phase: now Spec,TopP.

(142) $[_{TopP} [_{CP}$ PREDICATE $] [_{Top+Fin}$ is $[_{FinP} [_{CP}$ SUBJECT$] [t_{Fin} [t_{predicate}]]]]]$

EPP is not the impetus for predicate inversion, since all of the probing features are valued in the minimal domain of Fin. Information structure is instead the trigger for inversion. Since the counterweight is interpreted as the focal answer to a question denoted by the predicate, the predicate must be interpreted as a topic. When domain extension makes the Spec,TopP sufficiently local to the predicate, it raises in order to occupy its canonical discourse position.[24]

7.7   The root vs. non-root asymmetry and the anchoring function

In Chapter 5, I analyzed the amalgam pseudocleft copula as an instance of Fin, based on its interactions with the abstract functions of the displacement and context domains of the clause. Using the analytical approach developed in detail by Wiltschko (2014), I proposed that the functions of the displacement domain (dependent anchoring to a context) are not instantiated by the copular clause, unlike in an ordinary tensed clause. In section 7.3, I sketched how striking parallels emerge between elements of D and T that participate in dependent anchoring, as well as between elements of lexical category P spelled out in the nominal and verbal context domains, K and C, where they

---

[24]An alternative that I will not pursue here for reasons of space would feature a base-generated predicate in Spec,TopP and a null associate in the complement of Fin. This gets around the problem of motivating inversion, but it is difficult to determine what would distinguish the two analytical possibilities. Such an analysis of fronted argument clauses is motivated by Moulton (2013) since they fail to exhibit the full range of reconstruction effects that would expected if they had undergone A′-movement. Comparable reconstruction tests are not available for copular amalgams, because the whole counterweight clause c-commands the predicate in its base position. Since predicate inversion in amalgam pseudoclefts does not violate any constraints, I assume it can apply without an EPP-trigger.

encode anaphoric anchoring.

This section briefly discusses the relationship between the formalization of subject agreement in section 7.6 and the root/non-root asymmetry in CPs, with the goal of distinguishing the functions of dependent finite anchoring (displacement) from independent deictic anchoring. These functions are argued to belong to the displacement (DP/TP) domain and the context (KP/CP) domains, respectively (see also discussion in Wiltschko 2014). English, which lacks independently anchored nominals, and has extensive D-to-K and T-to-C movement, obscures this distinction. The copular amalgam clause, which removes the displacement domain from the structure, clarifies it.

7.7.1 The relationship between the displacement and context domains

As Chapter 6 points out, there is some confusion in the literature on finiteness about the distinction between a finite but formally dependent clause structure and a finite and independent clause structure. A characterization of finiteness that appeals only to the morphological specification of T/Infl does not offer a distinction; neither does a characterization that appeals only to the presence or absence of a positively specified head Fin in the structure. The functional implications of the root/non-root asymmetry in CPs can be formalized using the tools I developed for the analysis of copular amalgam subjects in section 7.6.

7.7.1.1 The displacement domain

The displacement domain, associated with T and D, carries out the *anchoring* function. It introduces a relation between a topical entity or situation/eventuality and a reference context.[25] A spatial metaphor for anchoring is useful: the event or individual either overlaps the utterance context, or it is removed from it. Temporal anchoring, familiar from ordinary clauses in most Indo-European languages, locates the event in time with respect to the context. As is well known, not all languages

---

[25] Anchoring is the main topic of Chapter 8, so the present discussion remains superficial.

anchor through tense (e.g., Bianchi 2003; Bittner 2005; Tonhauser 2011; Amritavalli 2013; Ritter and Wiltschko 2014)—the anchoring function can be realized by tense, or by other categories like location, mood, and person.

A finite clause must be anchored to a structurally local reference context, i.e., to a context variable in the CP/context domain. That context variable can be deictic, referring to the utterance context (the external logophoric center, in Bianchi's (2003) terms), or it can be anaphoric, bound by a structurally superordinate event associated with an embedding verb (an internal logophoric center). Anaphoric context variables are found in finite embedded clauses. The displacement predicate that mediates the anchoring relation signals dependent anchoring. Crucially for my approach, even present-tense anchoring, which does not intuitively involve displacement, involves formal displacement. The topic interval is located with respect to a context; the specific semantics of present tense happens to yield a meaning of overlap.

The displacement domain of the verbal projection has a parallel in the D-layer of the nominal domain. A fully specified displacement domain in the DP entails the presence of person features, which anchor the individual in the DP to the person elements of the utterance context. First and second person features, where the DP's referent overlaps with the utterance context, are the analogue of present tense, and third person features, where the DP's referent is not part of the utterance context, is the analogue of past tense. Similarly, demonstrative determiners in the displacement domain are interpreted as either proximate or distal. Bianchi (2003) and Sigurðsson (2004) propose that this parallel between temporal and nominal anchoring is at the heart of the grammatical/displacement domain's role in licensing independently referential subjects. A DP in the structural subject position of a finite clause has access to a deictic anchor in the context variable in Fin.

7.7.1.2 The context domain

The basic distinction in the verbal context domain is: [±deictic]. The context domain, as Bianchi proposes, encodes the logophoric center. The logophoric center is either deictic (referring to the utterance context in a finite clause) or anaphoric (bound by a superordinate context in a dependent clause). What is the counterpart of this alternation in the nominal domain?

Wiltschko (2014) points out that the deixis distinction in the verbal domain is formally similar to a deixis distinction manifested by certain DPs in languages like Squamish. In Squamish, there is a class of determiners that encode different logophoric centers. These are the so-called deictic determiners. The deictic determiner that is the counterpart of $\text{Fin}_{[+fin]}$ (*ta*) is used for referents that are part of the context (familiar from the discourse, present in the utterance context, or introduced into the context as novel referents), while the deictic determiner that is the counterpart of $\text{Fin}_{[-fin]}$ (*kwi*) is used for non-specific indefinites and non-referential noun phrases.

English lacks a functional category in the nominal domain that shows such an alternation. Demonstratives in both English and Squamish have a deictic meaning: they encode the proximate vs. distal location of noun phrase referents. These meanings correspond to those associated with the displacement head T. They locate noun phrase references with respect to the logophoric center—they do not encode distinctions in the *type* of logophoric center available in the context domain. It follows that English DPs are subject to dependent anchoring, like subordinate clauses. Dependent anchoring of a DP is structural case: the DP occurs in the environment of a verbal functional head that is anchored to the external logophoric center. Dependent anchoring in DPs is parasitic on dependent anchoring in the clause.

Distinguishing between dependent anchoring relations and logophoric center types predicts that a language like Squamish should not adhere as rigidly to the "finiteness-case connection" as a

language like English. The discussion of these issues in Wiltschko (2014) and Ritter and Wiltschko (2014) suggests that this hypothesis is on the right track.

7.7.2 T-to-C and D-to-K

The context domain and the displacement domain are intimately related in canonical structures. Dependent anchoring of a finite clause, in the Agree-based model laid out in this chapter, relies on T-to-C movement. A clause with an anaphoric context attracts the [φ]-bearing T-head (this feature bears the EPP property), so the dependent CP ends up with nominal features. The nominality of the dependent CP and its need to anchor to a context give it a similar distribution to DP arguments.

Similarly, the context domain in the noun phrase is intimately related to the displacement domain. Since English makes available only a non-deictic context variable in ordinary nominals, the D head is always remerged in K, like its counterpart in the verbal domain. This is consistent with the fact that it is the [φ]-features of D, and not some higher projection, that participate in Agree relations outside the DP.

An independent K domain is available in nominals that are used vocatively (see also discussion in Wiltschko 2014). Since the [φ]-probe on K lacks EPP, by analogy with its counterpart on C/Fin, no element of D is remerged in the edge of a vocative. At least in English, this is suggested by the fact that although definite, vocative expressions are not introduced by definite determiners.

(143)   a.   You guys, why don't you come in?

b.   *You the guys, why don't you come in?

c.   *The guys, why don't you come in?

Vocatives are by definition deictically anchored to the utterance context: to the addressee. Since they are deictically anchored, they cannot be dependently anchored; hence, they do not occur in

structural subject position, (144b).

(144) a. Hey John! You look hungry.

   b. *Hey John! look/looks hungry.

Dependently anchored CP and KP behave on a par with respect to the presence of a [φ]-bearing displacement domain element in their periphery. Independently anchored CP and KP, on the other hand, are compatible only with the context domain of a higher clause. I set aside the non-finite counterparts of the dependently anchored CP and KP discussed here, for reasons of space.

7.7.3 Root/non-root asymmetry in amalgam sentential subjects

Copular amalgams, where the sentential subject argument has the form of an independent (anchored) clause, despite occurring in a syntactically dependent position, provide a unique window into the function of the context domain. Recall that root and *that*-ful clauses are in complementary distribution in the structural subject position of canonical and amalgam clause types. I have offered an explanation for the ungrammaticality of root CPs as ordinary subjects, and their grammaticality as amalgam subjects. Given the proposal above, what rules out *that*-ful sentential subjects in amalgams?

(145) *That I need some coffee is what I need.

Sentences like (145) give rise to a very strong intuition that it is the nominalized, dependent status of the *that*-ful CP that is the problem. Its featural makeup, after all, is not the culprit: its Fin_2 bears both [φ] and [T] as a result of T-to-Fin movement, so it should be a licit goal for Fin_1's [T]-feature. The featural difference between licit root CP subjects and illicit non-root CP subjects is the presence of [φ], but it cannot be the case that a [φ]-bearing expression is barred from the

precopular position of an amalgam, because in predicate-inversion contexts, the indirect question CP or concealed question DP occurs there.

I propose that while the absence of [φ]-features is responsible for the ungrammaticality of root CP subjects in ordinary clauses, the absence of a deictically anchored context variable is responsible the ungrammaticality of non-root CP subjects in amalgams. Subordinating complementizer elements are associated with dependent anchoring; although *that*-ful clauses are finite in the traditional sense (they are anchored), they are not *deictically* anchored. When these elements of T (syncretic with dependent elements of D) come to occupy C, the resulting CP is no longer capable of serving as an independent utterance. As a dependently anchored expression, it must be in the scope of a deictic context variable—like a structural subject, it needs access to the external logophoric center. (Recall again the analysis in Sigurðsson 2004; the DP needs access to speech act participants in the logophoric center; the CP needs access to speech time.) Only deictically anchored expressions are licensed in the marked structural subject position of the amalgam clause.

7.8 Conclusion

The goal of this thesis is to bring data from copular amalgam sentences to bear on the theory of what makes a clause *finite*. Consequently, this chapter has focused on one of the key phenomena associated with finiteness: formal subject licensing. The ability of a clause to license an independently referential (nominative) structural subject is closely associated with both the formal features of finiteness ([T, φ]) and the semantic anchoring property of finiteness.

Copular amalgam sentences, however, present a problem for the model of subject licensing that depends on Agree between T and the subject, since a bare sentential subject is not compatible with the canonical subject position. Indeed, bare sentential subjects are in near-complementary distri-

bution with overtly headed CPs and nominal subjects. The task of this chapter has therefore been to account for the fact that bare sentential subjects occur always and only in copular amalgams, while also explaining why ordinary CP and DP subjects can also occur in amalgam pseudoclefts where predicate inversion has taken place.

I began the chapter by defending the existence of sentential subjects, in order to support an analytical approach to amalgams in which the distributional properties of root and non-root CPs were compared. I also discussed the formal similarities between the CP and the DP. The analogy served two purposes. It allowed the feature composition of CP arguments to be elucidated by comparison to their DP counterparts, and it provided a framework for discussing the implications of dependent subject licensing as a function of the displacement and context domains of the clause (Wiltschko 2014).

The derivation of copular amalgam sentences borrows from the insights of Chomsky (2001, 2008), Pesetsky and Torrego (2001, 2007), and den Dikken (2014b), who implement slightly different versions of the dependency relationship between C and T, and its interactions with EPP displacement. The version I proposed treats root C/Fin as bearing an EPP-specified [φ]-probe, and a non-EPP-specified [T] probe (contra Pesetsky and Torrego 2001); non-root C/Fin has EPP-specified [φ] and [T] probes. Root Fin satisfies the EPP property of [φ] in canonical clauses in the configuration permitted by feature inheritance, where [φ] is passed to its complement T. When this happens, Fin no longer bears [φ]. Non-root Fin does the same. In addition, the EPP property of its [T] probe is satisfied by the remerger of its complement head T. This operation re-imbues Fin with [φ], and results in both its *ability* to serve as a structural subject, and its *dependence* on a higher clause's T. Since root Fin has [T] without EPP, no T-to-Fin movement is warranted. A summary of the features of each expression type is given in Table 7.2.

	Bare-CP	*that*-CP	Indirect question	DP
[T]	✓	✓	✓	✓
[ɸ]	✗	✓	✓	✓

Table 7.2: Features of subject types

Ultimately, bare CPs were distinguished from non-bare CPs and DPs by the absence of [ɸ]-features on their Fin head. Finite root CPs were further distinguished from finite dependent CPs and DPs by their ability to be deictically anchored to the utterance context, without the mediation of a displacement domain. In this way, root clauses are like the verbal counterparts of vocatives and caseless deictic DPs in Squamish.

This proposal makes the most typologically marked feature of the copular amalgam clause—its ability to take a bare finite clause subject—compatible with Minimalist syntax. By making the mechanisms of feature inheritance and EPP satisfaction explicit, this chapter characterized the root/non-root asymmetry in a way that predicts accommodates finite sentential subjects only in the left periphery of clauses that lack T. The unusual syntax of the amalgam clause's spine in fact predicts that it should have a bare finite clause as a subject.

This analysis raises questions for the nature of the anchoring property of finiteness. In the absence of the T/tense anchoring domain, how exactly are copular amalgam clauses anchored to the utterance context? What are their temporal interpretations? These questions motivate the final chapter of this dissertation, which offers a syntactic-semantic account of deictic anchoring. Chapter 8 explicitly defends my central thesis: that finiteness does not entail tense.

Chapter 8

Anchoring in the left periphery

8.1 Introduction: Finiteness and tense

In this thesis, I have argued that copular amalgam sentences instantiate an unusual type of clause, where a finite element, the copula *is/was*, relates a sentential subject and an indirect or concealed question in the absence of an articulated verbal functional spine:

(1)   a.   What he needs is he needs a break.

        b.   He needs a break is what he needs.

Chapters 4 and 5 presented evidence that the copula, though morphologically inflected for tense and agreement, cannot combine with functional material from the middle functional field of the clause. The proposal that the T domain is absent in copular amalgam clauses finds support from the fact that these clauses are precisely the only environment in English where a bare finite clause can be a structural subject. On the basis of these facts, I argued that the finite copula in amalgams is located not in T, but rather in the left periphery. The sentence in (1b) thus has the structure in (2).

(2)

```
                              ForceP
                             /      \
                      Op-ASSERT     Force′
                                   /      \
                              Force        FinP
                                          /    \
                                        CP      Fin′
                                       /  \    /    \
                          He needs a break  Fin    CP
                                             |     /  \
                                            is   what he needs
```
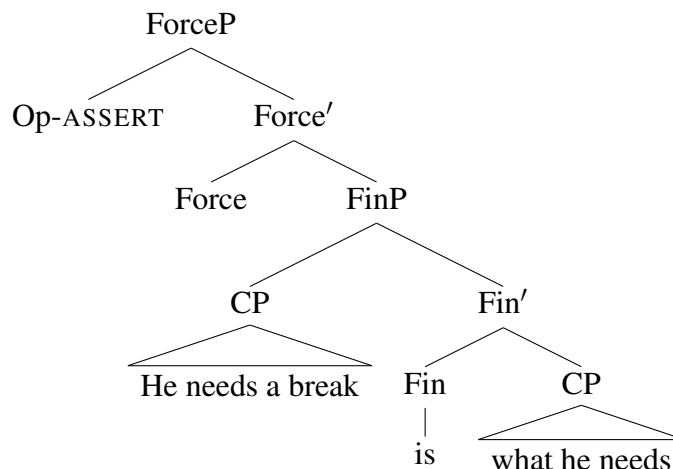
Chapter 7 showed that unlike overtly headed CP subjects, bare finite clauses lack [ɸ]-features, so

they cannot enter the canonical EPP-motivated structural subject relation with T. CPs do bear [T],

however, so they value the [T] probe on Fin through spec-head agreement.

This analysis of copular amalgams entails that a clause can be *finite* while lacking the syntactic

architecture for (nominal) argument licensing and verbal inflection. In other words, finiteness does

not entail the presence of Tense. This claim departs from current models of clausal architecture

in Minimalist syntax, which assume an implicational relationship between functional heads in the

C-domain and lower ones (see discussion in Chapter 6). To further examine the consequences of

the present proposal, this chapter focuses on the most basic property of finiteness: independent

anchoring to the utterance context.

The notion of clausal "independence" is a rather intuitive one, often invoked in traditional

grammars. It distinguishes between root and subordinate clauses. A finite clause's potential in-

dependence hinges on the presence of a context variable in a deictic relationship with the actual

utterance context. This deictic context variable is necessary for the proposition expressed by the

clause to be evaluated from the perspective of the participants in the speech event, and for it to be

bound by illocutionary force. An independent clause is formally characterized by the absence of a

subordinating complementizer: no dependently anchored T-domain element is remerged in Fin.

This deictic context variable is at the heart of a general requirement that propositions must be anchored. In early work on the syntax of anchoring, Enç (1987:642) proposes: "In main declarative clauses... events must be anchored to the utterance or some other salient reference point." The anchoring requirement is most often cast in temporal terms (Enç 1987) (emphasis mine):

(3)   Anchoring Principle

        "Each **tense** must be anchored." (Enç 1987:642)

The anchoring requirement holds of all clauses, but the anchor itself is different in different clause types. In independent, finite clauses, the anchor is deictic, while in dependent clauses, it is anaphoric to a reference context in the matrix clause. Truncated non-finite clauses have no anchor at all; their events depend directly on the superordinate clause for anchoring.

In most clause types and languages that have been investigated in the generative tradition, the anchoring relation is treated as temporal. Even in the absence of morphological tense, semantic tense (read off a projection of T) is implicated in other submodules of syntax, for instance, case and control theory. There is a tendency in the literature on finiteness to assume an implicational relation between finiteness (independent anchoring) and tense. Some works, however, (e.g., Comrie 1985; Bohnemeyer 2002; Bianchi 2003; Shaer 2003; Bittner 2005; Giorgi 2010; Jayaseelan 2013; Amritavalli 2013; Ritter and Wiltschko 2005, 2014; Wiltschko 2014) explore other means of anchoring. My contribution to this enterprise is the proposal that even in a tensed language like English, there can be finite, but tenseless declarative clauses.

This chapter takes the syntactic model for amalgams as a starting point, and examines its implications for anchoring at the two interfaces: syntax-semantics, and syntax-morphology. If the structure I propose for copular amalgam sentences is right, it is predicted that patterns of temporal interpretation associated with the copula and the embedded tense form in the copular amalgam

would behave differently from tense interpretation in other sentence types, particularly its closest relative, the canonical specificational copular sentence. The null hypothesis—that copular amalgams do in fact contain an ordinary T-layer—would predict instead that temporal interpretation in amalgam copular sentences behaves similarly to tense alternations in canonical copular sentences. In this chapter, I will provide experimental evidence that the null hypothesis must be rejected. Instead copular amalgams are anchored deictically to the speech event.

On the morphology front, the predictions of my analysis are less clear. We might expect two different things. First, the morphological tense form of the copula could be invariant, since it is not associated with syntactic T or temporal interpretation. However, as is clear from examples throughout this thesis, the copula's tense form *does* vary: it shows sensitivity to the tense form of the lexical verb in the weight or counterweight. On the other hand, since Fin's [T]-feature is valued by that of the counterweight (see Chapter 7, section 7.6), we might expect that the morphological form of the copula must match that of the subject. This is also not the case: the amalgam copula always has the option of taking present tense form. I will show that the unusual semantics-morphology mismatches associated with copular amalgams can be captured under my approach that dissociates morphological tense form from syntactic-semantic tense.

The present chapter has four main goals:

(i) to explicitly distinguish [tense] forms from semantic TENSE meanings in specificational copular sentences, including both canonical and amalgam clauses;

(ii) to argue that the copula of amalgams does not project temporal argument structure;

(iii) to argue that copular amalgam clauses are anchored to the utterance context in the modal domain via the *author* indexical rather than via *time*; and

(iv) to explain how morphological *tense* alternations on the copula in amalgams do not require a

   corresponding semantic alternation.

The chapter is structured as follows. First, in section 8.2, I describe a simple toolkit for analyzing patterns of temporal interpretation and anaphora resolution in ordinary finite clauses, focusing on English. Next, section 8.3 describes tense patterns in canonical specificational copular sentences, in order to establish a baseline for the analysis of amalgams. Since tense form and interpretation in canonical specificational sentences poses problems of its own, the discussion in this section makes an additional contribution by offering a new analysis. Section 8.4 compares canonical and amalgam copular sentences, presenting experimental support for the argument that the amalgam copula lacks temporal argument structure. Next, section 8.5 proposes that copular amalgams are anchored deictically, by way of the context variable in the left periphery of the clause. Lastly, section 8.6 sketches a concord account of the [present/past] form alternation on the amalgam copula, which is independent of semantic interpretation.

8.2   Tense in finite clauses: background

Although it is beyond the scope of this work to do justice to the vast literature on tense, some context is necessary for the discussion of anchoring in amalgams. The reader who is well acquainted with the pronominal and quantificational approaches to tense (Partee 1973; Enç 1987; Abusch 1988, 1997; Zagona 1990, 2003; Klein 1994; von Stechow 1995, 2002; Stowell 1996, 2007; Ogihara 1996; Guéron 2007; Kratzer 1998; Schlenker 1999, 2003b, 2004; Demirdache and Uribe-Etxebarria 2000, 2007, 2014; Lecarme 2004; Ogihara and Sharvit 2012) may wish to proceed to the interim conclusion in section 8.2.3.

   Since tense phenomena occur in the morphological, syntactic, and semantic components of

the grammar, I will use some typographical conventions throughout this chapter to indicate which module is under discussion. Occasionally, authors conflate *time* and *tense*; when possible, I will disambiguate these uses explicitly. Small capitals (e.g., PAST, PRESENT, TENSE) are used to refer to tenses in the semantic component; square brackets (e.g., [past], [present]) are used to refer to the morphological expression of tense; sentence case (e.g., Tense) is used to refer to the syntactic projection associated with temporal argument structure. Plain text is used to refer to the topic of tense in a more general way.

8.2.1   The primitives of tense and aspect

Syntactic and semantic theories of tense since foundational work by Reichenbach (1947) assume that natural language expressions can refer to *times*, objects that are ordered on a (dense) linear time-line, and to relations that order *times*.[1] In early work on the logic of tense (e.g., Prior 1957, 1967, 1969), simple TENSES were treated as weak sentence operators with existential quantificational force. They indiscriminately bind time variables in their scope.

(4)   John ate an apple.

        P[John eat an apple] $\longrightarrow$ There is some past time at which it is the case that John eats an apple.

   In an important development in temporal semantics, Partee (1973) observed that unrestricted existential force is too weak to capture the interpretations of TENSES, and that unrestricted universal force (e.g., when negation scopes over *tense*) is too strong. Consider the meaning of (5).

(5)   I didn't turn off the stove.

---

[1]Later, times are treated as tense intervals, but the ontology of instants vs. intervals is not central to the discussion in the present work, so it will be assumed henceforth that "times" are continuous, dense sets of intervals ordered with respect to a time-line.

    a. $\nrightarrow \neg \exists t[\text{past}(t) \, \& \, \text{I turn off the stove at t}]$

        'There is no past time at which I turned off the stove.'

    b. $\nrightarrow \exists t \neg[\text{past}(t) \, \& \, \text{I turn off the stove at t}]$

        'There is some past time at which I did not turn off the stove.'

(5a) is true only if I have never turned off the stove, and (5b) is trivially true if there exists at least one moment in the past where I didn't turn off the stove. Clearly, the meaning of (5) is that I have in mind some specific time at which I did not turn off the stove. Partee (1973) proposed that TIMES pattern with pronominals with respect to anaphoric dependencies: they can be bound or they can refer deictically to contextually determined intervals. This discovery led to the development of the pronominal approach to tense.

In current theories building on Partee's insight, two closely related approaches predominate. In the pronominal approach, *times* are treated as a pronominal expressions (e.g., Partee 1973; Dowty 1982; Guéron 2007; Demirdache and Uribe-Etxebarria 2000, 2007; Zagona 1990, 2003; Stowell 1996, 2007; Ogihara and Sharvit 2012). TENSES restrict the denotation of times, either as ordering predicates, or as presuppositional modifiers. In the quantificational approach, TENSES are existential quantifiers over times; the relative interpretation of a tense (i.e., PAST VS. PRESENT) is contributed by a presuppositional restriction (e.g., von Stechow 1995; Ogihara 1996; Kratzer 1998; Sharvit 2003). This is something of an oversimplification, as both approaches make use of both time arguments and quantification over times.

The present chapter will take a pronominal approach to tense: times are pronominal elements projected in specifier positions in the environment of T. TENSES are ordering predicates (prepositional elements). These predicates, existential quantification over times, and anaphora establish relations between time arguments within and across clauses.

8.2.1.1   Reference Time, Topic Time, and Event Time

What do times refer to, and how many are contained in a given clause? Following the tradition initiated by Reichenbach (1947) and developed in, e.g., Klein (1994), Stowell (1996), and Demirdache and Uribe-Etxebarria (2000, 2007), I assume that temporal and aspectual interpretations come about via the relationship between up to three time: the Reference Time (RT), a Topic Time (TT), and an Event Time (ET). These times are organized hierarchically in the clause. Reference Time is the anchor time: in tensed clauses, it encodes the time with respect to which a given eventuality is evaluated. In independent finite clauses, RT is deictic, referring to the time of the speech event (Utterance Time, UT). Topic Time (also called Assertion Time) encodes the contextually salient interval to which the assertion is confined—it represents, for instance, the interval the speaker of (5) has in mind when asserting that she did not turn off the stove (Klein 1994). Event Time (ET) is the interval associated with the event or situation restricted by the VP; it is located by (outer) aspect.

Consider the simple past tense sentence below:

(6)   Lou threw the ball.

It is obvious that the simple past orders the event of Lou's ball-throwing prior to UT; however, there is some disagreement in the literature as to whether all three times (UT, TT, and ET) are involved (Reichenbach 1947; Klein 1994; Kratzer 1998; Demirdache and Uribe-Etxebarria 2007), or whether the perfective aspect conveyed by simple tense with eventive verbs in English involves only UT and ET (Comrie 1985). The choice of analysis depends on how TENSE interacts with ASPECT in simple tense sentences.

8.2.1.2 Time in statives

Since the present chapter is concerned with tense in copular clauses, which are stative, let us examine a case of simple tense on a stative verb.[2]

(7)   Lou liked jazz.

Tense in (7) orders the interval at which Lou is asserted to like jazz prior to the time of utterance: $TT < UT$.

The relationship between TT and ET depends on what sort of aspectual interpretation is available in a stative sentence like (7). Such a sentence is potentially ambiguous between a perfective and an imperfective interpretation. Because *like* is stative, its default construal is imperfective. TT is a subinterval of ET (the run-time of Lou's liking jazz).

A useful metaphor for imperfective aspect is that the event is viewed from within (see, among others, Klein 1994; Kratzer 1998; Demirdache and Uribe-Etxebarria 2000, 2007, 2014):

(8)   Imperfective: TT WITHIN ET

       ⎯[⎯[⎯]⎯]⎯>

       ET   TT

---

[2]There is, of course, an agentive use of *be* in English, which forces an activity reading (Lakoff 1970; Partee 1977; Rothstein 1999):

(i)   a.   Sam is being noisy.
       b.   Pat is being an idiot.

This use of *be* occurs in ordinary predicational copular clauses, but is incompatible with specificational or equative copular sentences, e.g., (7), so I will not discuss it further in this work.

(ii)   a.   *What John likes is being pizza.
       b.   *Cicero is being Tully.   (unless *Tully* is a role in a game or play)

On the perfective interpretation, the event is viewed from the outside: its run-time is contained within the TT, or the two co-refer (Demirdache and Uribe-Etxebarria 2014).

(9)   Perfective: ET=TT

——[——]——>

ET/TT

8.2.1.3   Topic Time in copular sentences

Aspect will not play a large role in the discussion in this chapter, but it is useful to be explicit about the relationship between TT and ET in simple copular sentences in the interest of clarity. Consider (10):

(10)   Mary is smart.

In this sentence, the property of being smart is predicated of Mary. What is the role of tense and aspect in mediating this predication? What is the semantic contribution of the copula that tense associates with?

Rothstein (1999) proposes that the copula is a verb. Its function is to covert the state associated with the non-verbal (adjectival) predicate into an eventuality, an argument of a VP. This eventuality is then located by TENSE.

(11) $\lambda S \lambda e. \exists s \in S$ & locate(s)=e   (Rothstein 1999:372 (71))

        'There is a state in the domain of states that is instantiated in a located eventuality.'

Without an ET, for Rothstein, there can be no Tense. In Chapter 5, I argued that this semantic treatment of the contribution of the copula is flawed. The predicational small clause, without an overt verbal element, also includes a "located" eventuality/state expression.

(12)   [Hungry eight times a night], that newborn never let his parents sleep.

In Maienborn's (2005) view, copular clauses do not include eventualities at all. They are not amenable to spatial or manner modification, (13), or direct perception reports, (14), in contrast to canonical Davidsonian eventualities, (15).

(13)   a.   #John was hungry in my office.

        b.   *John was hungry efficiently.

(14)   *I saw John (be) hungry.[3]

(13a) has a peculiar depictive interpretation; *in my office* cannot felicitously modify the location of the state of hungriness.[4]

(15)   a.   John sat in my office.

        b.   John wrote efficiently.

        c.   I saw John sit/write.

Instead, Maienborn (2005) proposes that copular clauses include states.[5] These states are abstract objects, not real slices of space-time in the Davidsonian ontology. These abstract objects are identified as property-exemplifications, anchored to a property-bearer (the subject) and a time (the Topic Time). They can therefore be temporally modified, like events.

(16)   John was hungry for an hour.

---

[3]This is only acceptable on the agentive reading, where John is the agent of some event.

[4]When it is preposed, the locative modifier can have a "frame-setting function", in Maienborn's (2005) terms. It establishes a place where the proposition holds; it does not directly modify the state.

[5]This is a much simplified discussion of the proposal. My purpose here is simply to show that in copular sentences, there is no ET, only a TT. Moreover, canonical copular sentences include a TT.

The small clause introduces a property, a property bearer, and a state, and the copular projection, associated with T, provides a TT to pick out a relevant, temporally located portion of that state.

The tense predicate associated with the TT locates it with respect to the RT. In the copular clause *Mary is smart*, RT (in the context domain) is deictic. The simple present tense copula places RT/UT within the contextually located TT interval, where the state exemplifying the Mary-smart property is located (regardless of whether it is a subinterval of a larger *be smart* interval).

(17)   Simple present tense: UT WITHIN TT

—[——|—]———>

TT   UT

8.2.2   Tense at the interfaces

The preceding sketch introduced the basic relations between Reichenbachian time points, but tense in natural language has morphological, syntactic, and semantic components which need to be teased apart. Temporal structure in natural language, as described in the preceding subsection, involves time arguments, ordering relations, quantification over times, and anaphoric dependencies between times. Following the pronominal approach to tense, I treat times as ordered by predicates with the meanings WITHIN, BEFORE, and AFTER (Zagona 1990, 2003; Stowell 1996, 2007; Demirdache and Uribe-Etxebarria 2000, 2007, et passim).

Like pronouns, times can also be bound by quantifiers and lambda abstractors. Zagona treats times as expressions of category Z(eit)P: they are similar to nominals. The head of ZP is like D: it is a generalized quantifier, but it relates predicates of times (expressions of type i) rather than predicates of individuals (type e). I will now briefly sketch the syntactic implementation of tense I adopt in the present work.

8.2.2.1 The syntax of Tense

The role of T in the syntax is to relate the projection of the verb (vP or AspP), a property of times, to the located TT in its specifier.[6] This model of times in the clause follows that of Demirdache and Uribe-Etxebarria (2007) closely, but unlike their model, I do not locate the substantive temporal content in T; I locate it in a restrictive small clause modifier of TT. To take an analogy from this work, T serves as a Relator of temporal individuals and properties:

(18)

TP

ZP   T′

TT   T   AspP
before/after/within RT

property of some time within/after ET
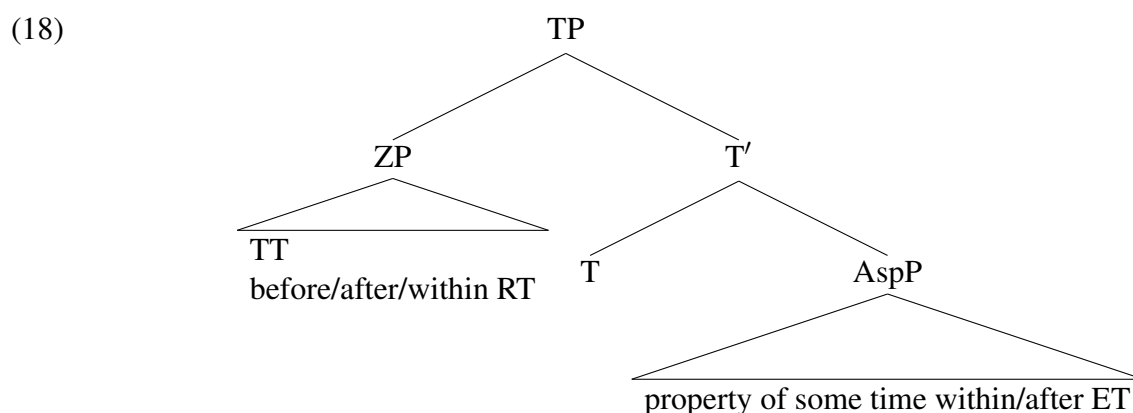
To make the analogy explicit, consider the parallel copular sentence in (19):

(19)

SubjP

He   Subj′

contextually restricted male individual
outside the utterance event   Subj   RP

is

property of some individual
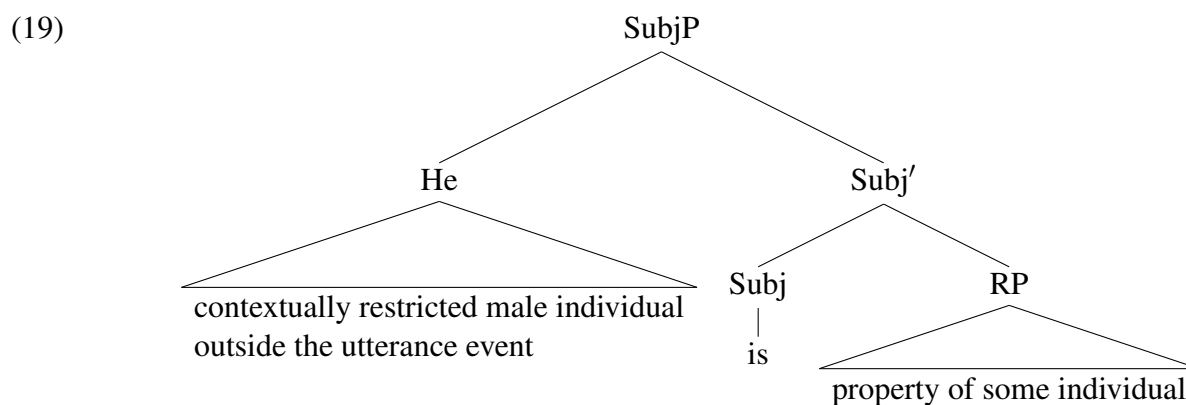
Notice in (18) that a time expression corresponding to TT is represented in two places: in the specifier of TP, and in the predicate. Since Aspect establishes a relation between ET and TT,

---

[6]The nominal structural subject must therefore be accommodated in a neighboring specifier position, made available in the dense functional field surrounding TP in the cartographic framework (e.g., Rizzi 2004, 2006; Cardinaletti 2004; Rizzi and Shlonsky 2007). I label this position Subj and abstract away from the details of subject vs. Topic Time positions in the present discussion, for reasons of space.

while Tense establishes a relation between TT and RT, a time argument corresponding to TT is introduced in the specifier of the verbal projection below T (for illustration, I use AspP).[7] Within AspP, ET (restricted by the denotation of VP) is related to TT, yielding a proposition denotation, but this proposition is not anchored. The time argument corresponding to TT is anaphoric: it has not been located with respect to the reference time. As von Stechow (1995) observes, this anaphoric character of this time expression requires it to undergo QR, like a PRO argument, leaving a variable *t* of type *i* and introducing a lambda abstractor (Heim and Kratzer 1998) that converts the AspP back into a property of times.[8]

(20)   a.   [$_{AspP}$ *t* [$_{Asp'}$ Asp [$_{VP}$ ET [$_{V'}$ ...]]]]

         b.   [$_{AspP}$ λt [$_{AspP}$ *t* [$_{Asp'}$ Asp [$_{VP}$ ET [$_{V'}$ ...]]]]]

Since AspP now denotes a property of times, it takes the referential TT projected in SpecTP as its subject. On the PRO analogy, this referential TT controls the anaphoric time variable in SpecAspP. The derivation of TP, where T is a vacuous relator (<<it><it>>), parallels the derivation of a predicational copular sentence like (19), where the copula is of type <<et><et>>.

(21)

TP: TT$_i$, T', T, AspP, λt, AspP, *t*, Asp', Asp, VP, ...ET...

---
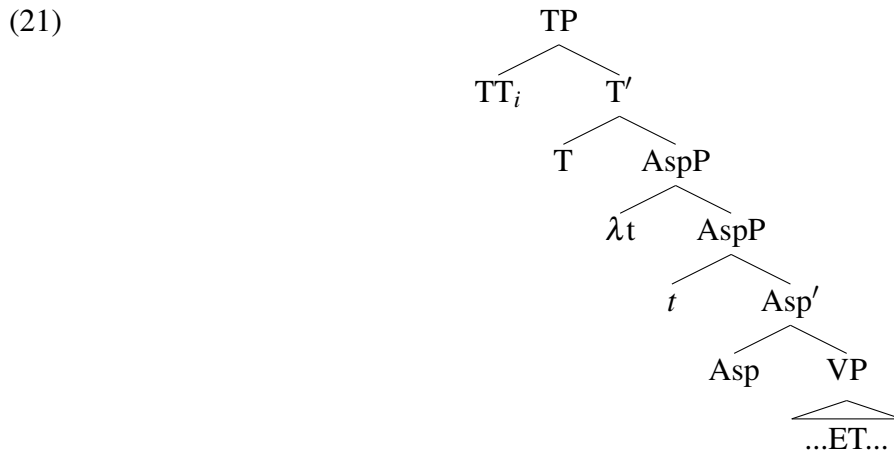
[7]Once again, I abstract away from the specifier positions occupied by A-moved nominal arguments, including the landing site of object shift.

[8]Demirdache and Uribe-Etxebarria (2014) also implement abstraction over ET, which allows TT to bind ET, giving rise to an imperfective interpretation when no aspectual predicate is present.

8.2.2.2 Anchoring TT

Demirdache and Uribe-Etxebarria (2000, et passim), henceforth D&U-E, treat the syntax of the clausal tense, aspect, and adverbial temporal modifiers as isomorphic, involving a small class of ordering predicates. The restrictor of a time can thus be either an AspP, a VP or a PP. Tense, Aspect, and Prepositions encode "spatio-temporal" relations between times.[9] For D&U-E, Tense contains a TENSE predicate that orders TT with respect to RT:

(22)

```
                        TP
                   ┌─────┴─────┐
                  ZP           T′
                   │      ┌─────┴─────┐
                  RT      T          AspP
                          │       ┌───┴────┐
                       TENSE     ZP       AspP
                                  │
                                 TT
```
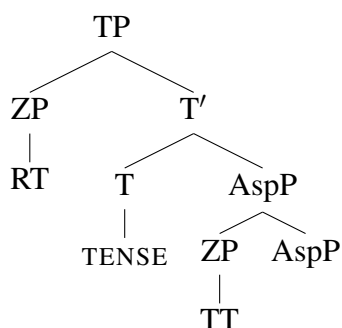
The class of temporal ordering predicates includes BEFORE, WITHIN, and AFTER. These predicates are sufficient for establishing temporal relations in the TP, the AspP, and the PP, thus allowing for parsimonious analysis of temporal modification in different syntactic environments.

(23)   BEFORE

    a.   She **left**.

        TT BEFORE RT

    b.   She **had** left.

        ET BEFORE TT

    c.   She left **before** Sunday.

        TT/ET BEFORE Sunday

---

[9]D&U-E use the term "Assertion Time", or Ast-T, for what I am calling Topic Time.

Preserving the main insights of this approach, I adapt the model slightly, in order to distinguish more clearly between semantic TENSE and morphological [tense]. My primary motivation for this is that copular amalgam sentences, where the copula is inflected for [tense], will be shown to lack TENSE. Not only do they lack temporal interpretation, but they do not pattern with environments (e.g., "fake" tense and Sequence-of-Tenses) where a [tense] form fails to associate with the corresponding TENSE interpretation. They have [tense], just like T, yet no temporal ordering predicate can be postulated.

An additional motivation for divorcing TENSE from [tense] is the fact that T itself is a functional head, a predicational Relator (den Dikken 2006), so imbuing it with the lexical properties of the prepositional predicate is undesirable. Since it behaves configurationally like the contentless head of a predication structure, it should be treated semantically as such, if possible.[10]

I propose the following syntactic structure for TENSES. T relates ZP to AspP, as in (18). ZP is a complex expression, headed by a generalized quantifier analogous to a definite determiner. This quantifier binds a phonetically empty time, which has a PP restrictive modifier locating it with respect to a reference time. The time bound by Z and the reference time with its own internal structure can be rendered overtly in expressions like *the Sunday before my birthday* (Demirdache and Uribe-Etxebarria 2007). Since ZP is definite, it is associated with a contextual uniqueness presupposition; it is not merely a weak existential quantifier. The definiteness of Z ensures that the TT Partee's *stove*-example is interpreted as a specific time. I assume that a contextually given assignment function determines which relevant interval it refers to, if it is not bound.

---

[10]Stowell's (2007) analysis of tense similarly dissociates tense form and interpretation, but like D&U-E, he situates the TENSE predicate in T.

(24)

```
                                  TP
                    _____|_____
                   ZP                            T′
             _____|_____                  _____|_____
          THE_i         Z′                 T         AspP
                   _____|_____          |        ___|___
                  Z            PP         λt      AspP
                  |        _____|_____            __|__
                  t_i    e_i        P′          ZP    Asp′
                               _____|_____       |      |
                              P          ZP    ...t    ...VP
                              |          |
                          BEFORE       PRO_RT
```
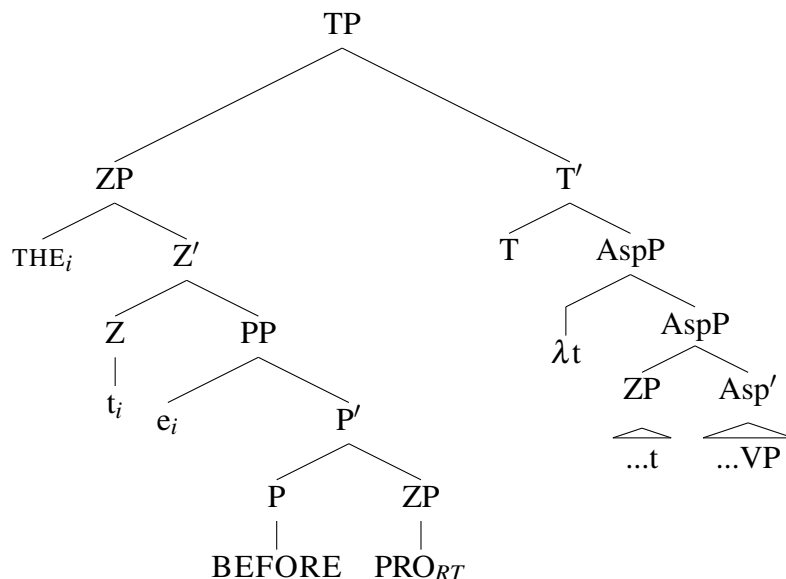
This tree can be paraphrased as:

(25)   'The contextually relevant time that is before RT has the property of being the time that

stands in an aspectual relationship with an Event Time that has the property VP.'

Using the *stove*-example:

(26)   'The contextually relevant past time lacks the property of being a time interval where an

event takes place that is a turn-off-the-stove event of which I am an agent.'

Under this analysis, TENSES, like PAST and PRESENT, are not predicates housed in T, nor are

they presuppositions associated with quantifiers over time; rather, they are predicates introduced

in a relative clause-like restriction of the time in Spec,TP. The determiner-like status of Z is com-

patible with the fact that the phonological form of the past-tense affix and the definite determiner

are similar in Germanic languages (see discussion in Chapter 7).[11] To summarize, TENSE is a

---

[11] Similarly, the relationship between ET and TT is established by way of a prepositional restrictive modifier. I have
abstracted away from the internal structure of the ZP in Spec,AspP, but it must be complex, just like its counterpart in
Spec,TP.

The *t* represented in the tree above in Spec,AspP is the structural counterpart of the anaphoric PRO_RT inside the TT
ZP—it serves as the reference time for the ET of the verb. *t* is bound by the lambda abstractor adjoined to Spec,AspP,
which ultimately allows it to be predicated of the TT in Spec,TP, whose reference is supplied by the assignment

prepositional predicate forming part of the restriction of a pronominal time expression.

In the structure I propose, RT is the complement of the TENSE predicate. RT is anaphoric, so ultimately, it raises and leaves an abstraction index, just like its counterpart in AspP. It is then bound by a context variable located in the C-layer. In independent clauses, this context variable is deictic: it refers to the actual utterance context. In finite dependent clauses, it is bound by a time in the superordinate clause—typically ET. In non-finite clauses where the C-layer is absent, there is no context variable, and so RT is bound directly by the superordinate clause. Some non-finite clause types may also have an anaphoric context variable. The reference of RT is thus established in a parallel fashion to pronominal elements (Stowell 2007), which are either deictic or bound.

The contribution of T in the clausal spine is functional, lacking substantive semantic content. It is a RELATOR, which simply serves to predicate a property of times (AspP) of a time (ZP), which is located with respect to RT. T clearly has a close relationship to substantive temporal content, because it relates a predicate of times and a time. It is thus the core of the temporal anchoring domain of the clause (see, e.g., Ritter and Wiltschko 2014). In the model sketched above, T's *specific* morphological [tense]-feature ([present] or [past]) is valued in the spec-head configuration when T merges with its TENSED external argument ZP. The featural make-up of T drives several combinatorial processes in the syntax. T is not, however, the locus of TENSE itself.

This approach takes the proposal of Demirdache and Uribe-Etxebarria (2007), who propose a homomorphism between ordering relations between times in PPs, AspPs, and TPs, to its logical extreme: the homomorphism exists not because of similarities between P and T, but because wherever there are "spatio-temporal" relations such as WITHIN, BEFORE, and AFTER, there are prepositional ordering predicates: the semantic type of these prepositions' arguments determines their specific interpretation. P can take ZP (temporal) arguments, but it can just as easily take DP

---

function.

(locative) arguments. In a language like English, T must take a ZP specifier in order to find a specific value for [tense], but the counterpart of T in a locative-anchoring language like Hakomelem Salish selects for place DPs instead (Ritter and Wiltschko 2009, 2014).

8.2.2.3 Embedded tenses

A central puzzle in tense theory deals with environments where TENSE interpretation does not match the local [tense] form. The correspondence between [tense] morphemes and TENSE meaning is superficially quite unstable, and varies across languages.[12] Consider, for example, the following [past]-under-[past] sentence in English:

(27)   Jill found out that Jack lived in Japan.

Applying the model above, where PAST is associated with the predicate BEFORE, the expected interpretation is one where Jack's living in Japan precedes Jill's finding out. The ET of the *found-out*-clause is the local binder of the RT of the *lived*-clause, so the embedded BEFORE locates the TT of *lived* prior to the ET of *found-out*. This interpretation is indeed available: in this case, (27) has a past-shifted interpretation like that of the pluperfect example in (28):

(28)   Jill found out that Jack had lived in Japan.

This interpretation is clearly not the only interpretation, however, nor indeed is it the preferred interpretation. The most natural interpretation of (27) is the *simultaneous* interpretation, where Jack's living in Japan includes the time of Jill's finding out. The embedded BEFORE seems to make no contribution to locating the embedded TT.

---

[12]Temporal interpretation in non-finite clauses makes for a particularly complex issue (see, e.g., Stowell 1982; Bianchi 2003; Landau 2004; Szabolcsi 2009; McFadden 2013), which is outside the scope of this chapter, since copular amalgam clauses are never embedded in nonfinite contexts.

There are a few options for analyzing the form-meaning mismatch associated with simultaneous [past]-under-[past]. This pattern is known as Sequence of Tenses (SOT). The embedded [past] could be a PRESENT-in-disguise, a zero-TENSE, or a deleted PAST. Referential and quantificational approaches to tense (e.g., von Stechow 2002; Kratzer 1998; Ogihara 1996) rely on binding to deliver simultaneous readings of tenses in complement clauses. Since time arguments participate in bound variable anaphora, it is expected that their formal features do not always map directly onto interpretation. As in the nominal domain, mismatches arise between the specific features of ZPs and their interpretation (Heim 1991, Heim and Kratzer 1998; von Stechow 1995, 2002; Schlenker 1999; Kratzer 2009).

8.2.2.3.1 Zero-tenses

Consider the following example, attributed to Heim (1991):

(29)   Only I did my homework.

(29) makes available an interpretation wherein *my* is a "fake" indexical (Kratzer 2009): it bears first person features, but is not interpreted as first person.

(30)   [Only I$_{[+first]}$]$_{[+first]}$ λx$_{[+first]}$ [x$_{[+first]}$ did x$_{[+first]}$'s homework]

The key insight here is that bound variables have their morphological features valued under binding: a moved expression transmits its morphological features to all of the variables it binds. Features that are transmitted in binding relations have no semantic content. They are purely morphological; hence, they make no contribution at LF. The result of this is that (30) can mean that 'I am the only person who did their own homework.'

Since times have the same structure as pronouns, we should expect bound ZPs to participate in the same kind of feature transmission/deletion under binding. Just as *my* in (30) does not have its

own first person features, a [past] ZP bound by a superordinate PAST ZP will be formally [past], but have no interpreted BEFORE-restrictor. Examples like (31) show that the analogy between pronouns and tenses holds up (von Stechow 2002).

(31)   Gerd*i* hopes[+pres] he*i* wins~~[+pres]~~.

The attitude verb *hopes* quantifies over a person, a world, and a time, and transmits its features to the corresponding bound variables in the embedded clause, so the bound time is spelled out with the morphological feature [present], but this feature is not interpreted at LF.

    Cross-linguistic variation is captured by proposing that the inventory of pronouns in a given language can be drawn from three types (Kratzer 1998):

(32)   a.   indexicals

          b.   variables

          c.   zero-pronouns

Indexical pronouns like *I, you,* and English PRESENT are interpreted with respect to the extra-linguistic context. (In the syntactically articulated model of the ZP that I propose, the indexicality of the PRESENT ZP comes from the NOW RT in its restriction.) Interpreted English PAST falls into the second category, since, like ordinary pronouns, it includes a restriction BEFORE-ZP responsible for its valued features, and its reference is determined by the assignment function (compare to *he*). The ordinary English PAST, as we saw above, is defined if it precedes some RT, and if the context maps it to a specific interval. SOT languages like English must also have a "zero"-tense: a ZP that enters the derivation without a restriction or features of its own. This pronoun is a bound anaphor, precisely the one that occurs in simultaneous [past]-under-[past] contexts.

(33)   a.   ZP-PRESENT: RT in the restriction clause is indexical

    b.   ZP-PAST: RT in the restriction clause is anaphoric

    c.   ZP-ZERO: ZP is PRO

Since English lacks a bound variable PRESENT tense, it shows the SOT pattern.[13]

(34)   a.   Petja said that Misha was crying.   simultaneous

        b.   Petja said that Misha is crying.   special reading

A non-SOT language like Russian, by contrast, has the following inventory.

(35)   a.   ZP-PRESENT: RT in the restriction clause is indexical

        b.   ZP-PAST: RT in the restriction clause is anaphoric

            ZP-PRESENT: RT in the restriction clause is anaphoric

It lacks zero-tense, so any embedded tense has a bound RT, but retains its BEFORE/WITHIN interpretation.

Since the syntactic model I proposed treats T, rather than ZP as the PF-locus of morphological [tense], I must assume that the features of T are valued by those of the ZP in its specifier, even when ZP's features are transmitted under binding. This is reminiscent of the [φ]-features on a T in an Agree relation with a bound pronoun.

8.2.2.3.2   Concord approach to SOT

An alternative approach to uninterpreted [tense] is put forth by Stowell (2007) (see also Zeijlstra 2004). As in the syntactic model sketched in the previous subsection, Stowell divorces formal tense from semantic tense interpretation. Morphological [present]/[past] tense is located on Z, while the predicates PRES/PAST occupy T. Stowell proposes that the morphological tenses are (anti-)polarity

---

[13]I defer detailed discussion of the special, double-access reading until the next subsection.

items: [past] must occur in the surface scope of PAST, while [present] must not occur in the scope

of PAST. This allows for the ambiguity between past-shifted and simultaneous readings of [past]-

under-[past] in English. The morphological configuration is licensed, because it obeys polarity, but

it can be associated with two different semantic representations: one in which the predicate on T is

PAST, which yields the shifted reading, and one in which the predicate on T is PRES, which yields

the simultaneous reading.

While this approach sacrifices one conceptually appealing parallel with the nominal domain

(feature deletion under binding), it gains another (scope-sensitive polarity items). Unfortunately, it

offers an overly simple derivation of the interpretation of [present]-under-[past], so polarity sensi-

tivity cannot be the whole story.

Consider (36):

(36)   John thought that Mary is in the room.

Since [present] cannot remain in the scope of PAST, it must raise out of the complement clause and

adjoin to the higher clause, where it is bound by UT. This yields the interpretation that Mary is in

the room at UT. Nevertheless, this sentence clearly entails something about the temporal location

of the content of John's thought. To capture this, Stowell proposes that [present] reconstructs at LF,

when the polarity requirement no longer holds. This yields a relative present interpretation, where

Mary's being in the room overlaps John's thinking. While this interpretation of [present]-under-

[past] is available, Abusch (1988) and Ogihara (1996) (among others) show that it is not entailed,

so the facts require a more nuanced semantic analysis. Polarity sensitivity alone is not sufficient.

In section 8.6, I will propose that concord is responsible for the appearance of [past] on the

amalgam copula, because it does not have a ZP in its specifier at all.

8.2.2.3.3 The double-access reading

[Present]-under-[past] in English has special properties, because of the conflict between the in-dexicality of the PRESENT and the intensional environments it occurs in. Consider the following example.

(37)   John said that Mary is in the room.   (Ogihara 1996)

The embedded clause in (36) has a "double-access" reading: it expresses a property of a TT interval that overlaps both UT and the local NOW of the subject. It is anchored to both the external and internal logophoric centers.

Unlike (38), where the embedded PRESENT is purely indexical, anchored to UT, the proposition *that Mary is in the room* does not necessarily hold at UT in (37) (see Abusch 1988, 1997; Ogihara 1996; Kratzer 1998; von Stechow 2002, among others).

(38)   [What I am eating right now] was hand-picked last week.

Rather, as Ogihara (1996) observes, some state of affairs that forms the basis for John's commit-ment to his statement must hold at his NOW, and extend to the actual NOW. Unlike in FUTURE-under-PAST, the subject is not attributed an attitude about a future time—it is *about* his present time, but it has some connection to UT. The PRESENT manages to be both indexical and bound, which is an odd state of affairs.

Ogihara observes that (37) is felicitous in a very specific, and perhaps surprising, set of circum-stances.

(39)   John said that Mary is in the room.

         a.   Mary is in the room. John says, "Mary is in the room." Mary remains in the room.

Speaker reports, "John said that Mary is in the room."

b.   Someone is in the room. John says, "Mary is in the room." Joe corrects him: "No, that's Sue." Sue remains in the room. Speaker reports, "John said that Mary is in the room."

c.   Someone is in the room. John says, "Mary is in the room." Joe corrects him: "No, that's Sue." John realizes his mistake. Sue remains in the room. Speaker reports, "John said that Mary is in the room."

d.   #Mary is in the room. John says, "Mary is in the room." Mary leaves the room. Speaker reports, "John said that Mary is in the room."

e.   #Someone is in the room. John says, "Mary is in the room." Joe corrects him: "No, that's Sue." Sue leaves the room. Speaker reports, "John said that Mary is in the room."

f.   #Nobody is in the room. John says, "Mary is in the room." Mary later enters the room. Speaker reports, "John said that Mary is in the room."

(39a) shows the simplest case, where the embedded proposition actually holds at both matrix ET and UT. This is not entailed by a PRESENT under an attitude verb, although it is available in simple factive contexts, where the proposition is true in all worlds:

(40)   John learned that two plus two is four.

In (39b), however, the state of affairs which John mistook for Mary being in the room remains at UT, but *Mary is in the room* is true at neither ET nor UT. In (39c), John's belief that Mary is in the room holds at matrix ET, but his belief state changes before UT—nevertheless, the tense of the object of the attitude can still be PRESENT. If the actual state of affairs forming the basis

of his belief changes, however, as in (39d) and (39e), the speaker can no longer report his belief with [present]-under-[past], even if the change does not affect the truth *in the actual world* of the proposition embedded under the attitude predicate. Conversely, if the property John's belief state does not hold consistently from ET to UT, even if it becomes true at UT, double-access is not available. Clearly, double-access sentences do not simply express that the embedded proposition holds at one or both times, nor do they simply express that the attitude-holder's belief holds at one or both times.

Although the details of the proposals go beyond the scope of the present work, the main idea in Abusch's (1988, 1997) and Ogihara's (1996) analyses of the double-access reading is that they express attitudes about *de re* states or times. Since the PRESENT is indexical, it must overlap the actual now. The attitude-holder is acquainted with the state or time under the description "a time that overlaps NOW".[14] The attitude-holder ascribes to that time the property expressed in the content of the attitude at the ET of the embedded clause.

Something blocks the pure indexical interpretation of PRESENT in an attitude complement. Intuitively, a person cannot be acquainted with something he knows under the description "now", and have a belief about it, it that thing has not yet come into existence. This is formalized as the "Upper Limit Constraint" (ULC) (Abusch 1988). In Schlenker's (2003) formulation:

(41)   "The time coordinate of a context variable c is an upper limit for the denotation of all time terms which are in its immediate scope, in the sense that these may not denote an interval which is entirely after the time of c."

The relevant context variable is the embedded clause's RT.

---

[14]This is analogous to *de re* beliefs about individuals; e.g., *The man in the yellow hat is my neighbor*. There is an actual individual that I am acquainted with under the description *that man in the yellow hat*, although I may very well be mistaken about his actual identity, and erroneously ascribe to him the property of being my neighbor.

Since the English PRESENT must overlap the speaker's NOW, it raises out of the embedded clause to avoid violating the ULC. It binds a temporal trace in the embedded clause, thus converting it into a property of times. The attitude-holder ascribes that property to the present time he is acquainted with. Since that time is bound by the indexical PRESENT, it must include both the attitude-holder's time and the matrix NOW. This derivation also allows for the truth of the embedded proposition and the attitude-holder's knowledge of its truth to be unconstrained, while requiring that the attitude-holder have access to some state of affairs holding of the time interval. For the example in (39), the paraphrase is:

(42)   There's a time John is acquainted with in the past, which he knows as "NOW". That time is an interval that overlaps the speaker's NOW. He ascribes to himself the belief that the interval has the property of Mary being in the room.

8.2.3 Summary

This section reviewed the primary approaches to the syntax and semantics of tense and time in generative grammar. It sketched a syntactic model of tense based on the pronominal approach to tenses. T is treated as a simple functional element relating a time and a property of times. TT, a pronominal expression of category ZP, occupies the specifier of T; the restriction of ZP that locates it in the PRESENT or PAST is situated inside a prepositional small clause modifier. English has an indexical PRESENT, a pronominal PAST, and an anaphoric zero-TENSE, while languages without SOT effects have a different inventory of pronominal tenses.

The section also illustrated some of the puzzles related to embedded tenses, in order to lay the groundwork for a discussion of temporal embedding in specificational copular sentences.

8.3 Tense in canonical specificational copular sentences

The central purpose of this chapter is to establish that amalgam specificational copular sentences are anchored without tense (i.e., the copula is not associated with a ZP), but the stage is not yet set for a close discussion of tense in amalgams. Tense in canonical specificational copular sentences remains mysterious: there is no clear baseline available for amalgams, until these canonical sentences are better understood. The present section therefore develops an analysis of canonical pseudoclefts' temporal embedding patterns. I will argue, contra Sharvit (2003), that the copula of canonical specificational sentences associates with an ordinary matrix TT. This conclusion allows for the comparison of canonical and amalgam specificational pseudoclefts, which I will show do not have ordinary matrix TT.

Specificational copular sentences, particularly pseudoclefts, present some unusual tense patterns, which have been described in detail in the foundational literature on specificational copular sentences (especially, Akmajian 1979; Higgins 1979; Declerck 1988), but which have only recently been examined from the point of view of formal semantics (Sharvit 2003; Romero 2004). The basic observations, from a morphological perspective, are that the copula may take [present] form, or it may match the [tense] form of the lexical verb.

(43)   a.   What John liked is coffee.

         b.   What John liked was coffee.

The latter observation is a curious reversal of the usual SOT dependency. The structurally superordinate element, the copula, seems to be conditioned by the structurally subordinate element, the lexical verb. Moreover, any meaning difference contributed by the different [tense] forms of the copula in (43a) and (43b) is difficult to detect, unlike in other contexts—even in other copular

clauses.

(44)   a.   The person I saw is (now) bald.

         b.   The person I saw was (then) bald.

On the semantics side, these authors observe that the TENSE of the copula is closely tied to that of the lexical verb. Unlike in ordinary embedded clause contexts, where the subordinate verb event's temporal reference can be shifted from the perspective of the superordinate verb (with its RT bound), the copula and the lexical verb in a specificational sentence cannot have disjoint temporal reference.

(45)   *[What John was a week ago] was important to himself yesterday.

Compare to:

(46)   [What John was a week ago] was important to him yesterday.

The temporal coreference between the matrix and embedded clause events of a pseudocleft is tied to the identity-like meaning established by a specificational relation. Intuitively, if the specificational sentence is asserting, at some particular time t, that an expression y with unknown reference has the denotation x, then it makes no sense to look for y at some other time t'—y doesn't refer to anything until its identity is specified at t.

This section addresses two questions. First, what is the best way to formalize this intuition? Secondly, does this formalization predict differences between canonical and amalgam specificational sentences?

The early literature on specificational sentences cited above (e.g., Akmajian 1979; Higgins 1979; Declerck 1988) gives informal descriptions of the basic observations above; these authors

catalogue some nuanced facts, but they make very light use of the tools offered by a formal temporal semantics.

Sharvit (2003) and Romero (2004) offer valuable starting points, formalizing the temporal dependency between the copula and the lexical verb in pseudoclefts. Nevertheless, I will show that their approaches lack empirical adequacy. They cannot accommodate some of the forms that speakers found acceptable in the experiments reported in Chapter 4. Moreover, they offer no clear route to characterizing cross-linguistic similarity between pseudoclefts in SOT and non-SOT languages. Relevant to the second question above is the fact that these approaches offer no insight into the relevant differences between canonical and amalgam pseudoclefts.

The analysis I give takes seriously Romero's (2004) observations about intensionality in specificational sentences, considering them in light of the fixed alignment between syntax and information structure in specificational sentences. The anchor in a specificational sentence is interpreted *de dicto*, so the specificational copula introduces an intensional environment: its reference is dependent on the world-time contributed by the copula. The anchor is also interpreted as a topic, however, which gives it access to the actual world. The anchor, as topic, sets the stage for the the predication relation contributed by the copula, because the topic individual and the Topic Time are part of the same topic situation (Klein 1994, 2014). In this presuppositional sense, the reference of the copula's TT is constrained by the TT of the anchor. This codependence following from the anchor's status as both *de dicto* and topical triggers the obligatory temporal coreference between the copula and the anchor.

Once tense form-meaning correspondences in canonical specificational copular sentences have been elucidated, the discussion in the next section will return to the chapter's primary purpose: proving that amalgam copular sentences are unique among (English) finite clauses in lacking a

syntactic T projection *and* the temporal semantic content associated with T's ZP arguments.

8.3.1 Temporal dependency in pseudoclefts

The basic paradigm to be explained is illustrated in (47).

(47)   a.   What he likes is coffee.

        b.   What he liked was coffee.

        c.   What he liked is coffee.

        d. *What he likes was coffee.

The two [tense] forms can either match, or the copula can be [present]. This paradigm differs from what is found in other subordinate clause contexts, where all combinations are allowed, (48).

(48)   a.   He says that he likes coffee.

        b.   He said that he liked coffee.

        c.   He says that he liked coffee.

        d.   He said that he likes coffee.

In non-SOT languages like Hebrew, the morphological and semantic pattern in (47) is identical, although the semantics associated with (48b) and (48d) differ from English. I will return to this issue when I present the analysis.

8.3.1.1 Morphological tense "harmony"?

Much of the literature on specificational sentences assumes that the [tense] form of the copula is dependent on the [tense] form of the lexical verb. For instance, Akmajian (1979) requires the tense of the copula to be "congruent" with the tense of the lexical verb. He treats this as a type of agreement relationship (p. 168): "It has been noted often that in pseudo-cleft sentences the tense

of the copula must agree with the tense of the verb in the clause." Higgins (1979:289–295) refers

to the pattern of congruency as "harmony".

(49)   a.   What he likes is coffee.

        b.   What he liked was coffee.

Although he notes that it is not completely obligatory, he proposes that "the tense of the copula in

a specificational pseudo-cleft sentence is not an independent variable" (p. 294).

The matching tendency is, however, just that: a tendency. The copula can always take [present]

form, even when the lexical verb does not:

(50)   a.   What he liked is coffee.

        b.   What he will like is coffee.

Clearly, the [tense] forms are not in a strict agreement relationship.

8.3.1.2   Simple form requirement

Another generalization in the literature is that the copula in pseudoclefts must take "simple form".

(Higgins 1979:290) asserts this requirement in no uncertain terms: "There are only two tenses, past

versus non-past, and the copula may only appear as *is* or *was*, regardless of what kind of verb ap-

pears in the *what*-clause." Higgins reports that in an out-of-the-blue context, only the sentence with

the simple copular form supports the specificational reading, although all are temporally *congruent*

(example from Higgins 1979:291 (50)):[15]

(51)   a.   What John has been is very rude.

        b.   *What John has been has been very rude.

---

[15]Of course, a predicational reading for (50b–c) is possible, if the expression *what John is/has been* is understood
referentially, e.g., as his profession, but *rude* is a somewhat odd predicate to ascribe to a profession, so the predicational
reading in these examples is not prominent.

    c. *What John is has been very rude.

In the same vein, Declerck (1988) proposes that the tense of the copula is often "obscured" by "tense simplification". The temporal contribution of the copula is "neutralized" in the environment of the tensed lexical verb, so it is impossible to express compound tense on both the variable expression and the copula. This is a semantic property, he argues, which is evident even in double-NP specificational sentences, where the variable expression does not bear morphological tense (Declerck 1988:82 (31)):

(52)   ??Smith had been killed, and the murderer had been John.

Since *the murderer*'s time argument is contextually restricted by the discourse establishing a *murder*-ing event in some anterior past (indicated by the past perfect), the second sentence specifying the identity of the murderer must be expressed in "simplified" or "neutralized" past.

    The empirical claim that the copula must take simple form is too strong, however. As we saw in Chapter 4, speakers do accept temporal auxiliaries on the copula, especially when the same auxiliary appears on the lexical verb in the anchor. The oddness of (52) is not the result of a syntactic or semantic ban on auxiliaries occurring with the copula, but rather, of ordinary temporal anaphora resolution.

    In unmodified out-of-the-blue contexts, past perfect forms in English are often odd, because there is no contextually salient TT interval from which to back-shift the ET.

(53)   a. #We had eaten.

         b. #Lou had written a book.

If there is a discourse antecedent or explicit restrictor of the TT, however, past perfect is felicitous.

(54) a. We had eaten when he arrived.

  b. Lou had written a book by Tuesday.

Given the nature of past perfect, there is an interpretation of (52) that poses a pragmatic problem. If the ET of *kill* establishes a contextual antecedent for the TT of *be*, such that the past perfect shifts the specification relation back from the *kill* event, then *the murderer* is undefined—at this time, the murder has not yet taken place, so the definite description *the murderer* is undefined. Since it is undefined, its denotation cannot contain the individual John.

The ET of *kill* does not obligatorily bind the TT of *be*, however, so this incoherent interpretation is not the only one available. If a later TT is provided by some other context (e.g., the TT associated with the auxiliary *had* in the first clause), then the back-shifted ET associated with *be* can coincide with the ET of *kill*, yielding a perfectly sensible proposition about the identity of the murderer.

If an appropriate past time is made salient in the discourse, then the copular clause can more easily be back-shifted. If we take Declerck's (1988) unacceptable example from (52) and modify the supporting context slightly, the past perfect form of the copula is not problematic:

(55) Context: I found Smith's body yesterday, and I suspected then that:

  The murderer had been John.

In this case, unlike in the original example, the contextually salient TT is not the ET of *kill*, but rather, the time of the body's discovery, posterior to the murder. It is fine to assert the identity of the murderer using the past perfect—identification of the murderer is back-shifted with respect to the discovery of the body.

An unbound or unmodified TT will naturally be interpreted as coreferential with an immediate discourse antecedent, just like a pronoun.

(56)   John$_i$ walked in. He$_i$ sat down.

It would be odd in (56) for 'he' to refer to someone other than John, just as it would be odd for the TT associated with the copula in 'the murderer had been John' not to take the ET of 'kill' in the preceding sentence as its antecedent. This is why the simple past form *was* is strongly preferred. A semi-formalized translation of the simple form sentence is given in (57), which abstracts away from the intensional properties of *the murderer*, for the time being.

(57)   Smith had been killed and the murderer was John.

'There is a contextually salient time t preceding UT, and some time t' before t, such that Smith was killed at t', and there is some time t'' preceding UT and included within t' such that the murderer was John at t''.'

The tense combinations of the copula with different modals will be described later (section 8.3.5.6), but let us note briefly that the copula in a specificational copular sentence need not be in simple form with respect to modals either. The following examples and judgments, due to Ross (1999), are cited in den Dikken (2005b: (120)):

(58)   a.   The one who will win will be one of us.

b.   What Sandy might have been reading {might have been/might be/has been} Tolstoy.

c.   What Sandy has been reading {might have been/?might be/has been/is/was} Tolstoy.

d.   What could have been happening {could have been/could be/was/is} that she {has been/ could have been/was/is} working at home.

The specification relation is much more amenable to modal modification than to aspectual modification, especially epistemic modal modification. It should be clear that no problem arises, in principle, when the copula is in a non-simple form.

8.3.1.3 Disjoint reference prohibition

The real restriction on [tense] and TENSE in specificational sentences is that the copula and the lexical verb's TTs cannot freely refer to different times. Even when explicit modifiers are included to restrict the TT, the dependency between the copula and the lexical verb persists.[16]

(59) a. *What John liked last year was coffee yesterday.

      b. *What John will like next year will be coffee tomorrow.

Compare these specificational sentences to other sentences with embedded clauses, where such disjoint reference is straightforwardly possible.

(60) a. I said yesterday that John liked coffee last year.

      b. I will say tomorrow that John will like coffee next year.

Disjoint temporal reference in (59) forces the implausible predicational reading of these sentences, paraphrased in (61).

(61) a. The thing that John liked last year (e.g., some flowers on a coffee plant) had the property being coffee yesterday.

      b. The thing that John will like next year will have the property of being coffee tomorrow.

The fact that the predicational reading is forced in these contexts is telling—it underscores the importance of the non-referentiality of the anchor in the specificational sentence. The lexical verb's time simply cannot be shifted from the time of the copula (Akmajian 1979; Declerck 1988; Sharvit 2003).

---

[16]Of course, the lexical verb can be back-shifted with respect to the copula with the aid of the perfect auxiliary *had*, but in this case, the TT associated with *had* corefers with the TT of the copula *was*.

The case of [past]-under-[present] would seem to constitute an exception to this generalization. Declerck (1988) observes, for example, that an absolute PRESENT on the copula is appropriate provided that the specification relation has "present relevance". Declerck (1988:83) argues that "the specificational relationship is by nature timeless", so tense on the copula has a similar meaning as tense in a statement of a universal truth, e.g., 'two plus two is four'. The PRESENT tense is therefore always available to the copula, implying that the specification relation has "present relevance" (i.e., present TT), even if the event or state denoted by the verb held in the past. Declerck's (1988) approach contrasts with earlier observations in Higgins (1979) and Akmajian (1979), which insist on a more robust morphological harmony between the copula and the lexical verb (although Higgins (1979) acknowledges that [past]-under-[present] indeed occurs).

8.3.2   Temporal dependency: the usual suspects

Tense patterns in pseudoclefts might reasonably be expected to follow from the constraints on temporal dependencies in other environments. For instance, specificational sentences might pattern with sentences with subject clauses, or sentences with relative clauses, since pseudoclefts include a free relative clause in the role of subject.

If the combinatorial possibilities of tenses in other embedded contexts run parallel to the patterns in pseudoclefts, then it is reasonable to analyze tense in the two domains using the same tools. For example, the tendency for a specificational pseudocleft with a [past] form verb to feature a [past] form copula should be treated as arising from the same source as SOT. In the discussion that follows, I will show that temporal dependency in pseudoclefts does not pattern with these other embedded tense environments.

8.3.2.1   Relative clause tense environments

Since pseudoclefts feature a (free) relative clause in subject position, we might ask whether its embedded RT is anchored to the copula's ET (as in subject and complement clauses) or to the UT (as in extensionally interpreted relative clauses). A quick look at a paradigm of [tense] combinations in specificational pseudoclefts vs. surface-similar predicational sentences makes it clear that the relative clause in the former patterns more closely with subject and complement clauses than relative clauses with respect to constraints on its embedded tense, despite the fact that in both cases, the relative clause is the structural subject.

(62)   Specificational

      a.   What John enjoys is baseball.

      b.   What John enjoyed was baseball.

      c.   What John enjoyed is baseball.

      d. *What John enjoys was baseball.

      e. ?What John enjoys will be baseball.

      f.   What John will enjoy will be baseball.

(63)   Predicational

      a.   What John enjoys is popular.

      b.   What John enjoyed was popular.

      c.   What John enjoyed is popular.

      d.   What John enjoys was popular.

      e.   What John enjoys will be popular.

      f.   What John will enjoy will be popular.

The [present]-under-[past] examples in (62d) and (63d) show the contrast clearly. (63d) means:

'The activity that John currently enjoys had the property of being popular at some past time.' The

relative clause is interpreted *de re*, and the *popular* property is predicated of it at some past time.

This interpretation is not available in the specificational sentence in (62d), however.

The interpretive difference between the two sentence types is highlighted in the following ex-

amples, with [past]-under-[past] in the predicational sentence.

(64)   a.   What John enjoyed yesterday was popular last week.

        b.   What John enjoyed last week was popular yesterday.

The PAST in the relative clause can be either anterior to or posterior to the PAST of the copula. The

relative clause's TENSE is interpreted outside the scope of the matrix copula's TENSE, while in the

specificational cases, the embedded TENSE is not free.

(65)   a.   *What John enjoyed yesterday was baseball last week.

        b.   *What John enjoyed last week was baseball yesterday.

8.3.2.2   Sequence-of-tenses: [past]-under-[past]

Since the embedded [past] in a specificational sentence cannot be posterior to the matrix PAST, it

may be like SOT environments. If it is, we would expect the specificational anchor to pattern with

complement clause tenses and subject clause tenses, which are subject to SOT.

If the anchor's TENSE behaves like other embedded TENSEs in English, then the simultaneous

interpretation of [past]-under-[past] is wholly unsurprising. Let us assume that, just as in other

contexts, it is the more embedded element (the anchor) that is dependent on the higher one (the

copula), and not the other way around, as assumed by Akmajian (1979), Higgins (1979), and

Declerck (1988).

8.3.2.2.1   The meaning of a PAST specificational copula

If the TT chosen by the speaker is in the past, the copula takes [past] form. Recall that Topic Time is "the time span to which the speaker's claim [...] is confined" (Klein 1994:4). A past TT does not entail that a state of affairs ceases to hold in the present. Klein (1994) presents the following example. In a courtroom, a judge can ask a witness to describe what he noticed when he entered a room. The witness may answer:

(66)   There was a book on the table. It was in Russian.

The TT in this case is anaphoric to the time specified by the judge in the discourse context, and the witness's response is confined to that time. Of course this does not imply that the book's being in Russian holds only in that interval; rather, the witness's assertion is only 'about' that interval.

The logic is the same when applied to specificational pseudoclefts—a person may make a claim that is restricted to a particular, salient interval, without implying that the property does not hold beyond that interval. Take the following specificational example, which, for the sake of simplicity, has a DP anchor.

(67)   Context: I went to my child's school last week to meet her teacher for the first time. I went into the classroom, and lo and behold:
        The teacher was Mary.

It is likely, given the real-world knowledge that teachers generally remain in their positions for longer than a week, that the teacher is *still* Mary at UT (one week later). Moreover, the specification *meaning* expresses a sort of lifetime property, if we construe it as analogous to the question-answer pair relation. The proposition that the answer to the question of the identity of the teacher, computed at that TT, is Mary remains true at all times, regardless of how long Mary remains in the

role of the teacher.

There is nothing anomalous, therefore, about restricting the specification assertion to a particular interval. The PAST tense of the copula in the specificational sentence in (67) therefore only serves to order the TT chosen for the assertion prior to the UT. It makes no claims about the boundaries of the interval at which Mary Smith is the teacher.

8.3.2.2.2 Embedded zero-TENSE

Let us now observe the relation between a [past] form copula and a [past] form lexical verb in pseudoclefts. If the speaker chooses a PAST for the copula, then according to the ordinary embedded tense pattern in English, the verb embedded under it should be either a zero-TENSE or a true PAST, as in (68) and (69).

(68)   They thought that she liked coffee.

(69)   She said that he WAS nice, but that he had gotten rather unpleasant lately.

In pseudoclefts with stative verbs, only the simultaneous interpretation is possible, as we saw above.

(70)   a.   What he liked was coffee.

        b.   What she needed was more money.

        c.   What he preferred was classical music.

In these examples, the PAST TT argument of the copula determines the reference of the embedded verb. For instance, (70a) means: 'There is a time t that is included in a contextually relevant interval t' preceding UT, such that the value of the unique entity x such that there is a time t'' included in t where he likes x at t'' is coffee at t.' This formalization assumes that the embedded simultaneous

TT is a zero-TENSE (i.e., t'' is not restricted by a local PAST tense—it is bound by the superordinate

TT t). In the framework described in the previous section, this means that ZP associated with the

embedded T is like a controlled PRO.

If the parallel between pseudoclefts and sentences with complement and subject clauses holds,

then [past] eventive verbs under a [past] copula should have a zero-TT and an anterior ET; that is,

a relative, rather than simultaneous past reading should be possible.

(71)   a.   What he drew was a circle.

        b.   A circle was what he drew.

A paraphrase is: 'It was the case at some contextually relevant time t prior to UT that the x such

that he drew it at some time t' prior to t had the value *a circle* at t.' Although this simplifies the

properties of perfective aspect somewhat, it captures the intuition that the identification of the

variable as *a circle* is posterior to the drawing event, just as in canonical sentences with embedded

eventive verbs, the embedded event is interpreted as anterior to the matrix event.

That this reading is specificational rather than predicational is highlighted by the interaction

between the aspectual properties of the verb and the referential status of the *value*. For example,

compare (71) with (72).

(72)   a.   What he ate was an apple.

        b.   An apple was what he ate.

In the case of *a circle*, the denotation of the value comes into being when the event associated with

the accomplishment verb *draw* is complete, while in the case of *an apple*, the value's denotation

ceases to exist in its *apple*-form once the accomplishment event of *eat*-ing is complete. Neverthe-

less, the same zero-TT reading is natural: in order for *an apple* to be identified at t as the x *such*

that John ate x, the eating event must have already taken place as some time prior to *t*.

8.3.2.2.3   Hebrew [present]-under-[past]

The counterpart of the simultaneous zero-TENSE form in English is the simultaneous PRESENT

in Hebrew. Hebrew is a non-SOT language, so embedded tenses are interpreted relative to the

local evaluation context; that is, the embedding predicate's ET. In (73a), the [present]-form *ohevet*

'loves' is interpreted as simultaneous to the matrix verb AMAR 'said'. The *loving*-time is therefore

interpreted as past from the perspective of the speaker.[17]

(73)   Dani amar   Se   Miriam ohevet   oto.
       Dani say.M.PAST COMP Miriam love.FEM.PRES him
       'Dani said that Miriam loved (lit. loves) him.'   simultaneous

    Curiously, in specificational pseudoclefts, a [present] form embedded in the anchor fails to

receive this relative PRESENT interpretation:

(74)   *ma Se   Miriam koret   haya   Hamlet.
        what COMP Miriam read.FEM.PRES be.PAST Hamlet
        'What Miriam is reading was Hamlet.'

If tense patterns in pseudoclefts could be reduced to SOT phenomena, the ungrammaticality of

(74) would be wholly unexpected.

8.3.2.2.4   Shifted PAST

By treating specificational pseudoclefts as no different from other sentences with embedded tenses,

we obtain a straightforward account of the "harmonious" [past]-under-[past] pattern in English, but

the unavailability of a simultaneous [present]-under-[past] in Hebrew casts doubt on this approach.

An additional problem arises when we consider the absence of the shifted reading with statives.

---

[17]I am grateful to Maayan Barkan, Itamar Kastner, Nadav Sabar, and Yael Ziv for help with the Hebrew data and judgments.

Recall that an English stative under a PAST can have an anterior interpretation when it is contrastively focused or contextually disambiguated, (69). If pseudoclefts patterned with SOT, then we should find the same shifted reading available. This prediction is not borne out, however. As we saw in section 8.3.1.3, the lexical verb cannot be interpreted as anterior to the PAST copula. (59) is reproduced in (75).

(75)   a.   *What John liked last year was coffee yesterday.

        b.   *What John will like next year will be coffee tomorrow.

The absence of the shifted reading is even more striking, if we again consider the case of Hebrew. In Hebrew, [past]-under-[past] by default triggers a shifted reading. Since Hebrew lacks a zero-TENSE, [past] is always interpreted as a true PAST with respect to the local RT (the superordinate verb's ET) (Kratzer 1998; Ogihara and Sharvit 2012). Compare the embedded [present], repeated from above, to the embedded [past].

(76)   a.   Dani amar   Se   Miriam ohevet   oto.
            Dani say.M.PAST COMP Miriam love.FEM.PRES him
            'Dani said that Miriam loved him.'   simultaneous

        b.   Dani amar   Se   Miriam ahava   oto.
            Dani say.M.PAST COMP Miriam love.FEM.PAST him
            'Dani said that Miriam had loved him.'   past-shifted

If temporal embedding in specificational sentences were like these complement clause tenses, then we would expect the Hebrew [past]-under-[past] to yield a past-shifted interpretation, like (76b). Once again, this interpretation is unavailable. Like English, [past]-under-[past] in a Hebrew pseudocleft must receive a simultaneous interpretation.

(77)   a.   *ma   Se   Miriam kar'a   lifney shavua haya   Hamlet etmol.
            what COMP Miriam read.FEM.PAST before week   be.PAST Hamlet yesterday
            'What Miriam read last week was Hamlet yesterday.'

b.   ma Se   Miriam kar'a   haya   Hamlet.
     what COMP Miriam read.FEM.PAST be.PAST Hamlet
     'What Miriam read was Hamlet.'

Since Hebrew lacks a zero-TENSE, what is the source of the simultaneous reading?

8.3.2.3   Intensional environments and temporal *de re*

Not all embedded TENSES are interpreted in-situ. Ogihara (1996), Abusch (1997), and others propose that an embedded TENSE (a ZP in the framework sketched in the preceding section) may be interpreted *de re* if it scopes out of its embedded context and adjoins to the matrix clause. The result is that the embedding attitude verb relates that temporal object (a "*res*-tense"), a property of times (the TT-less complement proposition), and an individual (the attitude-holder), via an acquaintance relation between the attitude-holder and the *res* (see section 8.2.2.3).

(78)   John believed that Mary was smart.

If the embedded PAST is interpreted *de re*, then *believe* expresses a relation between John, the property [Mary be smart], and a particular past time that John is acquainted with under a normal description like "the time before my now". John ascribes to himself the belief that the past time he is acquainted with is a time where 'Mary is smart' is true. This kind of temporal interpretation accounts for partially independent readings of TENSES embedded in intensionally opaque environments

The following illustrates the predictions of temporal *de re* for pseudocleft tenses; ultimately, it is shown that *de re*, like sequence-of-tenses, fails to account for specificational pseudoclefts.

8.3.2.3.1   Res-PAST

Ogihara and Sharvit (2012) observe that some Hebrew speakers, particularly younger ones, allow a simultaneous interpretation of a [past]-under-[past]. (79), for these speakers, allows the simulta-

neous reading that is most salient in its English counterpart.

(79) Dani amar   Se   Miriam ahava   oto.
      Dani say.M.PAST COMP Miriam love.FEM.PAST him
      'Dani said that Miriam loved him.'   simultaneous

They analyze this as an instance of temporal *de re* (Ogihara 1996; Abusch 1997). The embed-

ded PAST scopes out of the embedded clause. The meaning of this sentence has the semi-formal

translation in (80).

(80)   There is a past time *t*, which Dani is acquainted with under some description, and he

        commits himself to the truth of the proposition that *t* has the property Miriam-loving-him

        by asserting p at a past time t'.

Notice that the *de re* PAST interpretation does not establish an intrinsic ordering between *t* and

t'. Dani can certainly assert that a property holds of a time prior to his local NOW. The Upper Limit

Constraint (Abusch 1988, 1997; Ogihara 1996) prohibits the embedded PAST from being inter-

preted as posterior to the matrix PAST, however. TENSES embedded under attitude verbs cannot

be interpreted as posterior to the attitude-holder's NOW. The subject cannot be acquainted with the

temporal object he has a belief about under two contradictory descriptions: that it is a time in his

past, and that it is a time in his future.

Temporal *de re* is also behind the interpretation of embedded PRESENT in English. Unlike He-

brew, English does not have a relative PRESENT—its PRESENT is always interpreted with respect

to the utterance context. English [present]-under-[past] therefore receives the double-access read-

ing, as we saw in the previous section, where a speaker's attitude is toward an present time interval

with the description "NOW", and which he is acquainted with as his local NOW:

(81)   Dani said that Miriam loves him.

Given the availability of a *de re* interpretation of PAST in Hebrew, there is a plausible source for

the simultaneous [past]-under-[past] that emerges in pseudoclefts. The problem is far from solved,

however. We still do not have an explanation for why the usual back-shifted reading is unavailable.

Moreover, if *de re* PAST is available, then why is *de re* PRESENT unavailable in *both* languages?

8.3.2.3.2   Res-PRESENT

Recall the ungrammatical [present]-under-[past] in pseudoclefts, reproduced below.

(82)   *What Miriam is reading was Hamlet.

(83)   *ma   Se   Miriam koret   haya   Hamlet.
       what COMP Miriam read.FEM.PRES be.PAST Hamlet
       '*What Miriam is reading was Hamlet.'

The same pattern can be observed in specificational copular sentences that lack embedded

tensed clauses, so the ungrammaticality of PRESENT-under-PAST is clearly semantic in nature.

(84)   Specificational

       a.   The (current) principal is Joe.

       b.   The (former) principal was Joe.

       c.   The (former) principal is Joe.

       d.   *The (current) principal was Joe.

The PRESENT-under-PAST configuration forces a predicational interpretation, which is very odd,

since a name or other rigid designator is not easily interpreted as a predicate that can hold or not

hold at different times. (84d) must receive the rather bizarre interpretation that there is a person

who is the current principal and who formerly had the property of being Joe.

The unavailability of a double-access (or any) PRESENT-under-PAST interpretation in specificational copular sentences highlights the fact that it does not pattern with other embedded tenses.

To develop an analysis, let us examine tenses in relative clauses interpreted *de dicto*. The presence of the NPI in the following example ensures that the relative clause must be interpreted in situ:

(85)   I didn't find a single linguist who {will / would} chair any sessions. (Demirdache and
         Uribe-Etxebarria 2007:359)

Let us assume, following Abusch (1988), that *will* is the morphological realization of the FUTURE-shifting modal WOLL in combination with PRESENT, and that *would* is its counterpart under PAST. Since the form *would* in (85) indicates that WOLL is in the scope of a PAST, so the embedded verb in (85) must be anchored to the superordinate tense rather than to UT. Double-access (*will*) is similarly blocked in (85), because the relative clause must also remain in the scope of the NPI's licensor.

If the relative clause must be interpreted inside the opaque context created by the intensional verb, then its tense cannot be independently anchored; that is, a relative clause in an intensional context must have its tense bound. As Romero (2004) argues, if the "subject" position of a specificational copular sentence is an intensional context, then absolute PRESENT-under-PAST should be impossible. If this analysis is correct, then PRESENT-under-PAST fails in specificational copular sentences because the embedded (relative) clause (or the time argument of the NP) must be interpreted in situ. What remains to be explained is why the double-access reading is unavailable: there is no NPI in specificational sentences to block *res*-movement of an embedded PRESENT. I return to this issue shortly.

8.3.3 Unusual suspects

I have exhausted the ordinary sources of embedded simultaneous readings in English and Hebrew; the tense of pseudoclefts cannot be predicted by zero-TENSE (SOT) or by temporal *de re*. A large piece of the puzzle is the fact that the anchor is interpreted *de dicto*, which prevents an absolute PRESENT interpretation, but there must be more to the story to block shifted and double-access readings.

8.3.3.1 Atemporal copula: Sharvit (2003)

Sharvit (2003) argues that the specificational copula is atemporal. She assumes a framework where tenses are restricted existential quantifiers over times, indexed to a reference time and presupposed to be included within a contextually salient interval, and where the copula or any verb would ordinarily contribute an ET argument bound by that quantifier. The specificational copula contributes a tense quantifier, but lacks a time argument. Its tense therefore binds the TT of its subject, the anchor. This approach offers an account of the simultaneity requirement—the copula's tense must be the same as that of the lexical verb.

8.3.3.1.1 [Past]-under-[past]

A "harmonious" [past]-under-[past] has the following meaning, in Sharvit's model.[18]

(86)   What John was was a fool.

PAST-t_0-t_6 1 [[what 2 PAST-t_0-t_1 3 John be$_{PRED}$-t_3 e_2] be$_{IDENT}$ [a fool]]

"There is a t preceding now, such that $t \subseteq t_6$ and the unique property P such that there is a t' preceding now such that $t' \subseteq t$ and John has P at t', is the same as the property of being a fool." (Sharvit 2003:377)

---

[18]Sharvit (1999) analyzes specificational copular sentences as identity statements between properties. This aspect of her formalization is included in the cited examples here, but nothing hinges on it.

The restricted PAST quantifier associated with the copula binds the TT restriction of the PAST tense associated with the predicational copula in the anchor. The copula has no ET. This gives rise to the simultaneous interpretation.

The case here contrasts with ordinary PAST-under-PAST, which can support a shifted reading.

(87)   I heard that Mary was hungry (and that later she ate her fill).

Here, the matrix ET binds the embedded RT: clearly, *heard* is not atemporal.

8.3.3.1.2   [Past]-under-[present]

Sharvit (2003) presents PAST-under-PRESENT as a marked configuration in pseudoclefts, despite the fact that [past]-under-[present] is unmarked in other sentence types.

(88)   I think he liked coffee.

She proposes that a PAST-under-PRESENT pseudocleft like (89) requires a "special context".

(89)   What John was is a fool.   (Sharvit 2003:378 (7c))

Sharvit compares sentences like (89) to double-access complement clauses, pursuing the old intuition that it is the matrix tensed element that is sensitive to the embedded one, rather than the other way around.

Although Sharvit describes the reading in (89) informally as similar to the double-access reading, she states (p. 378): "although John doesn't need to be a fool at the current time, the current time must be made relevant for the sentence to be felicitous." Given this description, the comparison with double-access is unwarranted, as the sentence in (89) simply means that the specificational relation instantiated in the copula has present TT—"present relevance", in Declerck's (1988) terms. She builds this "special requirement" into the semantics of (89), which has the following semantic

representation:

(90)   PRES-t_7 1 [[what 2 PAST-t_0-t_1 3 John be$_{\text{PRED}}$-t_3 e_2] be$_{\text{IDENT}}$ [a fool]]

"There is a *t* overlapping now, such that $t \subseteq t_7$ and the unique property P such that there

is a t' preceding now such that $t' \subseteq t$ and John has P at t', is the same as the property of

being a fool." (Sharvit 2003:378)

Given this translation, the analogy with the double-access reading is much too strong, as Sharvit

(2003:378) herself acknowledges. John does not need to be a fool at UT in (90). UT is included in

the matrix TT, but nothing is asserted to hold at matrix TT: the matrix quantifier does not bind any

time argument in its own clause. This current time is made relevant in that it restricts the matrix

quantifier and contains the past interval of the embedded clause's TT, but no property is ascribed

to it. This state of affairs is much weaker than the true double-access reading, where something is

asserted to hold of a state that spans both at an embedded past time and UT. Perhaps PAST-under-

PRESENT in pseudoclefts does not require such a "special" context after all.

8.3.3.1.3   [Present]-under-[past]

For Sharvit (2003), the ungrammaticality of PRESENT-under-PAST in specificational copular sen-

tences is due to an incoherence in the truth conditions associated with binding into the subject (the

anchor). In this approach, the sentence in (91) must have the following meaning:

(91)   *What John is was a fool.

PAST-t_0-t_7 1 [[what 2 PRES-t_1 3 John be$_{\text{PRED}}$-t_3 e_2] be$_{\text{IDENT}}$ [a fool]]

"There is a *t* preceding now, such that $t \subseteq t_7$ and the unique property P such that there is

a t' overlapping now such that $t' \subseteq t$ and John has P at t', is the same as the property of

being a fool." (Sharvit 2003:377)

The time bound by the matrix PAST quantifier must both be included within a time preceding UT/now (past TT) and include a time overlapping UT/now. The time at which John has P, t', cannot both overlap UT and be included within an interval entirely prior to UT. Since these conditions cannot be met, PRESENT-under-PAST is not possible.

8.3.3.2 Ambiguous copula: Romero (2004)

Romero (2004) picks up where Sharvit leaves off, addressing the question of whether the anchor in a specificational sentence must be interpreted *de dicto*. Her proposal is that the specificational copula is not atemporal; rather, its time argument binds the TT of the lexical verb. This analysis yields roughly the same set of simultaneity requirements as Sharvit's, but they follow from the intensionality of the specificational copula's subject position.

Romero (2004) examines a subtle ambiguity in intensional contexts in order to demonstrate that world-dependence entails temporal dependence. She proposes that the anchor is always an intensional object—it is either a concealed question or a concealed question intension—and the specificational copula is an intensional verb (see also discussion in Chapter 4). An ambiguity arises due to the fact that the copula can satisfy its selectional requirement either with the extension (92a) or the intension (92b) of its subject, the anchor. In the intensional case, the anchor denotes a question-intension, a meta-question, and the value must be an intensional object as well: a question extension.[19] The example below is taken from Romero (2004:282–283):

(92)   Context: A group of 2-year-old girls from the Ukraine was given in adoption to several

---

[19]Romero (2004) uses bound-variable connectivity in the following examples to prove that they are truly specificational, and not predicational, but this test is not as reliable as she claims, as bound variable connectivity has been observed in predicational contexts as well (Heycock and Kroch 1999; Sharvit 1999, etc.).

(i)   What every student got was a nuisance to him.
       'For every student x, what x got was a nuisance to x.'   (Sharvit 1999:303 (20))

families in Barcelona. The director of the adoption program encouraged the biological relatives of each girl to keep in touch with her by writing letters, telling them though that they should not identify themselves using their name or family relationship. After a couple of years, the girls have developed some hypotheses on who every secret writer may or may not be. For example, no girl thinks that the one who writes to her the least can possibly be her mother. In fact, they are all right, since, for every girl, the one who writes to her the least is her uncle.

a. The anonymous writer that no girl thinks can possibly be her mother is (in fact) her uncle.   question-extension

b. The anonymous writer that no girl thinks can possibly be her mother is the one who writes to her the least.   question-intension

When the copula takes the intension of its subject, it binds the embedded TT; when the copula takes the extension of its subject, the embedded TT is free. The temporal contrast she argues for is illustrated in (93) (Romero 2004:287 (73)–(74)).

(93) a. The price that Fred thinks is $1.29 today was $1.79 yesterday.

Reading A: NP's extension, [present]-under-[past] possible

b. *The price that Fred thinks is $1.29 today was the price of milk.

Reading B: NP's intension, [present]-under-[past] impossible

This pattern is congruent with Abusch's (1997) analysis of the interaction between TENSE and *de re/de dicto* reference. If the NP cannot scope out (because it takes narrow scope with respect to the intensional verb), then neither can the TENSE of the relative clause inside the NP.[20] The *de dicto*

---

[20]Von Stechow (2002) also proposes that intensional verbs can bind into their complements, deriving similar temporal embedding patterns.

reading, Reading B, has bound tense, while the *de re* reading, Reading A, has independent tense.

Romero's (2004) analysis improves upon that of Sharvit (2003) in that it assumes the copula does have a time argument, but that it binds into its subject (requiring tense "harmony") when it behaves as an intensional verb. This correctly predicts that the TT of the specificational copula should freely reflect the contextually salient time chosen by the speaker, as presented in the previous subsection.

8.3.4 Some problems with TT-binding approaches

Here, I briefly point out some problems with Sharvit and Romero's approaches. I will then sketch an alternative.

8.3.4.1 Over-generation problems

8.3.4.1.1 Extension of an intensional anchor is temporally dependent

Romero (2004) analysis is that only the intensional environment created by the copula in Reading B triggers the temporal dependency. This means that PRESENT-under-PAST should be grammatical with Reading A: the reading where the anchor is a concealed question (or in my view, a predicate-intension). This over-generates:

(94)   *What I like today was yoga yesterday.

If the weight is interpreted *de re*, as a (concealed) question extension, its tense is still not independent.

Consider the examples in (95) against the backdrop of Romero's *price* example in (93).

(95)   a. *What John enjoys reading was *The Minimalist Program*.

        b. *John's current adviser was Professor Jones.

The use of names in the post-copular *value* position makes it difficult to construe these examples as predicational or as Reading B-type intensional specificational copular sentences. They are instances of Reading A, the more generic specificational reading; nevertheless, contra Romero (2004), PRESENT-under-PAST is impossible.

Romero's (2004) example of temporal independence in Reading A (93) is misleading. The value *$1.79* is in fact understood as a predicate, and so the example Romero uses to demonstrate temporal independence in Reading A is not specificational at all. Contrasting sentences are shown in (96) and (97).

(96)   Predicational

      a.   The price that Fred thinks is $1.29 today was $1.79 yesterday.

      b.   The menu item that Fred thinks is the special today was the special yesterday.

          Predicate: ...had the property of being P yesterday

(97)   Specificational

      a.   *What I like today was yoga yesterday.

      b.   *The person who answers the phones today was John yesterday.

It is clear after all that truly specificational copular sentences with Reading A do not allow for the superscriptional term's TENSE to be independently anchored, so the temporal dependency in pseudoclefts cannot be traced to a selectional quirk of the copula.

8.3.4.1.2   Reversibility

Another problematic component of Sharvit's and Romero's proposals is that they make specific reference to the structural subject position of the copula. Specificational copular sentences are of course reversible: they can have a linear order where the anchor is in its canonical predicate posi-

tion. In this order, the tense patterns are the same: PRESENT-under-PAST yields ungrammaticality. Witness (98) and (99).

(98)   a.   Joe is the (current) principal.

       b.   Joe was the (former) principal.

       c.   Joe is the (former) principal.

       d.   Joe was the (*current) principal.

(99)   a.   Important to himself is what John is.

       b.   Important to himself was what John was.

       c.   Important to himself is what John was.

       d.   *Important to himself was what John is.

If the copula's tense is binding into its *subject*, then there should be no problem with shifted readings. (98b), for example, would mean that there is a past stage of Joe that had the property of being the former principal; (99b) similarly would mean that there is a past stage of the property *important to himself* that previously had the property of being what John was. Nevertheless, we have seen ample evidence that such shifted readings are not available. The simultaneity requirement in pseudoclefts needs to be recast.

8.3.4.2   Under-generation problems

Now consider where these TT-binding approaches under-generate. First of all, Sharvit's atemporal analysis of the specificational copula does not match the intuitions I described above (see also Declerck 1988) that the copula's TT is freely chosen by the speaker. Next, it fails to accommodate the facts associated with the FUTURE example paradigms and non-SOT languages.

8.3.4.2.1 The future

The most neutral way of shifting the evaluation time into the future in a pseudocleft is to embed the FUTURE modal in the anchor, while keeping the copula in the simple [present], as in (100).

(100)   What John will like is coffee.

Sharvit's binding analysis of a sentence like (100) is given in (101).

(101)   What John will be is a fool.

PRES-t_6 1 [[what 2 PRES-t1 woll 3 John be$_{PRED}$-t_3 e$_2$] be$_{IDENT}$ [a fool]]

"There is a t overlapping now, such that $t \subseteq t_6$ and the unique property P such that there is a t' overlapping now, $t' \subseteq t$, such that there is a t'' after t', and John has P at t'', is the same as the property of being a fool." (Sharvit 2003:385)

The interpretation works out to be the same in Romero's (2004) intensional version, although the implementation is different. Clearly, (101) is predicted to be coherent: as long as the matrix TT, of which nothing is asserted to hold, overlaps both UT and the embedded future ET where John has property P, the sentence is acceptable. The matrix TT interval overlaps the present time, but nothing must necessarily hold at present time. This requirement is what Declerck (1988) calls "present relevance", although even in Sharvit's (2003) formalized proposal, it remains vague.

FUTURE-under-PRESENT is not the only way to convey a future meaning, however. The following are also acceptable, particularly when a supporting context is given.

(102)   Context: I have tried to anticipate the preferences of all my guests. Based on what I know about John's tastes, here is what refreshments I expect him to enjoy at my party.

    a.   What John will like will be coffee.

b.   What John likes will be the shrimp cocktail.

It is often claimed that these combinations are ungrammatical (Akmajian 1979; Higgins 1979; Declerck 1988; Sharvit 2003; Romero 2004). In Chapter 4, however, I presented evidence that speakers accept FUTURE-under-FUTURE and, to a lesser extent, PRESENT-under-FUTURE.

Both Sharvit's (2003) and Romero's (2004) binding analyses of the simultaneous evaluation time in pseudoclefts erroneously predict that the sentences in (102) must be incoherent. For example, in Sharvit's formalization of a sentence like (102), the matrix WOLL operator introduces an interval after UT, but since the copula contributes no ET, the future-shifting modal binds the present TT of the embedded clause, which must overlap UT. The embedded ET, whether PRESENT or future-shifted by a local WOLL, is thus dependent on a TT interval that is in two places at once. Since the requirement is contradictory, PRESENT or FUTURE-under-FUTURE is impossible. Similarly, in Romero's formalization, which does not explicitly discuss the future cases, the TT of the anchor must be both overlapping absolute present and bound by an interval that is entirely after absolute present (the copula's forward-shifted TT).

Additional examples of acceptable specificational sentences with the FUTURE copula are given below, where the use of the *self*-anaphor and the non-intersective reading of the two spider varieties mentioned ensures that these sentences really are specificational:

(103)   Context: John often worries about his wife. He has started drinking excessively to cope
          with his feelings. His friend Joe reports that if John keeps this behavior up...
          The person John worries about will be himself.

(104)   Context: Mary enjoys drinking a glass of wine with dinner whenever she goes out. She
          hopes to become pregnant soon, and her habit will have to change. Now whenever she

goes out...

What she drinks will be water.

(105)   Context: Lou plans to visit the spider exhibit at the American Museum of Natural History.

His friend explains what to expect:

What you see will be a tarantula and a black widow.

When there is a context that facilitates forward-shifting of the specification relation, *will* on the copula in a canonical specificational pseudocleft is acceptable (see also Ross 1999, 2000, who is similarly liberal with combinations involving the future). If the data are handled with care, therefore, it becomes apparent that the tenseless analysis of the copula developed in Sharvit (2003) is untenable.

8.3.4.2.2   Embedded pseudoclefts

If the canonical pseudocleft copula associates with its own TT, then when the entire pseudocleft is embedded, it should pattern with other statives. It should show a preference for SOT (zero-TENSE), but allow a bound, past-shifted PAST as well, just like the predicational example in (106). There is nothing that prevents ZP-PAST from being embedded in English.

(106)   Ordinary embedded stative

a.   I thought that John's house was blue.

thought-ET λt_{ZP-0} [blue(john's house) at t_{ZP-0}]

'I thought at some time t that John's house had the property blue at t.'

b.   I thought that John's house was blue.

thought-ET λt_{ZP-0} [[$_{ZP}$ TT BEFORE t_{ZP-0}] blue(john's house) at ZP]

'I thought at some time t that at some contextually defined time t' prior to t, John's

house had the property blue at t'.'

The prediction is borne out. It is possible for an embedded pseudocleft to be interpreted as past-shifted with respect to the embedding predicate.

(107)   Canonical pseudocleft

    a.   I thought that what John was was intelligent.

       'I thought at some time t that the unique property P such that John has P at t is the property of being intelligent at t.'

    b.   I thought that what John was was intelligent.

       'I thought at some time t that at some contextually defined time t' prior to t, the unique property P that John has P at t' is the property of being intelligent at t'.'

While the first interpretation is the default, the second is available as well, especially if it is supported by a context:

(108)   I had a belief about John that he was slowing down. I thought that what John was was intelligent, but that he had lost his edge lately.

It becomes easier to override the preference for the zero-TENSE by contrastively stressing the copula. The contrastive stress opens a domain of focus alternatives. In the case of the copula, since it has no lexical semantic content, contrastive stress is consistent with only two types of focus alternative sets (Rooth 1985): {true, false} (verum focus) and {zero-TENSE, bound PAST, indexical PRESENT, etc.} (see also Klein 1998).[21]

(109)   I thought that what John was [WAS]$_F$ intelligent.

---

[21]This is a simplification of verum focus, which is not always simply polarity focus (Lohnstein 2015). It suffices for the present purposes to identify environments where verum focus associates with the finite element, but tense focus does not.

    a.   Alternative set 1: {I thought that what John was **was** intelligent, I thought that what John was **was not** intelligent}

    b.   Alternative set 2: {I thought that what John was **was** intelligent **at ZP-0**, I thought that what John was **was** intelligent **at ZP-PAST**, I thought that what John was **is** intelligent **at ZP-PRESENT**, etc.}

Again, the past-shifted interpretation is not the default, but it is available.

If the empirical claim in (109) is accurate, then we have additional evidence that the canonical specificational copula is tensed. While Sharvit (2003) would have to attribute any past-shifted interpretation of an embedded pseudocleft to the most deeply embedded verb, the fact that contrastive stress on the copula makes temporal focus alternatives available casts doubt on this idea. It would have to be the case that only the most deeply embedded verb could have its tense focused:

(110)   I thought that what John WAS was intelligent.

This option is available, of course, since the most deeply embedded verb has is own TT, but it is not the only option. The availability of an interpreted, focused PAST TT associated with the specificational copula is another strike against Sharvit's atemporal analysis.

8.3.4.3 Non-SOT languages

Sharvit (2003) observes the absence of shifted readings in Hebrew PAST-under-PAST pseudoclefts. She takes this to be evidence for her binding approach to the TENSE of the lexical verb: if the copula lacks an ET argument, then the embedded PAST is not a zero-PAST, which does not otherwise occur in Hebrew, but rather, a PAST with a bound TT.

What does not receive an explanation is the ungrammaticality of the relative PRESENT-under-PAST. Since the PRESENT has no special indexical status in Hebrew, as we saw above, but is readily

interpreted as a simultaneous tense, the TT-binding approaches fail to account for the Hebrew facts.

8.3.5 Topic Time coreference

To capture the full force of the simultaneity requirement in pseudoclefts, we need to ensure that the TTs of the lexical verb and the copula corefer. The analysis needs to incorporate the *de dicto* interpretation of the anchor, the unavailability of zero-TENSES, and the unavailability of temporal *de re*. I propose here a route toward such an analysis, which exploits the mapping between syntax and information structure in specificational sentences.

8.3.5.1 The Topic situation

A hallmark of the specificational copular sentence is that its predicate is interpreted as a topic, whether it remains in-situ or raises to subject position via predicate inversion. Klein (2014) claims, "Topic entity and topic time go hand in hand; they are two components of the entire topic situation." This observation can be strengthened. Topic entity and the topic time corefer, unless one of them is deictic.

The topic entity is presupposed. It participates in setting the stage for the assertion. The topic tells us what the sentence is about—what entity, world, time, and location. The Topic entity contains a time argument, which participates in this stage-setting function. Taking this claim further, I hypothesize that the contextually determined matrix clause TT takes the Topic entity's TT as its antecedent by default. In this way, temporal anaphora between nominal(ized) and verbal elements associated with the same topic situation parallels anaphora between nominal elements, (111).

(111)   As for John$_i$, he$_i$'s great.

By default, pronominal elements in this configuration will corefer. This includes the TT pronominal within the topic, as illustrated in (112) and (113).

(112)   Context: Let me tell you about Einstein.

      a.   #Einstein is a great physicist.

      b.   Einstein was a great physicist.

(113)   Context: I'm trying to think of an example of a great physicist.

      a.   Einstein is a great physicist.

      b.   Einstein was a great physicist.

Because Einstein is the topic in (112), and Einstein has no present stages (he is dead), the matrix verb's PRESENT TT is unexpected. In (113), however, where *a great physicist* is the topic, the present is felicitous. The matrix verbal TT takes the topic entity's TT as its antecedent.

    This proposal makes several predictions. First, the TENSE patterns should be the same whether or not the specificational sentence is reversed. Secondly, SOT and non-SOT languages should show the same behavior. Thirdly, to the extent that specificational sentences accommodate the relevant tests, the covaluation relationship between the embedded and matrix TTs in pseudoclefts patterns with coreference, rather than with binding (Reinhart 1983). Finally, the same basic pattern (with some caveats) obtains in other instances of topicalization, like subject clauses and topicalized attitudes. I will show how TT-coreference, in conjunction with the crucial observation that the specificational anchor is an intensional object, predicts these properties.

8.3.5.2   Deriving simultaneity in pseudoclefts

8.3.5.2.1   Reversibility

As we have seen, the linear order of a specificational sentence does not affect its temporal interpretation: since the anchor is always a topic, its TT must always corefer with that of the copula:

(114)   a.   *Yesterday, what John liked last week was coffee.

b. *Yesterday, coffee was what John liked last week.

Rather than stating an ad hoc binding rule that applies at a different point in the composition in different sentences (what Romero would have to propose), I propose that the covaluation rule is sensitive to the syntax-information structure interface.

8.3.5.2.2 PAST-under-PAST

The simultaneous version of (114) is grammatical. In ordinary embedded contexts, the matrix ET binds the embedded RT or the embedded TT is a zero-TENSE. When the embedded TENSE is associated with a topic entity, however, I propose that the relation between them is one of pragmatically conditioned coreference, rather than binding. This relates to the second prediction from above, namely, that non-SOT languages should behave the same way. In Hebrew, which lacks a zero-TENSE, the two PAST TTs are pragmatically required to corefer.

(115)   ma   Se   Miriam kar'a   haya Hamlet.
          what COMP   Miriam read.FEM.PST   be.PST Hamlet
          'What Miriam read was *Hamlet*.'

8.3.5.2.3 PAST-under-PRESENT

The case of PAST-under-PRESENT requires an addendum to the coreference requirement.

(116)   What John liked is coffee.

If the copula's TT must corefer with that of *liked*, then the PRESENT should not be possible.

I suggest an additional component of the conference requirement, based on an analogy with nominal topic coreference. Deictic elements can escape the coreference requirement, referring to an element of the utterance context rather than an element in the topic situation.

(117)   a.   In John$_i$'s opinion, he$_i$ is smart.

    b.   In John$_i$'s opinion, he$_j$ is smart.

English PRESENT, as we have seen, has the special property of being indexical: it must be interpreted with respect to the actual utterance time. The present copula is thus always allowed in pseudoclefts, as observed by Declerck (1988). The "special reading" associated with [past]-under-[present] arises from the fact that there are two deictic centers, just as in (117b), one contributed by the contextually determined TT, and one contributed by the indexical TT.

    Corroborating this amendment is the fact that Hebrew, whose PRESENT is not strictly indexical, requires the PAST copula in the environment of a PAST anchor. (Note that *hu* is the pronominal copula, which occurs in PRESENT contexts.)

(118)   a.   ma Se   Miriam koret   hu   Hamlet.
                  what COMP Miriam read.FEM.PRES be.PRES Hamlet
                  'What Miriam is reading is Hamlet.'

        b.   ma Se   Miriam kar'a   *hu/haya   Hamlet.
                  what COMP Miriam read.FEM.PRES *be.PRES/be.PAST Hamlet
                  'What Miriam was reading is Hamlet.'

Similarly, PRESENT-under-PAST in Hebrew is correctly ruled out, because the two TTs do not corefer. A binding approach to pseudocleft tense covaluation should predict that Hebrew PRESENT-under-PAST has its usual simultaneous interpretation. I return to the ungrammatical PRESENT-under-PAST in English pseudoclefts shortly.

8.3.5.2.4   FUTURE-under-FUTURE

The future-patterns that could not be accommodated by Sharvit's and Romero's models can be captured with coreference between TTs. Normally, a FUTURE under a FUTURE receives a shifted reading. Future shifting is contributed by a modal WOLL in the scope of a PRESENT.

(119)   She will believe that he will smoke.

The shifted reading is incoherent in specificational sentences, as Sharvit and Romero show.

A simultaneous reading with coreferential TTs is unproblematic, however.

(120)   What she will read will be *Hamlet*.

The paraphrase is: at some contextually supplied time t after NOW, the unique member of the set of things that she reads at that time is *Hamlet* at that time.

Likewise, in Hebrew, which has a true FUTURE tense rather than a modal future.

(121)   ma Se   Miriam yikra   yihye Hamlet.
          what COMP Miriam read.FEM.FUT be.FUT Hamlet
          'What Miriam will read will be *Hamlet*.'

In fact, it is precisely this simultaneous reading that is salient in sentences with sentential subjects, which are interpreted as topics. Consider the interpretation of (122):[22]

(122)   That Mary will live in France will surprise John.

The most natural interpretation is one of simultaneity: John's surprise overlaps with Mary's living in France. This sentence does not mean that at a future time t a state of affairs that will obtain at a time t' posterior to t will surprise John at t. Contrast this with a paraphrase where the embedded clause is not in a topic position, and where the shifted reading is obligatory, as predicted.

(123)   It will surprise John to learn that Mary will live in France.

Here, John's surprise is about a future state of affairs.

The TT coreference analysis thus predicts simultaneous readings in both English and Hebrew in all three tenses.

---

[22]Note that this example involves embedding as the subject of a factive predicate, so the TENSES are predicted to be interpreted *de re*, unlike in specificational sentences.

8.3.5.3 Blocking PRESENT-under-PAST

If, as has been argued in the foregoing discussion, disjoint reference between the copula's TT and that of the superscriptional term is allowed when one of the TTs is indexical, then we should expect the PRESENT-under-PAST combination to be licit in English. In actuality, the indexical PRESENT is ungrammatical. The reason for this is tied to the intensionality of the specificational clause environment, which so far this analysis has not exploited.

Ordinarily, the ET of the intensional verb binds the RT of the embedded proposition, but since the matrix TT and embedded TT must corefer, this canonical relationship would set up a clash, in e.g., FUTURE-under-FUTURE.

(124)   What you'll want will be coffee.

The copula's ET, subsequent to a PRESENT TT for WOLL would bind the RT of the embedded WOLL, which shifts *want* into its future—but the TT coreference condition requires the two WOLLs' TT to corefer. This is nonsensical. Instead, I assume that matrix RT binds embedded RT.

Binding of RT is familiar from ordinary embedded clauses with shifted readings: superordinate ET binds embedded RT. Where RT binding in the pseudocleft differs from these ordinary clauses is that it is superordinate RT, rather than superordinate ET, that binds RT. As long as the embedded RT is bound, the intensional expression can be evaluated with respect to a time and world, so there is no problem from the perspective of the intensional object. D&U-E propose that this RT-RT binding is the default embedded tense relation. It occurs in PRESENT-under-PRESENT in English.

In other embedded contexts, e.g., PAST-under-PAST, RT-RT binding is insufficient. It fails to establish an intrinsic ordering between the matrix and embedded ETs. D&U-E propose that when no intrinsic ordering is established, the embedded RT must be re-anchored to the superordinate

ET, yielding a shifted reading. In PAST-under-PAST, therefore, RT-RT binding does not occur. According to the present proposal, however, pseudoclefts should not require this re-anchoring: the coreference requirement suffices to satisfy the intrinsic ordering constraint. Because of the TT coreference requirement, the default RT-RT binding configuration remains unaltered.

8.3.5.3.1   Opacity

Recall that the superficially similar predicational clause, where the subject free relative is interpreted *de re*, allows the tense of the verb inside the free relative to be totally independent of that of the copula. Predicational sentences behave differently from specificational sentences in this regard because the TT that sets the stage for the copula's TT is not that of the lexical verb inside the relative clause, but rather, the time argument of the (null) DP head itself. A relative clause in a transparent context can freely scope out and have its tense interpreted *de re*.

(125)　　a.　　What John likes is gross.

　　　　　b.　　What John liked was gross.

　　　　　c.　　What John liked is gross.

　　　　　d.　　What John likes was gross.

(125d) introduces a time associated with the *thing*—at some past stage, it had the property of being gross, and in the present, it has the property of being liked by John.

　　　The relative clause associated with a specificational anchor cannot scope out, because it is in an intensional context. The lexical verb inside the anchor thus serves as a proxy for the covert TT argument of relative's null head. There is no special coreference condition just for specificational sentences—it is a condition on topics. If the topic happens to be in an opaque environment and to contain a lexical verb, the lexical verb will wear the coreference requirement on its sleeve.

8.3.5.3.2   No double-access

In other embedded contexts in English, PRESENT-under-PAST is perfectly licit. For example, factive subject clauses allow it, despite the coreference condition.

(126)   That Mary lives in France surprised John.

(126) can be analogized to the relative clause cases above. Since the sentential subject in (126) is interpreted *de re*, the tense embedded inside it has access to utterance time.

More importantly, complement clauses in intensional contexts allow PRESENT-under-PAST, which yields the double-access reading.

(127)   John believed that Mary lives in France.

Why should the double-access reading be blocked in specificational contexts?

Consider what the double-access reading means. Recall from the discussion of double-access readings in section 8.2 that a sentence like (127) does not entail that Mary actually lives in France, or that she ever did—simply that there be a state to which John ascribes the property of Mary living in France, and with which he is acquainted under a present-time-oriented description. The moved PRESENT is interpreted relative to John, and relative to the utterance context.

Two crucial ingredients of the double access-reading are the bearer of the belief and the matrix speaker (Giorgi 2010). In specificational copular sentences, these two ingredients are not distinguished. Unlike a propositional attitude predicate, the copula is a functional element. It does not relate attitudes, temporal objects, and attitude-holders, so the PRESENT cannot be interpreted its embedded position. There is only one "access"-point in copular sentences: that of the speaker. The opacity of the copula requires that the embedded clause be interpreted in its scope, however, which sets of up an irreparable clash between the PAST intensional predicate and the PRESENT embedded

clause.

Let us apply this to our now-familiar specificational pseudocleft example:

(128)   *What John likes was coffee.

Here, the copula relates three simple pieces: a world-time, an object of type X, and an intensional predicate of type $<s<X,t>>$. There is no attitude-holder in the argument structure of the copula to serve as the subject of the acquaintance relation and to allow the PRESENT to be interpreted in its base position. Because it can only refer to the absolute PRESENT, it triggers an automatic ULC violation.

The double-access reading is complex in attitude-ascription contexts, but it is much more straightforward in factive environments, where the embedded clause must have both the simultaneous/bound TENSE interpretation and the absolute PRESENT tense interpretation. Let us examine why a factive analysis of the double-access reading in pseudoclefts fails.

The bound component of the reading is:

(129)   There is a past time *t* at which coffee was a member of things that John liked at *t*.

The absolute component of the reading is:

(130)   There is a past time *t* at which coffee was a member of the set of things that John likes at
          UT.

These informal paraphrases make the incoherence clear: it is the intensionality of the anchor that is the culprit. It is not an object with a history, such that we can check, for a given interval, whether or not it contains coffee. Intensional objects have no history—evaluation *time* is subordinate to evaluation *world*. As Romero (2004:287) puts it: "a *de dicto* NP is necessarily temporally dependent,

though a temporally dependent NP can be understood *de dicto* or *de re*." Relevant here is the first part: since the superscriptional term of a pseudocleft (regardless of whether it receives Reading A or Reading B) is an intensional object, it is temporally opaque: it cannot directly access UT.

The absolute (and hence the embedded PRESENT) part of the double-access reading is available to an intensional object under attitude predicates, but not under intensional predicates that do not relate belief-worlds to the actual world via an attitude-holder. The lexical semantics of the copula is simply not rich enough to support the reading.

8.3.5.3.3 PRESENT-under-FUTURE

What about PRESENT-under-FUTURE? The PRESENT in English is indexical, so if it occurs under FUTURE, it triggers the double-access reading, which we have already seen is unavailable in specificational sentences:

(131)   What you see will be a man and a woman.

     a. Simultaneous: The value of the plural entity x such that you see x at time t will be a man and a woman at time t, ordered after UT.

     b. *Indexical: The value of the plural entity x such that you see x at time t overlapping UT will be a man and a woman at time t', ordered after t.

Nevertheless, (131) is grammatical for many speakers.

An explanation is available if we exploit the fact that the English (and other Germanic) future consists of the combination of TENSE with a future-shifting modal operator WOLL. The [present] form under *will* is in fact the familiar zero-TENSE. Recall that the zero-TENSE agrees morphologically with the superordinate [tense]. When bound by *will* (PRESENT + WOLL), the zero-TENSE receives [present]-features. The zero-TENSE is a bound variable, so it does not manage to intro-

duce an antecedent for the matrix verb's TT. The parallel case in the nominal domain is illustrated in (132), where the bound variable in the topic does not constrain the reference of the referential expression coindexed with it.

(132)   As for their own$_i$ mothers, those boys$_i$ would never insult them.

The absence of the coreferential pronoun may be behind the decreased acceptability of [present]-under-[future] pseudoclefts for many English speakers.

This account of [present]-under-[past] accurately predicts that Hebrew, which lacks zero-TENSE and has a genuine FUTURE, rather than only a modal future, should require the FUTURE copula when the lexical verb of a pseudocleft is FUTURE.

(133)   a.   ma   Se   Miriam yikra   yihye   Hamlet.
               what COMP Miriam read.FEM.FUT be.FUT Hamlet
               'What Miriam will read will be *Hamlet*.'

         b.   *ma   Se   Miriam koret   yihye   Hamlet.
               what COMP Miriam read.FEM.PRES be.FUT Hamlet
               'What Miriam is reading will be *Hamlet*.'

Since the future is not a modal in Hebrew, it does not participate in feature deletion/transmission under binding. The competing requirements of RT binding and TT coreference render (133b) incoherent. The embedded TT is required to overlap utterance time (by RT binding) and to follow it (by TT coreference: TT corefers with the matrix future TT).

A related problem is posed by Hebrew [future]-under-[present], which is perfectly acceptable in English, but degraded in Hebrew. Since Hebrew's PRESENT is not indexical, it does not escape the TT coreference condition, so (134) runs into the same contradiction as (133b). The copula's PRESENT RT binds the present RT of the FUTURE read, but the copula's TT, which overlaps the utterance context, must also corefer with the TT of read, which follows the utterance context.

(134) *ma Se Miriam yikra hu Hamlet.
      what COMP Miriam read.FEM.FUT be.PRES Hamlet
      'What Miriam will read is *Hamlet*.'

8.3.5.4 Coreference vs. binding

In the nominal domain, there are a number of tests to distinguish pragmatically conditioned coreference from semantic binding. Many of these tests can be brought to bear on temporal covaluation patterns as well. For example, temporal pronouns in VP-ellipsis sites support strict and sloppy identity readings, just like their nominal counterparts (Reinhart 1983). A temporal pronoun embedded within a VP-ellipsis site can be interpreted either as coreferential with its antecedent, which is bound by the event time c-commanding the antecedent, or as bound by its local superordinate event time. (135) shows the strict/sloppy identity contrast in the nominal domain, and (136) shows it in the temporal domain.

(135)   a.   John$_i$ loves his$_i$ wife, and Joe does <love his wife> too.

        (i)   Strict: Joe loves John's wife.

        (ii)   Sloppy: Joe loves Joe's own wife.

(136)   John thought that Mary was pretty, and he still does <think that Mary BE pretty>.

    a.   Strict: John thought at time t that Mary was pretty at time t, and he still thinks at utterance time that Mary was pretty at time t.

    b.   Sloppy: John thought at time t that Mary was pretty at time t, and he still thinks at utterance time that Mary is pretty at utterance time.

This test can be applied to specificational pseudoclefts, but specificational pseudoclefts must be handled with care when it comes to VP-ellipsis. Value-initial pseudoclefts fare better in elliptical contexts, so the test will rely on the value-initial order, with focus prosody disambiguating the

structure in favor of the specificational reading.[23] The predictions of the test are as follows. If the embedded TT in specificational sentences must be a bound variable (Sharvit 2003; Romero 2004), then an elliptical anchor's TT receives a sloppy interpretation: it must be bound by its local copula's TT, rather than coreferential with its antecedent. On the other hand, if the embedded TT is a coreferential pronoun, which I have argued here, strict identity under ellipsis should be required. This prediction is borne out.

The pair of examples in (137) illustrates that the unelided counterpart of the test sentences is grammatical. I attribute the second clause (which will contain the ellipsis site in the test examples) to a second speaker, to avoid the infelicity arising from violating the exhaustivity implicature of the specificational sentence. With a tense difference between the two clauses, a single speaker suffices, since exhaustivity is not violated, (138).

(137)   A: COFFEE$_F$ is what Mary likes.

         B: BEER$_F$ is what Mary likes too!

(138)   COFFEE$_F$ is what Mary likes, and BEER$_F$ was what she liked.

(139) is included to demonstrate that the baseline ellipsis example is grammatical. In the baseline example, since the remnant and the antecedent share the same tense, there is no way to tell which one is responsible for the TT's value—strict and sloppy identity cannot be distinguished. (140), where the remnant and antecedent tenses differ, shows that the elliptical tense receives a strict identification: it is a conferential pronoun, not a bound variable (whether the example is attributed to a single speaker or two).

(139)   a.   A: COFFEE$_F$ is what Mary likes.

---

[23]Connectivity effects could also be used, but such examples are difficult to process in combination with ellipsis.

        B: BEER$_F$ is <what Mary LIKE> too.

    b.   A: COFFEE$_F$ was what Mary liked.

        B: BEER$_F$ was <what Mary LIKE> too.

(140)   a.   *COFFEE$_F$ is what Mary likes, and BEER$_F$ was <what Mary LIKE> too.

    b.   A: COFFEE$_F$ is what Mary likes.

        B: *BEER$_F$ was <what Mary LIKE> too.

If a sloppy reading were available for (140), it would be grammatical, because the elliptical TT would be a PAST (or zero-TENSE) locally bound by the PAST copula; [past]-under-[past] is grammatical. Since the elliptical TT is not bound, but rather, coreferential with its antecedent, it is a PRESENT, and therefore it is ungrammatical embedded under the local PAST. This brief subsection has marshaled a well established diagnostic from the domain of nominal anaphora to show that Topic Time anaphora in specificational sentences is due to coreference, rather than binding.

8.3.5.5 Topicalized propositional attitudes

Consider briefly some corroborating evidence from other topicalization environments. This analysis of TT coreference predicts that when an attitude complement is topicalized, the reading introduced by TT coreference will become salient, in both English and Hebrew. While in-situ attitude complements in English and Hebrew show very different tense patterns, some of these differences are leveled when the clauses are fronted. Like the pseudocleft's embedded clause, the fronted attitude clause shows mixed opaque and transparent properties. Nevertheless, the parallel is not perfect. Since complement clause topicalization is an A′-relation, while predicate inversion in pseudoclefts is an A-relation, the former allows additional readings under reconstruction that are not available in pseudoclefts.

This topicalization structure requires a supporting context, and is a very infrequent construction, so speakers find it rather difficult to judge. To the extent that I could verify the responses, the prediction is somewhat borne out.

(141)   John will believe that Mary will be pregnant.   shifted reading obligatory

(142)   That Mary will be pregnant, even John will believe.   simultaneous reading preferred

In Hebrew, TT coreference predicts that speakers who require a shifted reading for in-situ PAST-under-PAST will prefer a simultaneous reading when the complement is topicalized. Most speakers I consulted accept the simultaneous reading of complement PAST, so this prediction could not be completely verified; however, these speakers preferred the simultaneous reading when the clause was topicalized.

(143)   Afilu Dani gila   Se   Miriam ahava   oto.
        even Dani find.out.M.PAST COMP Miriam love.FEM.PAST him
        'Even Dani found out that Miriam loved him.'   shifted reading preferred

(144)   et   ze Se   Miriam (adain) ahava   oto, afilu Dani gila.
        DOM this COMP Miriam (still)   love.FEM.PAST him, even Dani find.out.M.PAST
        'That Miriam (still) loved him, even Dani found out.'   simultaneous reading preferred

This analysis also predicts that PRESENT-under-PAST in Hebrew will have a salient double-access when the PRESENT is topicalized. The topicalized TT is interpreted with respect to the utterance time in the actual world, while in its base position, it is in an opaque context, where its relative PRESENT is bound. This yields the double-access reading, in contrast to the in-situ PRESENT-under-PAST, where only the simultaneous reading is salient.

(145)   Lifney mataim   shana, afilu Dani gila   Se   Miriam ohevet
        before two.hundred years, even Dani find.out.M.PAST COMP Miriam love.FEM.PRES

> oto
> him
> 'Two hundred years ago, Dani found out that Miriam loved him.' simultaneous reading

(146) #Lifney mataim   shana, et   ze Se   Miriam ohevet   oto, afilu Dani
       before two.hundred years, DOM this COMP Miriam love.FEM.PRES him, even Dani
       gila
       find.out.M.PAST
       'Two hundred years ago, Dani found out that Miriam loves him.' double-access reading

       preferred

Because the complement clause in (146) is topicalized, it must receive both a relative PRESENT

interpretation and an absolute PRESENT interpretation. Because of the normal lifespan of humans,

the double-access reading triggered by the structure in (146) yields an implausible sentence.

This corroborates the findings for pseudoclefts in Hebrew, where PRESENT-under-PAST cannot

be salvaged, even though in-situ PRESENT-under-PAST simply receives a simultaneous interpre-

tation. If topicalization of a PRESENT generated in the scope of a PAST forces the double-access

reading, and the double-access reading fails in pseudoclefts because of the absence of an attitude

holder, then the similarity between English and Hebrew pseudoclefts follows.

8.3.5.6 Modals and special cases

The tense "harmony" behavior of modals in specificational copular sentences presents a challenge,

since they do not show the same morphological harmony pattern as lexical verbs. The temporal

properties of English modals are complex, as their morphological [tense] does not always clearly

correlate with their semantic TENSE (see, e.g., von Stechow 1995; Abusch 1997; Condoravdi 2002;

Demirdache and Uribe-Etxebarria 2007).

English modals include the following formally [present]/[past] pairs (*shall* is archaic, for many

speakers).

(147) a. shall / should

       b. can / could

       c. will / would

       d. must / must

       e. may / might

       f. have to / had to

The fact that the second member of each pair is [past] can be illustrated by their unmarked occurrence in embedded SOT contexts.

(148) a. I think I shall leave.

       b. I thought I should leave.

(149) a. I think I can sing.

       b. I thought I could sing.

(150) a. I think I will sing.

       b. I thought I would sing.

(151) a. I think I have to sing.

       b. I thought I had to sing.

(152) a. I think I may sing.

       b. I thought I might sing.

The SOT-respecting form of these modals gives rise to the usual simultaneous/bound TENSE interpretation. On the other hand, if the embedded modal takes [present] form, it forces the double-access reading. Consider, for example, (153):

(153) a. I thought I must sing.

    b. I thought I can sing.

In (153a), the deontic requirement holds both at *thought* time and at UT, and in (153b), the possibility requirement holds both at *thought* time and at UT.

 Now consider modals in specificational pseudoclefts.

(154) a. What he can eat is fish.

    b. What he could eat is fish.

    c. *What he can eat was fish.

    d. What he could eat was fish.

This example demonstrates the expected pattern: (154c), on the specificational reading, could only be interpreted with the double-access reading, but as we have seen, the double-access reading in a specificational sentence is unavailable.

 The special temporal properties of modals become apparent upon considering (154b), which is ambiguous. The fact that *could* can occur easily under both PRESENT and PAST rests on an ambiguity in the temporal orientation of *could*. Higgins (1979:291) observes that [present] form *is* occurs with *could* on its conditional reading, while *was* occurs with *could* on its past abilitative reading: (154b) can mean either, 'It was the case that what he was able to eat at some past time was fish,' or 'It is the case now that what he could hypothetically eat is fish.'

 In Condoravdi's (2002) decompositional approach to English modals (see also Demirdache and Uribe-Etxebarria 2007, 2014), there are two relevant temporal parameters that group modals: temporal perspective and temporal orientation. Temporal perspective gives the time where the possibility or necessity is evaluated, while temporal orientation relates to the time where the relevant

state of affairs may obtain. The former is an open interval extending forward in time. *Could* is a "modal for the present"—its temporal perspective is always the local evaluation time—but it can take either a future orientation or a past orientation. Therefore, since PAST-under-PRESENT is grammatical in pseudoclefts, *could* under *is* is ambiguous, while *could* under *was* can only have the past orientation reading. The past-perspective *conditional* reading can only be given by the modal + perfect: *could have*.

Compare (154), contrasting the *can/could* pair with *shall/should* in (155).

(155) a.   What he shall eat is fish.

     b.   What he should eat is fish.

     c.   *What he shall eat was fish.

     d.   *What he should eat was fish.

Here, the pattern is different. We might predict *should* under *was* to be grammatical like ordinary zero-TENSE, as in (148b), but instead, it gives rise to the familiar incoherent double-access reading, just like *can* under *was* in (154c). Higgins (1979:291) observes that "the 'past tense' forms of modal verbs may behave as present tenses with respect to this tense harmony rule." That is, certain formally [past] tense modals must occur with the formally [present] copula in pseudoclefts; the [past] copula must combine with a modal + perfective *have*. Because of the delicacy of the data, Higgins does not provide a formal treatment of the pattern; his descriptive conclusion is that "the tenses that the copula must harmonize with are closer to 'semantic' tenses" (p. 294).

Two routes to an analysis present themselves. Either modals interact differently with the copula in specificational sentences than they do elsewhere (and differently than lexical verbs do), or *should*'s temporal properties differ from those of *could* (and lexical verbs). Assuming that the latter explanation will be more fruitful, let us take another look at (148b), repeated below:

(156)   I thought I should sing.

If *should* is a present-perspective modal and does not shift evaluation time (unlike *could*), then it should not be able to support a past-shifted interpretation on its own, unlike other PAST tenses under PAST. This prediction is borne out:

(157)   *I thought yesterday that I should sing last week.

It is, however, compatible with UT-oriented adverbial modification:

(158)   I thought yesterday that I should sing right now.

The decompositional approach to modal time offers insight into the difference between *should* (along with other non-shifting, present-perspective modals like *might* and *must*) and *could*. Modal TT encodes temporal perspective, introduced by an open-ended MODAL-TIME interval. The opening bound of MOD-T can participate in temporal anaphora—including the TT coreference requirement—but since MOD-T is open-ended, it can always extend into the present (Condoravdi 2002).

In embedded contexts, where a matrix time binds the opening bound of the present-perspective modal's MOD-T (in a pseudocleft, the copula's RT is the binder), a simultaneous interpretation arises, but this configuration also supports an interpretation where MOD-T overlaps UT, since it is open-ended. Thus, the simultaneous (SOT) interpretation of *should*-under-PAST also includes something like double-access. Provided the MOD-T interval is not closed off before UT by a perfective auxiliary, it will extend past UT. This double-access-like reading is wholly unproblematic in ordinary embedded contexts, but in the intensional context of the superscriptional term of a pseudocleft, it runs into the same problem as its lexical verb counterparts: the opacity of the superscriptional term in the specificational copular sentence blocks access to UT.

8.3.5.7 *Used to*

An analysis along these lines finds support from the fact that by adding to the copula of specifica-
tion, the reading of the predicate emerges where its history is "accessible": the reading is supposed
by *used to*:

(159)   What John likes used to be coffee.

A possible objection, raised by Higgins (1979), is to argue that (159) is simply a predicational
"pseudocleft", with the paraphrase: 'There is a thing that John likes and it used to have the prop-
erty of being coffee.' If this is right, then *what John likes* is interpreted as an entity, and *coffee* as
a predicate. We can overcome this objection by observing examples with *used to* that are unam-
biguously specificational (although note that *used to* can only combine with a state, so not all of
the disambiguating examples are available).

(160)   a.   John's girlfriend used to be Mary.   (Higgins 1979:225 (74))

          (i)   'It used to be the case that John's girlfriend was Mary.'

          (ii)   *'There is a person who is John's girlfriend, and she used to have the property
              of being Mary.'

    b.   The one who does most of the work used to be John.   (Declerck 1988:85)

          (i)   'It used to be the case that the one who did most of the work was John.'

          (ii)   *'There is a person who does most of the work, and he used to have the prop-
              erty of being John.'

    c.   The person everyone$_i$ loves used to his$_i$ mother.

          (i)   'It used to be the case that everyone loved his mother.'

          (ii)   *'There is a person everyone loves, and this person used to be his mother.'

Further evidence that these are specificational, rather than predicational, is that the precopular term must antecede a property anaphor *it*, rather than an individual anaphor with gender features—if the gendered anaphor is used, only the rather odd predicational "name-change" or "substance-change" reading is available (see also Mikkelsen 2004b, 2005 on this diagnostic):

(161)   a.   John's girlfriend used to be Mary, but now it/she is Sue.

           b.   The one who does most of the work used to be John, but now it/he's Bill.

           c.   The person everyone loves used to be his mother, but now it/she's his wife.

The modal *used to* contributes a long TT interval, which "makes room for" a correspondingly long stage of its subject, thereby giving access to its extension (e.g., {x: John likes x}) evaluated at different times. Since a large enough stage of the superscriptional term is available, encompassing both UT and the copula's past ET, PRESENT-under-PAST becomes acceptable. The PAST part of the reading is the ordinary PAST-under-PAST bound reading, and, since the PRESENT clause is not an assertion but a free relative, the PRESENT reading simply consists of an existential presupposition associated with the subject (the predicative expression), computed at a topic interval large enough to include UT.

(162)   What John likes used to be coffee.

           a.   'There is an interval t preceding UT and included in t' (TT), and there is an interval t'' included in t, such that at t, coffee is the unique member of the set of things that John likes at t''. AND

           b.   There is a t' overlapping UT and overlapping t, such that the set of things that John likes is not empty at t'.'

The conflict between the embedded clause's *de dicto* interpretation and multiple evaluation times is thus resolved under *used to*.

8.3.6   Conclusion: Canonical pseudoclefts are tensed

This concludes the analysis of tense combinations in canonical pseudoclefts. After reviewing previous approaches to TENSE in pseudoclefts, I provided evidence that the canonical copula does indeed have a TT argument. Since the anchor in the pseudocleft is interpreted *de dicto*, it is temporally dependent. Like other subordinate intensional clauses, its RT is bound. Assuming D&U-E's constraint that anchoring to the deictic RT is the default option, RT is bound to matrix RT.

Since this binding configuration does not entail the covaluation relationship between the copula's TT and the anchor's TT, the analysis includes an additional component. This component is based on the fixed information structure of the specificational pseudocleft. The topic entity's TT pronominal introduces a local contextual antecedent for the matrix verb's TT.

Since specificational pseudoclefts are associated with a fixed information structure, where the anchor is always a topic, its TT is responsible for establishing the copula's TT. Matrix RT binds the anchor's RT, while embedded TT corefers with matrix TT, so the two times must be interpreted as simultaneous, whether they are PAST or FUTURE. I demonstrated with data from Hebrew that this analysis has the advantage of accounting for the simultaneity requirement in non-SOT languages, which show the same pattern as English. Moreover, the anchor's TT argument receives a strict identity under ellipsis, indicating that it is not bound by the local copula's TT.

Independent evidence for topic time coreference was drawn from topicalized attitude clauses in both English and Hebrew. These environments show similar patterns of interpretation, where the formally embedded TT provides a center of deixis for the whole proposition. Topicalized attitude clauses show some more flexibility than embedded tenses in pseudoclefts, which follows from the

fact that the former occupy an A′ position and undergo reconstruction, while the latter do not.

This chapter now turns to a comparison of canonical and amalgam pseudoclefts. Although the two sentence types have similar morphological [tense] patterns, the copula of canonical specificational sentences is TENSED, while the copula of amalgams is not.

8.4   Temporal properties of canonical vs. amalgam pseudoclefts

Now that I have given an analysis of canonical pseudoclefts, establishing a baseline for comparison, the discussion now returns to its original task: to characterize anchoring without Tense in copular amalgam clauses.

At the beginning of this chapter, independent anchoring was presented as the core property of clauses that are *finite*. Languages like English typically anchor their clauses, including canonical specificational copular sentences, through TENSE. The rest of this chapter is dedicated to arguing that although amalgam copular clauses are *finite*, they are not TENSED. The copula has morphological [tense], but since it does not occupy T, it associates with no ZP time arguments.

I order to make this point, a confound must be addressed. Amalgam copular clauses exhibit the same morphological [tense] patterns as their canonical counterparts, so at a superficial level, there seems to be no difference between them. This similarity suggests the null hypothesis that their TENSE properties are also the same. While the morphological patterns in the two sentence types are similar, their interpretive patterns are different. The canonical specificational copula takes an ordinary matrix ZP in its specifier, anchored to the utterance context, but the amalgam specificational copula does not.

First, I describe the [tense] patterns that are available in amalgam specificational pseudoclefts. In section 8.4.2, I describe contexts where the semantic TENSE differences between different sen-

tence types can be tested. Section 8.4.3 then provides experimental evidence in support of the claim that the canonical copula has a time argument, while the amalgam copula does not.

8.4.1   Morphological tense patterns in canonicals vs. amalgams

Just as in canonical pseudoclefts, only three of the four logically possible simple [tense] combinations are available in simple amalgam pseudoclefts:

(163)   a.   What John likes is John likes coffee.

        b.   What John liked was John liked coffee.

        c.   What John liked is John liked coffee.

        d.   *What John likes was John likes coffee.

(164)   a.   What John likes is coffee.

        b.   What John liked was coffee.

        c.   What John liked is coffee.

        d.   *What John likes was coffee.

The same two generalizations hold: the copula may always take the simple [present] form, and the copula may take the same [tense] form as the lexical verb. [Present]-under-[past] is ungrammatical. I will argue, however, that the tense patterns have different semantic consequences, and so the ungrammaticality of (164d) and (163d) has a different source.

In section 8.3.5.6, I showed that canonical pseudocleft embedded tenses behave differently from other embedded tense environments with respect to, for instance, modal *should*. Modal *should* can be embedded under a PAST:

(165)   I thought that you should leave.

The modal time—the interval where the obligation is evaluated—has its opening point bound by the superordinate verb, locating it in the past, simultaneous to the time of thinking. The interval extends into the actual present. It has an indexical component, unless a perfective auxiliary *have* closes the interval prior to UT. In pseudoclefts, bare *should* under a PAST copula is ungrammatical:

(166)   *What he should try was coffee.

I showed that this is because the construal of *should*-under-PAST requires access to the actual NOW: it has a double-access component. Since the opaque environment of a specificational sentence makes the double-access reading unavailable, bare *should* cannot occur under *was* in a pseudocleft.

Once again, the pattern is the same for the amalgam pseudocleft.

(167)   a.   *What he should try was he should try coffee.

         b.   *He should try coffee was what he should try.

Given that the morphological forms are the same, it may at first seem unwarranted to provide tense patterns in the two sentence types with a different semantic analysis. However, since there is independent syntactic evidence that the Tense domain is absent in amalgams, a difference in temporal semantics is also expected.

8.4.2   Identifying the presence of TT in pseudoclefts

Since the copula and the lexical verb in canonical pseudoclefts cannot have disjoint reference, straightforward diagnostics for tenselessness are unavailable. That is, under normal circumstances, we might expect to find a contrast between the canonical and amalgam copulas with respect to whether they support independent temporal modification. The hypothesis would be that if the amalgam copula is tenseless, in contrast to the canonical copula, it cannot be temporally modi-

fied independently of the embedded clause. Part of this prediction is of course accurate: the copula in amalgams cannot have an independently modified time.

(168) a. *Yesterday, what he liked years ago was he liked coffee.

b. *Yesterday, he liked coffee years ago was what he liked.

Nevertheless, this property does not distinguish between amalgam and canonical pseudoclefts. The matrix and embedded TTs must corefer in canonical pseudoclefts, so disjoint temporal modification is independently ruled out.

(169) a. *Yesterday, what he liked years ago was coffee.

b. *Yesterday, coffee was what he liked years ago.

Some speakers find a contrast between canonical and amalgam pseudoclefts with respect to the following environment.

(170) *He liked coffee was what he liked, but not anymore.

(171) Coffee was what he liked, but not anymore.

This pair suggests that temporal modification of a TT is possible in canonicals, but not in amalgams. The contrast is delicate, however. Some speakers find it very weak—it is conceivable that the *not anymore* modifier could associate with the counterweight proposition. The distribution of temporal modifiers therefore does not offer very robust support for treating the canonical copula as tensed and the amalgam copula as tenseless.

There is one context, however, where the difference between canonical and amalgam pseudocleft tense emerges clearly. When the pseudoclefts themselves are embedded, the tense properties of their copulas can be manipulated. Although [past]-under-[past] embedded stative predicates tend

to be interpreted as simultaneous to the superordinate tense, English also has a past-shifted reading for [past]-under-[past]. The simultaneous reading is the default, but contrastive stress on the tensed copula can make the past-shifted reading more prominent. Section 8.3.4.2 showed that the PAST TT of the copula in an embedded pseudocleft can support a past-shifted reading, particularly when the tensed copula is contrastively stressed.

The question is: when the amalgam pseudocleft is embedded, does its copula show the same patterns of form and interpretation available to embedded statives in other sentences types, including the canonical pseudocleft? If the canonical copula does show the expected patterns, while the amalgam copula does not, I will have found semantic evidence for treating the amalgam copula as tenseless.

Before I can test this hypothesis, a confound must be addressed: amalgam pseudoclefts tend to resist embedding (see Chapter 4). For the purposes of the data reported in this section, the slight degradation some speakers note in finite embedded contexts will be ignored. The reason for this is that the relative judgments speakers give are sensitive to the [tense] contrasts under study, even where the acceptability of the embedded pseudocleft sentence type is globally decreased.

8.4.3 Experimental evidence for lack of semantic tense in amalgams

The prediction concerning the availability of the past-shifted reading in canonical vs. amalgam pseudoclefts was tested experimentally. If the amalgam copula supports a PAST-shifted reading in embedded contexts where temporal alternatives are introduced by contrastive stress, then it is tensed. If it does not, then it is atemporal. The expected pattern, where the amalgam pseudocleft differs from the canonical one, is illustrated in (172).

(172)   I said last week that...

  a. what you liked **was** coffee.   [*past-shifted, ✓simultaneous]

  b. what you liked **WAS** coffee.   [✓past-shifted, ✓verum focus]

  c. what you liked **was** you liked coffee.   [*past-shifted, ✓simultaneous]

  d. what you liked **WAS** you liked coffee. [*past-shifted, ✓verum focus]

Because the judgments are delicate, and even in the canonical pseudocleft baseline, quite gradient, a controlled survey was conducted to test the availability of past-shifting in the two sentence types.

8.4.3.1 Experiment 8 design

Unlike the other experiments presented in this thesis, the present experiment measures participants' reactions to a semantic factor, rather than a syntactic one. That is, the hypothesis cannot be tested with a simple acceptability judgment task. In order to test whether a sentence supports an ambiguity between a preferred reading and a dispreferred reading, a sensitive task is needed. For this reason, the experiment assessed whether participants were able to override the inferences associated with the preferred reading, in favor of the dispreferred reading.

  For example, the zero-PAST-under-PAST gives rise to an inference that the embedded TT overlaps the matrix one:

(173) I heard that John's house was blue. $\longrightarrow$ John's house was blue at the time of my hearing.

The interpreted PAST-under-PAST entails that the embedded TT precedes matrix ET. This interpretation implies that the truth of the embedded proposition no longer holds at matrix ET.

(174) I heard that John's house was blue. $\longrightarrow$ John's house was no longer blue at the time of my hearing.

  In the pseudocleft experiment, participants were presented with pairs of sentences. The first

was a pseudocleft embedded under a PAST-tense verb. The second was a sentence compatible with

the PAST-shifted interpretation, but incompatible with the zero-TENSE (simultaneous) inference.

Participants were instructed to consider the second sentence in the context of the first one, and

asked to rate their level of surprise at the second sentence on a 7-point Likert scale.

If the PAST-shifted reading is available, then the cancelation of the simultaneous reading should

give rise to a lower surprise rating than sentences where PAST-shifting is not available. Some

degree of surprise in both cases is expected, since the simultaneous reading is always preferred. To

manipulate the degree of surprise, half the items had a contrastively stressed copula, which should

make the past-shifted reading more accessible. Participants were trained on practice items (and in

fillers) to interpret capital letters as contrastive stress.

The experiment had a 2x2 design, where the type of pseudocleft (canonical or amalgam) was

crossed with the form of the copula (unstressed or stressed) (see Table 8.1). Four versions of the

experiment were constructed using a Latin square design, where each of the four conditions above

was represented four times in each version, but no version contained any repeated lexicalizations

(for a total of sixteen lexicalizations). The order of item presentation across blocks and versions

was also counterbalanced.

	Unstressed was	Stressed WAS
Canonical	What he liked was coffee	What he liked WAS coffee
Amalgam	What he liked was he liked coffee	What he liked WAS he liked coffee

Table 8.1: Factor design for past-shifting experiment

Because the experimental item type is so salient, each version contained double the number

of fillers as experimental item types, for a total of 48 items per version. The fillers were evenly

distributed and counterbalanced among the versions. Half of the fillers contained a felicitous in-

ference, and half contained a highly infelicitous inference. For example, a low-surprise filler item

is:

(175)   a.   Bob says: I heard that Lee only takes TAXIS in Manhattan.

        b.   Bob also knows that Lee doesn't use any other type of transportation in Manhattan.

(175b) is implicated by the exhaustive focus on *TAXIS*, so participants are predict to find this example unsurprising. A high-surprise filler item is exemplified in (176):

(176)   a.   Bob says: I said that not everyone got an A on the test.

        b.   Bob also knows that nobody got an A on the test.

(176b) cancels the strong scalar implicature associated with (176a), so participants are predicted to find it highly surprising. Filler items used different types of inference, in order to avoid repetition. In addition to exhaustivity and scalar implicatures, fillers tested a range of conventional and syntactically conditioned implicature types. A pilot study revealed outlier filler items, which were discarded and replaced in the final version. The filler mean was used to normalize the results for the experimental items.

An additional control was incorporated in the form of an acceptability judgment sub-task, embedded at random points throughout the task. This sub-task checked that participants found baseline amalgam pseudoclefts at least somewhat acceptable. This control ensured that high surprise ratings of pseudoclefts was not due to syntactic unacceptability.

The experiment was presented using the IbexFarm platform. 36 participants were recruited using Amazon Mechanical Turk (AMT). They were all native speakers of English, raised in the United States, ranging in age from 22 to 60 (median age: 31). Participants controlled the pace of task by clicking or pressing any key. The first key stroke revealed the pseudocleft; the second added the inference and the rating scale. The screenshot in Figure 8.1 shows an example of experimental

item presentation, as it appears on the after the first key stroke.

Participants were instructed to attribute the sentences to a speaker, "Bob", so that they would not insert themselves in the role of speaker. The goal was to eliminate any confounds like the real-world plausibility of these propositions, or the participants' likelihood to use the sentence types in question. The first sentence in the example above shows a canonical pseudocleft with contrastive stress on the copula, and the second sentence is an inference associated with the past-shifted reading, but incompatible with the simultaneous reading. Since the copula is stressed, participants are predicted to be able to accommodate the past-shifted inference with a relatively low degree of surprise.

Bob says:
Yesterday, I found out that what Jane was studying WAS art history.

Bob also knows that:
Jane used to study art history, but she changed her major to political science last semester.

Rate how surprising you find the second sentence.

not surprising :) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | :O very surprising

click any box to continue

Figure 8.1: Example item for past-shifting experiment

8.4.3.2 Experiment 8 results and analysis

The results of five participants were excluded after outlier analysis, since they failed to distinguish accurately between low-surprise and high-surprise filler items. The judgments for the experimental items were normalized against the filler mean. The results are summarized in Figure 8.2.
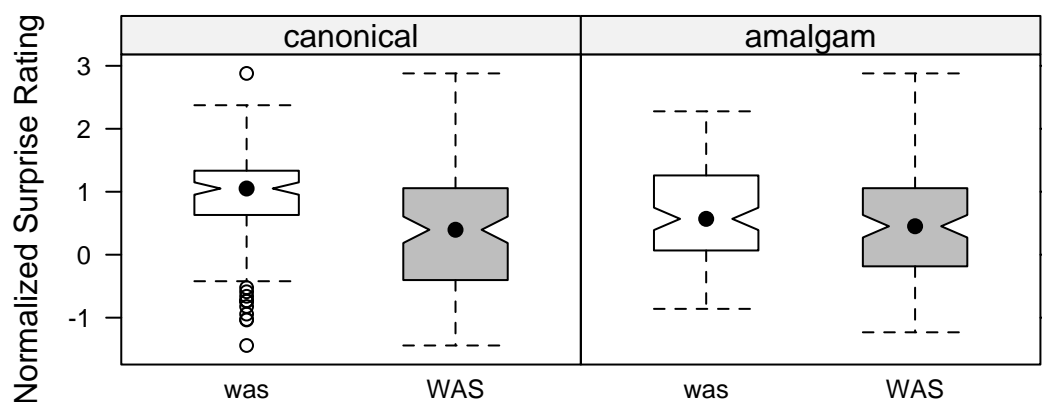

Figure 8.2: Surprise rating for past-shifted reading by sentence type and copular stress

Normalized results were analyzed using a linear mixed effects model, with random intercepts for item, lexicalization, and subject. An ANOVA of the test model with a null model found a significant interaction between sentence type and stress ($p < 0.05$). As the plot shows, stressing the canonical copula significantly reduces the surprise rating, while stressing the amalgam copula has little effect. In all four conditions, the surprise rating has a positive z-score, but this is consistent with the fact that the simultaneous reading is always the default.

This study demonstrates that experimentally obtained data can be brought to bear on even subtle semantic predictions. It also supports the claim that the amalgam copula is atemporal, in contrast to its canonical counterpart.

8.4.4   Interim conclusion: Copular amalgam clauses are tenseless

This section has illustrated that although the morphological [tense] properties of canonical and amalgam pseudoclefts are similar (where the copula is in a simple form), their semantic TENSE properties differ. I used embedded contexts, where the tense properties of the copula can be manipulated, to illustrate the difference between the two sentence types. While the canonical copula supports the PAST-shifted interpretation available to embedded statives in English, the amalgam copula does not. Contrastive stress on the copula, which introduces a set of temporal focus alternatives if it is tensed, was used to facilitate the shifted reading. I presented the results of a semantic inference task that showed a significant contrast between the two sentence types.

I conclude that the amalgam copula projects no ZP Topic Time argument, because it fails to generate a set of temporal alternatives when stressed. This finding is further supported by the syntactic incompatibility of the amalgam copula with T and V-domain phenomena, as discussed at length in Chapter 5. Since the copula can associate with verum focus alternatives, however, it does have some assertive content, namely, it is bound by Force. How that assertion is anchored in the absence of Tense/TENSE is the topic of the next section.

8.5   Deictic anchoring in copular amalgams

Copular amalgams are peculiar in their instantiation of finiteness. Unlike other finite clauses, they lack projections of V and T. Since only the C-domain of the clause is projected in copular amalgams, they allow the deictic component of anchoring to be isolated.

8.5.1   Copular amalgam sentences are anchored

The copula's inability to associate with tense in amalgam pseudoclefts might raise the suspicion that it is not anchored at all; that is, that it is not finite. In Chapters 2 and 5, I provided evidence

against this notion, based on the copular clause's association with illocutionary force. While the copular clause is not the most informative component of the sentence, it contributes an assertion, namely, that the counterweight clause provides the answer to the presupposed question predicate.
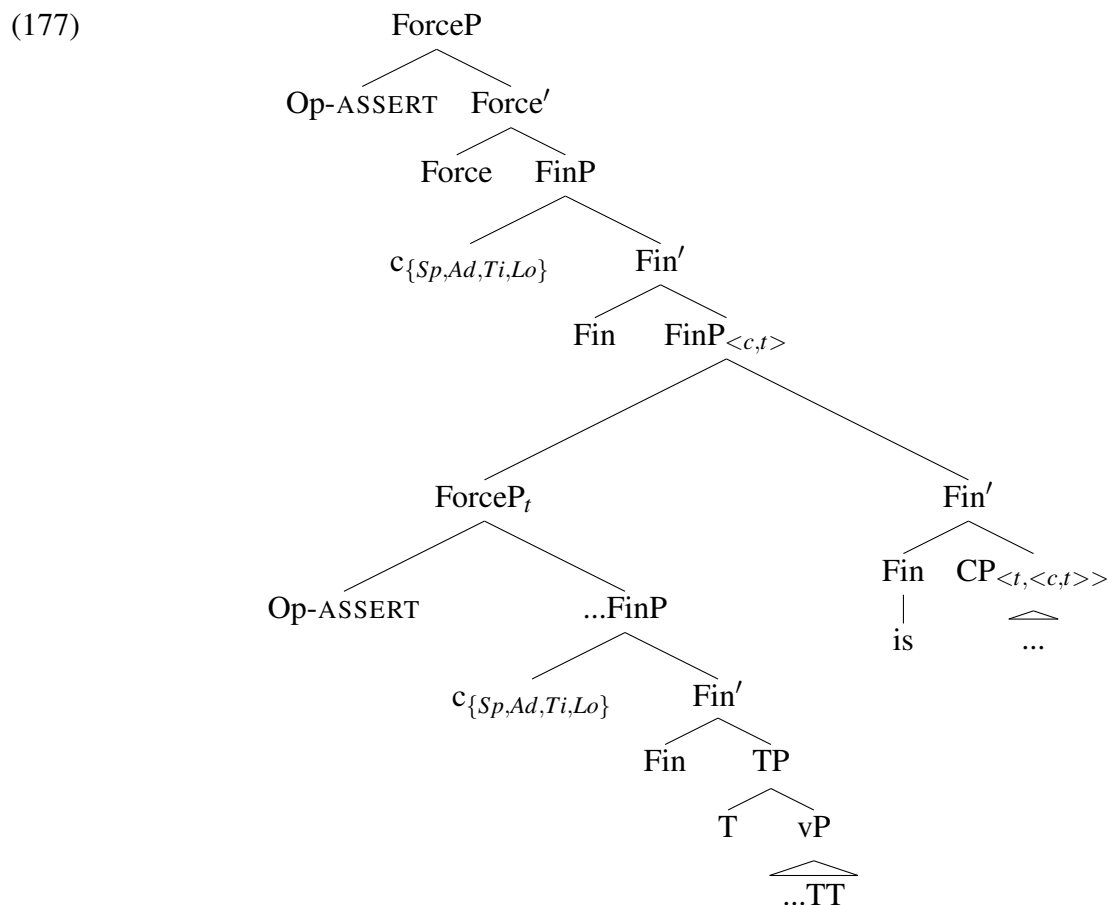
Additional evidence that the copula is finite can be found in the experimental results described in the previous section. The copula's ability to project verum focus alternatives shows that it has basic assertive force. Indeed, Klein (1998) proposes that finiteness is essentially decomposable into "some claim component" (assertion) and a TT. In the copular amalgam, the TT component is stripped away, leaving only assertive force.

In the results of the experiment, the surprise rating with stress on the amalgam copula is almost identical to the unstressed version. This suggests that the verum focus alternatives are readily available when the copula is stressed. If they were not, I would expect the surprise rating to increase, because there would be no coherent way to interpret the stressed copula. An informal consultation with a (non)-linguist informant corroborates this. I asked the informant for her impressions of sentences with stress on the copula, in both [present] and [past] form, without context. She said, "I probably wouldn't say it that way." I asked what it would mean in a context, and she replied, "It means I'm agreeing with someone." Her interpretation is consistent with the verum focus alternatives.

If not a Topic Time, what is the entity in an amalgam pseudocleft that needs to be anchored? Unlike eventualities, which are spatio-temporal entities, or the states in canonical copular sentences, which are identified by temporally located properties, the entity in an amalgam pseudocleft is only world-dependent. More specifically, I take it to be utterance context-dependent. The basic semantic ingredients of amalgam pseudoclefts are: a question and an answer. These ingredients are sentential: they are world-dependent propositions. [24] The amalgam pseudocleft proposition ex-

---

[24]In embedded contexts, like the ones used in the past-shifting experiment, the evaluation world is bound by the

presses a property of the utterance context. The utterance context is deictic, so it does not undergo

further contextual anchoring. The tree in (177) illustrates the deictic anchoring relation.

(177)

```
                        ForceP
                       /      \
                 Op-ASSERT    Force′
                             /      \
                          Force     FinP
                                   /     \
                          c_{Sp,Ad,Ti,Lo}  Fin′
                                          /     \
                                        Fin    FinP_{<c,t>}
                                              /          \
                                        ForceP_t          Fin′
                                       /       \         /    \
                                 Op-ASSERT   ...FinP   Fin    CP_{<t,<c,t>>}
                                             /    \     |       ‿
                                   c_{Sp,Ad,Ti,Lo}  Fin′  is     ...
                                                  /    \
                                                Fin    TP
                                                      /   \
                                                     T    vP
                                                           ‿
                                                          ...TT
```

The copula is the matrix, finite element of the amalgam pseudocleft. Its proposition is not tem-

porally anchored, since T and its local ZP are unavailable; instead, it is anchored by way of a

deictic utterance context element in the context domain. The copula's interpretation always has an

indexical component; for instance, it is never PAST (in root contexts). Compare (178) and (179).[25]

(178)   Her hair, prior to now, was blonde.

(179)   *She needed a break, prior to now, was what she needed.

_____

embedding verb.

[25](179) only allows *prior to now* to be interpreted as modifying the *need*-clause.

Since the copular clause is anchored deictically, it cannot be temporally displaced.

8.5.2 Anchors in the left periphery

How can we implement this kind of deictic anchoring syntactically? A proposal that has existed in many forms since Ross (1967) is that elements of the speech act itself are projected in the left periphery of the clause. While semantic studies of the utterance context often assume that natural language includes utterance context expressions, which encode not just time, but also world, participants, and location, the syntax literature did not use them much until the minimalist era.

Sigurðsson (2004) offers a somewhat unconventional syntactic analysis of the relationship between the speech context and the lower domains, which rings very true for the findings related to copular amalgam sentences. Sigurðsson (2004) partitions the clause into three functional domains.

(180)   [speech event [grammatical domain [event domain]]]

These correspond to context, displacement, and classification, in model used in this dissertation. Sigurðsson proposes that the speech event domain contains syntactically projected features associated with the indexical elements of the speech event:

(181)   Speech event: $\{S_T, S_L, \{\Lambda_A, \Lambda_P\}\}$ (Sigurðsson 2004:226 (15))

S_T is the speech time, S_L is the speech location, Λ_A is the "logophoric agent", and Λ_P is the "logophoric patient". These features are the anchors for tenses, existential and expletive expressions, agreement markers and pronouns. Since they are interpreted at the LF interface, and have consequences for PF, Sigurðsson argues that speech event features are represented in narrow syntax, just like the more traditionally recognized features of overt elements.

For Sigurðsson, the function of the grammatical domain is to anchor the event domain to the speech domain by mediating a matching relation between event time, location, and participants and speech time, location, and participants. Positive and negative matches are interpreted differently by the morphological component. For example, first person agreement is the result of a match between a participant in the event and the logophoric agent. Similarly, present tense results from a match between an event time and a speech time. In copular amalgams, which contain no event and no event participants, but only speech acts and speech act participants, it is therefore totally unsurprising that the grammatical domains should be absent.

8.5.2.1 Anchoring to the speaker

Giorgi (2010) (see also earlier work; Giorgi and Pianesi 2004, etc.) focuses on the syntactic representation of the speaker. She proposes that the speaker variable in the left periphery of the clause contains coordinates (essentially, indices) for time and location. Therefore, the speaker in the C-layer of the clause provides the external center of deixis or anchor for tense. A dedicated speech time need not be projected. The speaker's coordinates are privileged in interpretation; for example, double-access readings come about because an embedded proposition is interpreted with respect to the temporal perspective of the embedded speaker (or thinker) and the matrix one.

Giorgi's (2010) model treats the left peripheries of embedded attitude propositions as complex. The T-layer includes the coordinates of the matrix event (specifically, the coordinates of the matrix attitude-holder, although the attitude-holder is ultimately spelled out as the subject of the higher clause), while the C-layer includes the coordinates of the actual speaker.[26] The event in the embedded clause is then interpreted twice: once with respect to the attitude-holder's coordinates (RT) in Spec,TP, and again with respect to the speaker's coordinates (UT) in Spec,CP. The double

---

[26]While Giorgi situates the matrix event variable directly in Spec,TP, I would treat it as the binder of RT.

interpretation is facilitated, for Giorgi, by remerging T in C. This structure has the consequence that all attitude complements of this form have double-access readings.[27]
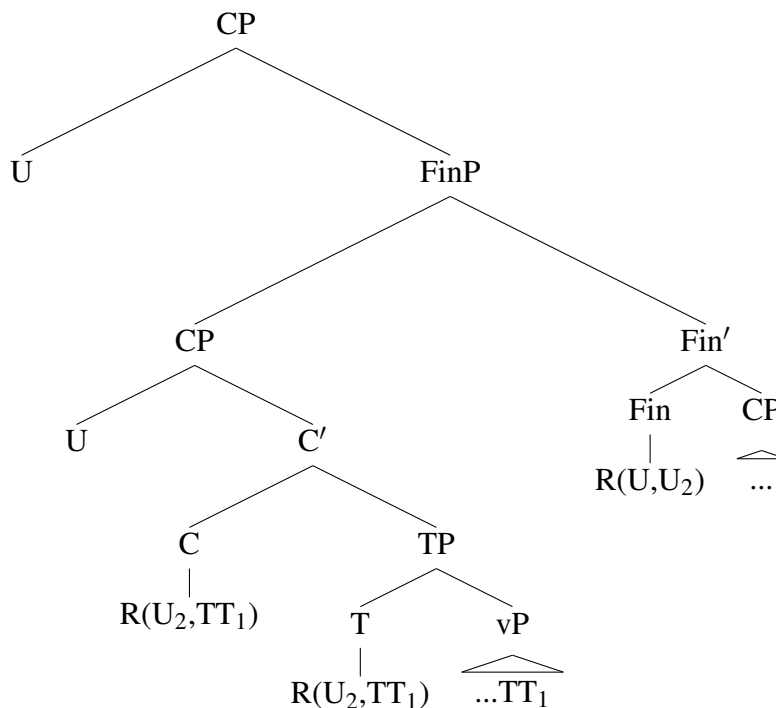
Consider how this model applies to amalgam sentences, which lack an attitude-holder. The only speech event in the matrix (copular) clause is the actual speech event—encoded in the coordinates of the actual speaker, which occupies Spec,Fin.[28] The embedded clause (the counterweight) includes the higher event's speaker coordinates as its reference variable, which happen to be those of the actual speaker, since no other lexical verb of saying is present. Double-anchoring of the clause to the coordinates of the actual speaker, separately projected in the edge of the embedded C-layer, does not entail displacement: the internal speaker and the external speaker are one and the same. This is the desired interpretation.

A syntactic model of anchoring incorporating Giorgi's (2010) speaker variables is sketched in the tree below. The variable U represents the coordinates of the utterance event—those of the actual speaker.

---

[27]Giorgi analyzes different tense embedding patterns by classifying the lexical properties of the verbs that select them. For example, verbs of saying require double-evaluation of the embedded event, while verbs like *dream* do not include a representation of the speaker in their complement clauses, so the content of the dream is evaluated only with respect to the subject.

[28]Just as in the analysis of TP, I assume that there must be two specifier positions, one to accommodate the anchoring variable and one to accommodate the structural subject.

(182)

```
                              CP
                     ┌────────┴─────────┐
                     U                 FinP
                              ┌──────────┴──────────┐
                             CP                    Fin′
                        ┌─────┴─────┐           ┌────┴────┐
                        U          C′          Fin       CP
                              ┌─────┴─────┐      │        ◁
                              C          TP   R(U,U₂)    ...
                              │       ┌────┴────┐
                           R(U₂,TT₁)  T        vP
                                      │       ◁
                                   R(U₂,TT₁)  ...TT₁
```

In the embedded clause (the counterweight), a temporal relation is established between the embedded TT and the local RT. In this adaptation of Giorgi's (2010) formulation, the RT is provided by the coordinates of the speaker in the event of saying encoded in the matrix verb (U_2). In ordinary embedded clauses, these are the coordinates of the event of saying, believing, etc. contributed by the matrix verb and its subject. In amalgams, the matrix event is the actual speech event. There is no temporal displacement in the Fin-assertion. Fin behaves like a simple Relator. The result is that the utterance context of the root counterweight clause is identified with the utterance context of the copular clause.

8.5.2.2 Anchoring to embedded speaker

The type of speaker context-anchoring in amalgams exhibits some interesting contrasts with free indirect discourse with respect to indexical shift. For instance, while free indirect discourse preserves person indexicals and the present tense indexical, it allows locative and adverbial indexicals to shift to the embedded speaker's perspective.

(183) a. She$_i$ wondered what she$_i$ would do tomorrow.

 b. *She$_i$ wondered what I$_i$ will do tomorrow.

In amalgams, since [tense] is not TENSE, a [present] copula is interpreted simply as speaker-anchored—it can be anchored to a local speaker in embedded contexts, and escapes the res-movement required of the [present] when associated with an actual TENSE. (184) has no particular present-time orientation, while the copular clause in (185) must receive the double-access reading, where the relevant interval where the pseudocleft is asserted to hold spans John's present and the speaker's present.

(184) John said that what he wanted is he wanted coffee.

(185) John said that what he wanted is coffee.

These judgments are subtle, but informal consultations with three skilled non-linguist informants yielded a reliable contrast between the interpretations of the two sentence types in different embedded environments.

8.5.2.3 Other sentence types with deictic anchoring

The deictic-type anchoring found in amalgam copular clauses may seem to be rather anomalous—in a tensed language, a proposition is interpreted as deictically anchored, by way of indexicals in the left periphery. There are other phenomena, however, which could fall under a similar explanation.

For example, "Presentative Dative" clauses, analyzed by Horn et al. (2015), are a good candidate for a deictically anchored sentence type.

(186) Here's you a cup of coffee.

Presentative datives, like copular amalgams, include a copula with limited functional structure. For

example, the copula cannot be negated (187) or temporally shifted (188):

(187)   a.   *Here's not you a cup of coffee.

          b.   *Here isn't you a cup of coffee.

(188)   *Here was (you) a cup of coffee.

Like the amalgam pseudocleft, the dative argument in the presentative sentence must be first or second person—bound by the speech act participant indexicals in the left periphery of the clause. The structural subject of the presentative dative clause is also a deictic element, which independently refers to the utterance context. This presents another interesting parallel with copular amalgams, where the sentential subject is also deictic: it is an independent finite clause. The structure of the presentative dative is thus consistent with the structure I have proposed for amalgam pseudoclefts, where a purely deictic structural subject is licensed in the absence of the anchoring domain of the clause.

The copula in the presentative dative may well be atemporal, merged in the left periphery, just like its counterpart in the amalgam pseudocleft. Given the similarity between the two sentence types, the presentative dative provides useful indirect evidence that [past] morphology on the amalgam copula arises via concord. Where [past] is illicit in amalgams, it is unclear on the surface whether it is because of a semantic clash with the embedded [present] or because there is no valued instance of [past] in the structure. In the presentative dative, there is no other [tensed] verb in the structure, and yet the copula must take [present] form. The unavailability of [past] is not plausibly due to a clash with a more deeply embedded [present] in this case. Therefore, by analogy, [past] on the amalgam copula must be morphologically dependent on another [past]. It is not surprising that [present] should be the default realization of Fin's [T]-feature in amalgams (and presentative

datives): [present] is after all the indexical tense in English.

8.5.3   Free indirect discourse

I suggested above that since amalgam pseudoclefts feature an embedded clause which has root indexical properties, they are reminiscent of free indirect discourse, a literary style where some indexicals seem to be interpreted with respect to an internal context, rather than the external context (e.g., Schlenker 2004; Sharvit 2008; Giorgi 2010).

(189)   Tomorrow, she thought, she would leave here.

In (189), *tomorrow* is evaluated from the subject's perspective, as is the locative expression *here*, rather than from the perspective of the speaker (or writer). The embedded subject is ascribed some mental event, and the indexical expressions are part of the content of that event. In free indirect discourse, some proposition with root-syntax is focal, and typically some material setting up the indirect speech (or thought) report is backgrounded.

   Amalgam pseudoclefts have a similar character. The counterweight proposition has root syntax, and is focal with respect to the backgrounded open question. The crucial distinction between free indirect discourse contexts and copular amalgam clauses is that the embedded and matrix speakers are the same: by relating the speakers, Fin relates the two utterance/mental events. In (190), for example, *tomorrow* is interpreted the same way from the perspective of the internal and external speaker, because they are identical.

(190)   He should leave tomorrow is what he should do.

   Amalgam pseudoclefts, where an embedded speaker is anchored to a matrix one, are thus like a peculiar type of free indirect discourse. Intuitively, the speaker reports her own utterance or mental

event inside the copular clause. This intuition is reflected in the explicit cases where the predicate is a comment on the embedded assertion:

(191)   He should leave tomorrow is what I mean.

Their formal similarity with free indirect discourse contexts might be behind the intuition that the sentential subject of an amalgam is like a quotation; however, as I argued in previous chapters, it does not pattern semantically or syntactically with direct speech reports.

From an on-line perspective, this type of free indirect discourse is useful for making metalinguistic comments. The speaker asserts something, and then projects a structurally minimal clause around it (the FinP), to make the assertion a subordinate part of a larger proposition.[29]

The subject of an amalgam is not just a proposition, but a speech act.

8.5.4   Against an alternative: DISTAL/COMP-valuation

Clearly, [past] is not in free variation with [present] throughout English, but perhaps it need not always associate with a TENSE. A possibility worth entertaining is that [past] in an amalgam pseudocleft could signal the presence of a more general category, like DISTAL (Iatridou 2000) (the [−coincidence] specification of the anchoring category) (Ritter and Wiltschko 2014). This category expresses lack of coincidence with the reference situation (e.g., Hale 1986; Demirdache and Uribe-Etxebarria 2007; Ritter and Wiltschko 2014), and is not necessarily temporal. It can anchor to different components of the deictic utterance situation—time, participants, locations, or worlds. If [past] associates with a non-temporal version of DISTAL, then it could indicate that the topic

---

[29]A somewhat playful colloquial construction with a similar metalinguistic commenting function has become popular recently, (i).

(i)   I don't like chocolate—said nobody ever.

While the "counterweight" clause in this structure is a quotation, the function of this construction is to pretend to assert some proposition, and then metalinguistically negate it.

situation holds of a non-coincidental person (third person/switch reference), a different location, or a different world (counterfactuality), rather than a different time.

Ritter and Wiltschko (2014) and Wiltschko (2014) propose that the "substantive content" of elements in the complementizer domain can value the feature housed on T, the locus of anchoring, if T itself lacks anchoring content ("COMP-valuation"). COMP-valuation accounts for environments where non-declarative Force is in complementary distribution with substantive anchoring. In these cases, morphological [tense] is sensitive to PROXIMATE/DISTAL content in the C-domain, rather than to a temporal ordering predicate located in the T-domain. Since the amalgam copula heads a special clause type, is located in the C-domain, and lacks a T-domain, this line of analysis ought to be considered.

Ritter and Wiltschko (2014) demonstrate that COMP-valuation is at work in imperatives and counterfactuals. In imperatives, which in many languages, including English, lack [tense] (or the analogue of [tense] sensitive to other anchoring categories), T's feature is valued as [+coin] by COMP; that is, anchoring is PROXIMATE. Similarly, in Halkomelem Salish, locative auxiliaries, which in declarative finite clauses encode the anchoring function, must be absent from imperatives (see their Section 4). Counterfactuals, by contrast, receive [−coincidence] marking: in English, this is [past], and in Halkomelem, this is [distal], the non-coincidental locative auxiliary.

Suppose that copular amalgam clauses also involve a type of COMP-valuation. They are declarative, so it would be surprising to find COMP-valuation, but they are after all anchored in the C-domain. In this exploration of COMP-valuation as applied to amalgams, I depart from Ritter and Wiltschko's (2014) formulation by maintaining the proposal developed earlier in this thesis on syntactic grounds that no T/Infl domain is projected in amalgams. COMP-valuation therefore could not involve substantive content in COMP transmitting its value down to Infl in these cases,

but rather, the substantive content would value a morphological feature expressed locally on Fin, where I assume the copula is situated. To determine whether COMP-valuation is a viable explanation for the [present]/[past] alternation in amalgams, we must test whether there is any interpretive difference between sentences with the two morphological specifications.

Consider (192):

(192)   a.   What Joe did is he complained to the manager.

         b.   What Joe did was he complained to the manager.

If the difference is not temporal, then it must be either participant-based, locative, or modal. Clearly, it is not the first: the alternation between (192a) and (192b) does not mean that in the former, the embedded proposition is ascribed to the same speaker as the copular proposition, while in the latter it is not—(192b) is not an indirect speech report. In both cases, without further modification, the embedded and matrix speakers are one and the same. This can be further demonstrated by adding context-shifting modifiers, which are insensitive to the [tense] form of the copula, and always shift the matrix context.

(193)   a.   What Joe did, according to Lulu, is he complained to the manager.

         b.   What Joe did, according to Lulu, was he complained to the manager.

The difference between [present] and [past] also does not signal a locative difference. This is obvious.

The last possibility can also be rejected after consideration. If the [tense] morphology indicates modal anchoring, then the [past] tense should give rise to a counterfactual or conditional interpretation: it should indicate that context where the counterweight specifies the answer to the question in the weight is non-coincidental with the actual world. This is ordinarily difficult to tease apart

from PAST tense, but in the case of amalgams, the matter is more straightforward. [Past] on the copula does not indicate counterfactuality: the proposition *he complained to the manager* serves to answer the question of *what Joe did* in the actual world.

Although COMP-valuation offers an analysis of "fake" morphological [tense] (and other anchoring morphology) that gives insight into the relationship between clause-typing and finite anchoring, as well as into robust cross-linguistic patterns, it appears that it cannot account for the form/meaning mismatch in [past]-marked amalgam copular sentences. COMP-valuation, therefore, must work as Ritter and Wiltschko (2014) propose: it contributes a type of dependent anchoring, which requires the presence of both the context and displacement domains.

When the displacement domain of the clause is not projected, there can be no proximate/distal contrast in the anchoring mechanism—only deictic anchoring is possible. The dependent anchoring content of the T-domain can creep into the C-domain in ordinary clauses, through feature dependencies like COMP-valuation contexts described above, or through T-to-C movement, where a dependent element is remerged in C and spelled out as a subordinating complementizer. When anchoring domain content is completely severed from the C-domain, as in copular amalgams (and perhaps presentative datives), then anchoring is necessarily deictic.

8.6 Accounting for [tense] form alternations on the copula

I have argued that while the copula of canonical pseudoclefts is semantically TENSED, the copula of amalgam pseudoclefts is only morphologically [tensed]. This distinction has previously gone unnoticed, primarily because amalgam pseudoclefts have not been studied closely in the literature, but perhaps also because the two sentence types feature superficially similar [tense] combinations. Section 8.3 offered a semantic explanation for the constraints on TENSE in canonical pseudoclefts,

building on work by Sharvit (2003) and Romero (2004). Amalgam pseudoclefts, however, are not amenable to the same explanation, since their copula does not make any contribution to their temporal semantics. The restrictions on their [tense] forms must, therefore, have a morphological source. This brief section accounts for the morphological alternations on the copula.

8.6.1 Possible sources for uninterpreted [tenses]

[Tense] form is often distinct from the TENSE meaning. The obvious example of this is SOT, where the morphological features of a zero-TENSE are valued under binding.

A "fake" past [tense] occurs in non-final serial verbs in Malayalam (Dravidian):

(194)   ñaa<u>n</u> oru maanga poTT-iccu muR-iccu ti<u>nn</u>-um.
        I   a   mango pick-PAST cut-PAST eat-FUT
        'I will pick, cut, and eat a mango.'   (Jayaseelan 2013:193 (4))

Although the verbs are interpreted with FUTURE tense, all of the non-final verbs in the series must inflect for the [past] and are formally identical to verbs interpreted with PAST tense.

Another environment where we find so-called "fake" tense is in counterfactuals and subjunctives; the presence of the future-oriented adverbial *tomorrow* makes it clear that the [past] form *left* cannot be interpreted as a simple past:

(195)   If you left tomorrow, I would be sad.

The counterfactual fake past, which occurs in languages other than English as well (see e.g., Iatridou 2000; Bjorkman and Halpert to appear; Arregui 2009; Ritter and Wiltschko 2014), has been analyzed as contributing a meaning of remoteness: remoteness from the actual world (Ritter and Wiltschko 2014) or temporal remoteness. This is the COMP-valuation pattern described in the previous subsection, which was shown not to be a possible source for [past] on the amalgam pseu-

docleft copula.

The most basic morphological explanation for matching [tenses] relies on Agree. Agree does

not offer an explanation for [tense] form in amalgam pseudoclefts, because, as the previous sections

showed, the matching requirement is optional: the copula can always occur in [present] form.

Agree, however, is not optional, so it is not responsible for the matching tendency.[30]

I also must reject a polarity account like that of Stowell (2007), where morphological [past]

must occur in the scope of PAST. The only PAST available in the amalgam pseudocleft is the ZP

in Spec,TP of the counterweight clause (or in the *wh*-initial version, Spec,TP of the weight clause).

Since the PAST-ZP is inside a specifier island, it is too deeply embedded to license a polarity-

sensitive morpheme in the matrix clause.

Instead, I propose that [past] on the amalgam copula can be licensed in a concord relation.

8.6.2   [tense]-concord

Let us assume that the specification of the [tense]-feature on the amalgam copula can come about

in one of two ways: (i) default or (ii) concord.

The default is [present]—as I suggested above, [present] form is associated with deixis in its

temporal environment, so it is the likely candidate for the default [tense] in a deictic anchoring

environment.

In a concord relation, multiple occurrences of a morphological feature depend on only one

occurrence of that feature's semantic correlate. For instance, in a negative concord sentence like

(196), there is both a sentential negation morpheme and a constituent-level negative morpheme,

but only one negation.

---

[30]In Chapter 7, I proposed that Fin's [T]-feature is valued under spec-head agreement with the counterweight. I
assume that this configuration values it as [+T], but does not endow it with a specific [present]/[past]-feature.

(196)   She don't have no food.

The advantages of treating [past] in amalgam pseudoclefts as arising from concord are that it entails nothing about semantic binding (unlike feature transmission) and it is optional (unlike Agree) (Khanjian 2013). The [past] form is available to the copula simply when there is a local [past]. In the copular amalgam, this is the spec-head relation with a [past]-tense CP (where C Agrees with T in its own clause):

(197)   [She **wanted** some cookies] **was** what she wanted.

8.6.3   [past] without [past]

The concord analysis of [past] on the amalgam copula also offers an explanation for why some predicates that do occur in [past]-under-[past] SOT contexts nevertheless do not license [past] on the copula:

(198)   I said that he should call the manager.

(199)   *What he should do was he should call the manager.

Somehow, the formal system needs to capture the intuition the copula cannot be [past] when the main verb is not notionally PAST. In the tense syntax model I presented early in the chapter, the modal time associated with *should* is a zero-ZP: it includes no anteriority predicate. As such, its local T is not valued for [past] in (199). Since its local T (and C) are not [past], they cannot provide a [past]-feature for the copula.

In the SOT environment (198), where the zero-ZP is bound by a PAST, it receives the features of the higher verb, and its T is [past]. This is why in the embedded pseudocleft in (200), the past form copula becomes available:

(200)   I said that what he should do was he should call the manager.

If the perfective auxiliary *have* is included, the [past] copula is licensed:

(201)   What he should have done was he should have called the manager.

This is unsurprising if the perfective includes the PAST predicate BEFORE. The scope relations involved in computing modal time with respect to aspect are quite complex (Condoravdi 2002; Hacquard 2006; Demirdache and Uribe-Etxebarria 2014), but given (201), it must be the case that the ZP in Spec,AspP scopes high enough to value [past] on T.

Related to this observation, in languages like German and French, a formally present tense auxiliary verb conveys past semantics; these languages use present perfect as their unmarked past tense form.

(202)   J 'ai   fini   les devoirs.
          I have finish.PART the homeworks
          'I finished my homework.'

(203)   Ich habe das Wasser getrunken.
          I   have the water   drink.PART
          'I drank the water.'

In embedded contexts under a present perfect form verb in German, a simple [present] form is licit, and receives a simultaneous interpretation.

(204)   Ich habe gedacht,   dass du so klug **bist**.
          I   have think.PART, that you so smart be.PRES
          'I thought you were this/so smart.'

We might expect the present tense auxiliary verb form to be compatible only with a present tense copula in a pseudocleft, since it does not transmit [past] in complement clause environments. This prediction is not borne out, however. The notional PAST of the present perfect verb is compatible

with the morphological [past] on the pseudocleft's copula.

(205)   Was ich gegessen habe **ist/war** ich habe einen Apfel gegessen.
          what I   eat.PART have **is/was** I   have an   apple eat.PART
          'What I ate is/was I ate an apple.'

It is clear from these data that tense forms on the copula and lexical verb in a specificational pseudocleft do not adhere to the canonical SOT pattern. Yet, the copula in amalgams is atemporal, so it should not be sensitive to semantics. I maintain, therefore, that the pattern is due to concord. As in the modal environment, the ZP including the BEFORE predicate must raise high enough to value T as [past]. This is particular to languages that use present perfect form with simple PAST meaning; in English, the counterpart of (205) does not license a [past] copula:

(206)   *What I have eaten was I have eaten an apple.

8.6.4   Summary

This section proposed that the mismatch between TENSE and [tense] in amalgam copular sentences is due to a concord process. The amalgam copula is not associated with TENSE—or with an anchoring category housed in the C-domain that supports a morphologically expressed [proximate]/[distal] alternation—so its morphological [tense] alternation must be licensed by something in a different clause.

Agreement and polarity were rejected as sources for the morphological dependency, and concord was adopted instead. A PAST-ZP containing the predicate BEFORE values its local T (and indirectly, C) as [past], enabling the counterweight CP to value Fin's T feature as [past] via a spec-head concord relation.

8.7 Conclusion

This chapter has examined how the amalgam pseudocleft is anchored to the utterance context without a structural or semantic instance of Tense. I sketched a model of tense syntax where semantic TENSES are pronouns in specifier positions, while the heads Asp and T are vacuous relators with only morphological instances of [tense]. I then applied this pronominal model of tense to the patterns of tense embedding in canonical specificational pseudoclefts.

Tense in canonical pseudoclefts is a contentious issue. At the outset of this chapter, there was no established baseline analysis that could be used for comparison with the tense properties of amalgams. I proposed that the intensional properties of the weight clause, along with the information structure properties of specificational sentences, contribute to the requirement that the TTs of the copula and the lexical verb be covalued. I showed that such an approach improves upon atemporal or strict binding analyses in predicting similarities between SOT and non-SOT languages, and between specificational sentences and sentences with topicalized attitude complements.

Having established that the canonical specificational copular sentence has a TT argument, I then argued that the amalgam pseudocleft clause lacks a TT argument. I presented experimental and informal evidence that a distinct TT interpretation cannot be coaxed out from the amalgam copula, even under contrastive focus.

Given the syntactic and functional analysis of the context domain of the clause in copular amalgams, developed in Chapters 5 and 7, I proposed that copular amalgams are anchored deictically, by way of an indexical context/speaker pronominal in a specifier of Fin,P. Fin's function is simply to relate an internal speaker indexical (the counterweight's speaker) to an external speaker indexical (the amalgam clause's speaker). The presence of an internal and external speaker, where both associate with indexical elements, suggest a parallel between copular amalgams and free indirect

discourse constructions. Crucially, however, since the internal and external speakers in amalgams

are one and the same, there is no displacement. Copular amalgams are strictly deictic.

Chapter 9

Conclusion

In this dissertation, I have investigated the structure of the English copular amalgam sentence, and its implications for the theory of finiteness in generative grammar. I have argued that the structural domain of Finiteness alone can constitute an independent clause, even in a tensed language like English. The copular amalgam, which lacks a lexical Verb and has a logical subject that is neither syntactically nor semantically dependent on grammatical Tense, shows that the structural domain of Finiteness is sufficient to render a clause fully finite.

This work began with the observation that the finite copula in English can relate root clauses, the syntactic environment ordinarily occupied by coordinators. While root clauses can be embedded in complement positions, they cannot ordinarily serve as subjects of predication or structural subjects. In a series of formal acceptability surveys, I compared amalgam copular sentences with their canonical specificational pseudocleft counterparts. These experiments showed a clear contrast between the amalgam and canonical sentence types with respect to restrictions on the form of the copula. Despite variation in the acceptability of different sentence types, formal judgement experiments can clarify the empirical status of much-studied phenomena, like the canonical pseudocleft, and under-studied phenomena, like the copular amalgam.

On the basis of these findings, I argued that the copula in amalgam copular sentences associates

directly with Fin or Top. Copular amalgam clauses are projected from a functional head in the clausal left periphery, rather than from T or V. I showed that the functional spine of the clause can be truncated from above or below at different levels of structure, leading a to different distribution of finiteness phenomena in different copular amalgam clause types.

The absence of the T-domain in amalgams was motivated on empirical grounds, but also on the basis of theoretical considerations regarding sentential structural subjects. The Agree relation between Fin and its structural subject is similar to the canonical grammatical subject relation, but it can be established in the minimal domain of the predication structure. While ordinary "tense–case" connection phenomena depend on the relation between a dependently anchored clausal T-domain and a dependently anchored argument D-domain, the parallel connection in amalgams is established between two deictically anchored C-domains. The parallelism between the two relations is aligned with recent efforts in Minimalist syntax to reformulate the "tense–case" connection as a relation between the dependent (e.g., temporal) anchoring features of the clause and the nominal argument.

Finally, the absence of the T-domain makes a prediction about how the amalgam clause is anchored to the utterance context. Without Tense, the clause lacks a Topic Time argument and temporal ordering predicate, so the copular proposition cannot be interpreted as temporally displaced, like an ordinary finite clause. While the canonical specificational pseudocleft proposition has a Topic Time, the amalgam copular clause does not. In the absence of Tense, the copular amalgam clause is anchored *deictically* to the utterance context. The proposition is independently anchored to the utterance context—a hallmark of finiteness, but it is not displaced from it. The copular amalgam clause is used to assert a property of the utterance context, rather than a property of an eventuality. This relation is behind the intuition that copular amalgams encode a peculiar

type of free indirect discourse, in which the speaker simultaneously *utters* and *reports* her own assertion.

The analysis in this dissertation focused entirely on copular amalgam clauses, only scratching the surface of the range of strategies available in natural language for encoding displacement. It may be fruitful to extend a Tense-less approach to other clauses that lack a dependent subject grammatical relation and temporal displacement, like presentational clauses. An avenue of future research is to examine how deictic anchoring manifests itself in languages that have non-temporal anchoring in canonical finite clauses.

Bibliography

Abney, Steven. 1987. The English noun phrase in its sentential aspect. Doctoral Dissertation, Massachusetts Institute of Technology.

Abusch, Dorit. 1988. Sequence of tense, intensionality and scope. In *Proceedings of the West Coast Conference on Formal Linguistics*, volume 7, 1–14. Stanford, California: CSLI Publications.

Abusch, Dorit. 1997. Sequence of tense and temporal *de re*. *Linguistics and Philosophy* 20:1–50.

Adger, David. 2007. Three domains of finiteness: A minimalist perspective. In *Finiteness: theoretical and empirical foundations*, ed. Irina Nikolaeva, 23–58. Oxford University Press.

Akmajian, Adrian. 1979. *Aspects of the grammar of focus in English*. New York: Garland.

Alrenga, Peter. 2005. A sentential subject asymmetry in English and its implications for complement selection. *Syntax* 8:175–207.

Amritavalli, R. 2013. Separating tense and finiteness: Anchoring in Dravidian. *Natural Language & Linguistic Theory* 32:283–306.

Andersen, Gisle. 2002. Corpora and the double copula. *Language and Computers* 40:43–58.

Arche, María J. 2006. *Individuals in time: tense, aspect and the individual/stage distinction*, volume 94. John Benjamins Publishing.

Arregui, Ana. 2009. On similarity in counterfactuals. *Linguistics and Philosophy* 32:245–278.

Asher, Nicholas. 2000. Events, facts, propositions, and evolutive anaphora. In *Speaking of events*, ed. James Higginbotham, Fabio Pianesi, and Achille C Varzi, 123–150. Oxford University Press.

Authier, Jean-Marc. 1992. Iterated CPs and embedded topicalization. *Linguistic Inquiry* 329–336.

Baker, Mark C. 2003. *Lexical categories: Verbs, nouns and adjectives*, volume 102. Cambridge University Press.

Barbiers, Sjef. 2005. Word order variation in three-verb clusters and the division of labor between generative linguistics and sociolinguistics. In *Syntax and variation: Reconciling the biological and the social*, ed. Leonie M.E.A. Cornips and Karen P. Corrigan, 233–264. John Benjamins

Publishing.

Bard, Ellen Gurman, Dan Robertson, and Antonella Sorace. 1996. Magnitude estimation of linguistic acceptability. *Language* 72:32–68.

Bayer, Josef. 1999. Final complementizers in hybrid languages. *Journal of Linguistics* 35:233–271.

Becker, Misha. 2004. *Is* isn't *be*. *Lingua* 114:399–418.

Benmamoun, Elabbas. 2008. *Clause structure and the syntax of verbless sentences*, volume 45. MIT Press.

Bianchi, Valentina. 2003. On finiteness as logophoric anchoring. In *Temps et point de vue/Tense and point of view*, ed. Jacqueline Guéron and Liliane Tasmovski, 213–246. Nanterre: Université Paris X Nanterre.

Biswas, Priyanka. 2013. The role of tense and agreement in the licensing of subjects: Evidence from participial clauses in Bangla. *Natural Language & Linguistic Theory* 32:87–113.

Bittner, Maria. 2005. Future discourse in a tenseless language. *Journal of Semantics* 22:339–387.

Bittner, Maria, and Ken Hale. 1996a. Ergativity: Toward a theory of a heterogeneous class. *Linguistic Inquiry* 27:531–604.

Bittner, Maria, and Ken Hale. 1996b. The structural determination of case and agreement. *Linguistic Inquiry* 27:1–68.

Bjorkman, Bronwyn. 2011. Be-ing default: The morphosyntax of auxiliaries. Doctoral Dissertation, Massachusetts Institute of Technology.

Bjorkman, Bronwyn, and Claire Halpert. to appear. In search of (im)-perfection: The illusion of counterfactual aspect. In *Proceedings of NELS*, volume 42.

Blanchette, Frances. 2013. Negative concord in English. *Linguistic Variation* 13:1–47.

Bliss, Heather, Elizabeth Ritter, and Martina Wiltschko. 2011. A comparative analysis of theme marking in Blackfoot and Nishnaabemwin. In *Papers of the 42nd Algonquian Conference*, ed. Monica Macaulay and Rand Valentine.

Bohnemeyer, Jürgen. 2002. *The grammar of time reference in Yukatek Maya*. Lincom Europa.

Borer, Hagit. 2005. *Structuring sense*. Oxford: Oxford University Press.

Bošković, Željko. 1997a. Pseudoclefts. *Studia Linguistica* 51:235–277.

Bošković, Željko. 1997b. *The syntax of nonfinite complementation: An economy approach*. 32.

MIT Press.

Bošković, Željko, and Howard Lasnik. 2003. On the distribution of null complementizers. *Linguistic Inquiry* 34:527–546.

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24:591–656.

Brenier, Jason, and Laura Michaelis. 2005. Optimization via syntactic amalgam: Syntax-prosody mismatch and copula doubling. *Corpus Linguistics and Linguistic Theory* 1:45–88.

Broekhuis, Hans. 2013. Feature inheritance versus extended projections. Ms., Meertens Institute, Amsterdam.

Bybee, Joan, Revere Perkins, and William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. University of Chicago Press.

Calude, Andreea S. 2008. Demonstrative clefts and double cleft constructions in spontaneous spoken English. *Studia Linguistica* 62:78–118.

Caponigro, Ivano. 2003. Free not to ask: On the semantics of free relatives and wh-words cross-linguistically. Doctoral Dissertation, University of California, Los Angeles.

Caponigro, Ivano, and Kathryn Davidson. 2011. Ask, and tell as well: Clausal question-answer pairs in ASL. *Natural Language Semantics* 19:323–371.

Caponigro, Ivano, and Daphna Heller. 2007. The non-concealed nature of free relatives: Implications for connectivity in specificational sentences. In *Direct compositionality*, ed. Chris Barker and Pauline Jacobson, 237–263. Oxford University Press.

Cardinaletti, Anna. 2004. Towards a cartography of subject positions. In *The structure of CP and IP: The cartography of syntactic structures*, ed. Luigi Rizzi, volume 2, 115–165. Oxford University Press.

Carnie, Andrew. 1995. Non-verbal predication and head-movement. Doctoral Dissertation, Massachusetts Institute Of Technology.

Carnie, Andrew. 1997. Two types of non-verbal predication in Modern Irish. *Canadian Journal of Linguistics* 42:57–74.

Carstens, Vicki. 2003. Rethinking complementizer agreement: Agree with a Case-checked goal. *Linguistic Inquiry* 34:393–412.

Chierchia, Gennaro, and Raymond Turner. 1988. Semantics and property theory. *Linguistics and Philosophy* 11:261–302.

Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in English transformational grammar*, ed. Roderick Jacobs, 184–221. Ginn.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Barriers*. MIT press.

Chomsky, Noam. 1995. *The minimalist program*, volume 28. Cambridge University Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. Jean-Roger Vergnaud, Robert Freidin, Carlos Peregrín Otero, and Maria Luisa Zubizarreta, volume 45, 133–166. MIT Press.

Chung, Sandra. 1991. Sentential subjects and proper government in Chamorro. In *Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda*, ed. C. Georgopoulos and Roberta Ishihara, 75–99. Springer.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford University Press.

Cinque, Guglielmo. 2004. "restructuring" and functional structure. In *Structures and beyond: the cartography of syntactic structures, volume 3*, ed. Adriana Belletti, 132–191. Oxford University Press.

Citko, Barbara. 2008. Missing labels. *Lingua* 118:907–944.

Comorovski, Ileana. 2008. Intensional subjects and indirect contextual anchoring. In *Time and modality*, ed. Jacqueline Guéron and Jacqueline Lecarme, 39–57. Springer.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*, volume 2. Cambridge University Press.

Comrie, Bernard. 1985. *Tense (Cambridge textbooks in linguistics)*. Cambridge University Press.

Condoravdi, Cleo. 2002. Temporal interpretation of modals. In *The construction of meaning*, ed. David Beaver, Stefan Kaufmann, Brady Clark, and Luis Casillas, 59–88. CSLI Publications.

Coon, Jessica. 2013. Predication, tenselessness, and what it takes to be a verb. In *Proceedings of the 43rd Annual Meeting of the North East Linguistics Society*.

Coppock, Elizabeth, Jason Brenier, Laura Staum, and Laura Michaelis. 2006. ISIS: it's not disfluent, but how do we know that? In *Proceedings of the 32nd Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA.

Coppock, Elizabeth, and Laura Staum-Casasanto. 2004. Origin of the English double-is construction. Manuscript, Stanford University.

Cowart, Wayne. 1997. *Experimental syntax*. Sage.

Cowper, Elizabeth. 2010. Auxiliary verb insertion. Handout from ACL/CLA Conference.

Culicover, Peter. 1977. Some observations concerning pseudo-clefts. *Linguistic Analysis* 3:347–375.

Culicover, Peter, and Ray Jackendoff. 1999. The view from the periphery: The English comparative correlative. *Linguistic Inquiry* 30:543–571.

Culicover, Peter, and Ray Jackendoff. 2001. Control is not movement. *Linguistic Inquiry* 32:493–512.

Curzan, Anne. 2012. Revisiting the reduplicative copula with corpus-based evidence. In *The Oxford handbook of the history of English*, 211. Oxford University Press.

Davies, Mark. 2008–. The Corpus of Contemporary American English (COCA): 400+ million words, 1990–present. Online: http://corpus.byu.edu/coca.

Davies, William, and Stanley Dubinsky. 2009. On the existence (and distribution) of sentential subjects. In *Hypothesis A/Hypothesis B: Linguistic explorations in honor of David M. Perlmutter*, ed. Donna B. Gerdts, John C. Moore, and Maria Polinsky, 111–128. Cambridge, MA: MIT Press.

Dayal, Veneeta. 1997. Free relatives and "ever": Identity and free choice readings. In *Proceedings of Semantics and Linguistic Theory*, volume 7, 99–116.

Déchaine, Rose-Marie. 1995. One *be*. In *Linguistics in the Netherlands*, ed. Marcel den Dikken, volume 12, 73–88. John Benjamins Publishing Company.

Declerck, Renaat. 1988. *Studies on copular sentences, clefts and pseudo-clefts*, volume 5. Walter de Gruyter.

DeGraff, Michel. 1992. The syntax of predication in Haitian. In *Proceedings of the 22nd meeting of the North-East Linguistic Society*, 103–117.

Delahunty, Gerald. 1982. *Topics in the syntax and semantics of English cleft sentences*. Indiana University Linguistics Club.

Delahunty, Gerald. 1983. But sentential subjects do exist. *Linguistic Analysis* 12:379–398.

Demirdache, Hamida, and Myriam Uribe-Etxebarria. 2000. The primitives of temporal relations. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. Noam Chomsky, Roger Martin, David Michaels, and Juan Uriagereka, 157–186. Cambridge (Mass.), MIT Press.

Demirdache, Hamida, and Myriam Uribe-Etxebarria. 2007. The syntax of time arguments. *Lingua* 117:330–366.

Demirdache, Hamida, and Myriam Uribe-Etxebarria. 2014. Aspect and temporal anaphora. *Natural Language & Linguistic Theory* 3.

Demonte, Violeta, and Olga Fernández-Soriano. 2009. Force and finiteness in the Spanish complementizer system. *Probus* 21:23–49.

Déprez, Viviane. 2003. Haitian Creole *se*: A copula, a pronoun, both, or neither? On the double life of a functional head. In *Recent development in Creole studies*, ed. Dany Adone, 135–173. Walter de Gruyter.

den Dikken, Marcel. 2005a. A comment on the topic of Topic–Comment. *Lingua* 115:691–710.

den Dikken, Marcel. 2005b. Specificational copular sentences and pseudoclefts. In *The Blackwell companion to syntax*, ed. Martin Everaert and Henk C van Riemsdijk, 292–409. Wiley Online Library. Cited page numbers are from the 2012 revised version.

den Dikken, Marcel. 2006. *Relators and linkers: the syntax of predication, predicate inversion, and copulas*. Cambridge, Mass: MIT Press.

den Dikken, Marcel. 2013. Predication and specification in the syntax of cleft sentences. In *Cleft structures*, ed. Katharina Hartmann and Tonjes Veenstra, 35–70. Amsterdam/Philadelphia: John Benjamins Publishing Company.

den Dikken, Marcel. 2014a. The attractions of agreement. Talk given at the Agreement ϕ Workshop in Honor of Professor Maria Denilda Moura. February 6–7, 2014. Recife, Brazil.

den Dikken, Marcel. 2014b. On feature interpretability and inheritance. In *Minimalism and beyond: Radicalizing the interfaces*, ed. Peter Kosta, Steven Franks, Teodora Radeva-Bork, and Lilia Schürcks, volume 11, 37–55. John Benjamins Publishing Company.

den Dikken, Marcel, André Meinunger, and Chris Wilder. 2000. Pseudoclefts and ellipsis. *Studia Linguistica* 54:41–89.

Doherty, Cathal. 1997. Clauses without complementizers: Finite IP-complementation in English. *Linguistic Review* 14:197–220.

Doherty, Cathal. 2000. *Clauses without "that": The case for bare sentential complementation in English*. Taylor & Francis.

Donati, Caterina, and Carlo Cecchetto. 2011. Relabeling heads: A unified account for relativization structures. *Linguistic Inquiry* 42:519–560.

Doron, Edit. 1986. The pronominal copula as agreement clitic. *Syntax and Semantics* 19:313–332.

Dowty, David. 1982. Tenses, time adverbs, and compositional semantic theory. *Linguistics and Philosophy* 5:23–55.

Dryer, Matthew S. 1980. The positional tendencies of sentential noun phrases in universal grammar. *The Canadian Journal of Linguistics / La Revue Canadienne de Linguistique* 25:123–195.

Ekulona, Tabia Janine. 2000. The BA morpheme in Garifuna. M.A. thesis, Department of Linguistics, UCLA, Los Angeles, CA.

Emonds, Joseph. 1970. Root and structure-preserving transformations. Doctoral Dissertation, Massachusetts Institute of Technology.

Emonds, Joseph. 1976. *A transformational approach to English syntax: Root, structure-preserving, and local transformations*. New York: Academic Press.

Enç, Mürvet. 1987. Anchoring conditions for tense. *Linguistic Inquiry* 633–657.

Erdmann, Peter. 1988. On the principle of 'weight' in English. *On Language* 325–339.

Erlewine, Michael Yoshitaka, and Hadas Kotek. to appear. A streamlined approach to online linguistic surveys. *Natural Language & Linguistic Theory* .

Faraci, Robert. 1971. On the deep question of pseudo-clefts. *English Linguistics* 6:48–85.

Frana, Ilaria. 2010. Copular questions and concealed questions. In *Proceedings of Sinn und Bedeutung 14*, ed. Martin Prinzhorn, Viola Schmitt, and Sarah Zobel. Vienna.

Frazier, Lyn, and Keith Rayner. 1988. Parameterizing the language processing system: Left-vs. right-branching within and across languages. In *Explaining language universals*, ed. John A. Hawkins, 247–279. Blackwell.

Garrett, Andrew. 2012. The historical syntax problem: Reanalysis and directionality. In *Grammatical change: Origins, nature, outcomes*, ed. Diane Jonas, John Whitman, and Andrew Garrett, 52–72. Oxford University Press.

Geist, Ljudmila. 2007. Predication and equation in copular sentences: Russian vs. English. In *Existence: semantics and syntax*, ed. Ileana Comorovski and Klaus Von Heusinger, 79–105. Springer.

Gibson, Edward, Steve Piantadosi, and Kristina Fedorenko. 2011. Using Mechanical Turk to obtain and analyze English acceptability judgments. *Language and Linguistics Compass* 5:509–524.

Giorgi, Alessandra. 2010. *About the speaker: Towards a syntax of indexicality*. Oxford Studies in Theoretical Linguistics. Oxford University Press.

Giorgi, Alessandra, and Fabio Pianesi. 2004. The temporal perspective of the speaker and the subject: from semantics to morphosyntax. In *The syntax of time*, ed. Jacqueline Guéron and Jacqueline Lecarme, 129–152. MIT Press.

Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10:279–326.

Grimshaw, Jane. 1991. Extended projections. Manuscript, Rutgers University.

Grimshaw, Jane. 1997. Projection, heads, and optimality. *Linguistic Inquiry* 28:373–422.

Grimshaw, Jane. 2000. Locality and extended projection. In *Amsterdam Studies in the Theory and History of Linguistic Science Series 4*, 115–134. Amsterdam, John Benjamins.

Groenendijk, Jeroen, and Martin Stokhof. 1984. On the semantics of questions and the pragmatics of answers. *Varieties of Formal Semantics* 3:143–170.

Groenendijk, Jeroen, and Martin Stokhof. 1997. Questions. In *Handbook of logic and language*, ed. Johan van Benthem and Alice Ter Meulen, 1055–1124. Elsevier.

Grohmann, Kleanthes K. 2003. *Prolific domains: On the anti-locality of movement dependencies*, volume 66. John Benjamins Publishing.

Grosu, Alexander, and Sandra Thomson. 1977. Constraints on the distribution of NP clauses. *Language* 53:104–151.

Guéron, Jacqueline. 2007. On tense and aspect. *Lingua* 117:367–391.

Gundel, Jeanette. 1977. Where do cleft sentences come from? *Language* 543–559.

Hacquard, Valentine. 2006. Aspects of modality. Doctoral Dissertation, Massachusetts Institute of Technology.

Hacquard, Valentine. 2009. On the interaction of aspect and modal auxiliaries. *Linguistics and Philosophy* 32:279–315.

Hacquard, Valentine. 2010. On the event relativity of modal auxiliaries. *Natural Language Semantics* 18:79–114.

Haddican, Bill, Anders Holmberg, Hidekazu Tanaka, and George Tsoulas. 2014. Interrogative slifting in English. *Lingua* 138:86–106.

Haeberli, Eric. 1999. Features, categories and the syntax of A-positions. Doctoral Dissertation, Université de Genève.

Haegeman, Liliane. 1992. *Theory and description in generative syntax: A case study in West Flemish*. Cambridge University Press.

Haegeman, Liliane. 2004. Topicalization, CLLD and the left periphery. *ZAS Papers in Linguistics* 35:157–192.

Haegeman, Liliane. 2009. Parenthetical adverbials: The radical orphanage approach. In *Conference on Dislocated Elements in Discourse*, 331–347. Routledge.

Haegeman, Liliane. 2013. The syntax of registers: Diary subject omission and the privilege of the root. *Lingua* 130:88–110.

Haegeman, Liliane, and Marjo van Koppen. 2012. Complementizer agreement and the relation between C^0 and T^0. *Linguistic Inquiry* 43:441–454.

Haider, Hubert. 2010. *The syntax of German*. Cambridge University Press.

Hale, Kenneth. 1986. Notes on world view and semantic categories: some Warlpiri examples. In *Features and projections*, ed. Pieter Muysken and Henk van Riemsdijk, 233–254. Dordrecht: Foris.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces in inflection. In *The view from building 20*, ed. Kenneth Hale and Samuel J. Keyser, 111–176. Cambridge, MA: MIT Press.

Halvorsen, Per-Kristian. 1978. The syntax and semantics of cleft constructions. Doctoral Dissertation, University of Texas at Austin.

Hankamer, Jorge. 1974. On the non-cyclic nature of WH-clefting. In *Proceedings of the Chicago Linguistic Society*, volume 10, 221–286.

Hartmann, Katharina, and Malte Zimmermann. 2012. Focus marking in Bura: semantic uniformity matches syntactic heterogeneity. *Natural Language & Linguistic Theory* 30:1061–1108.

Haspelmath, Martin. 2007. Pre-established categories don't exist: Consequences for language description and typology. *Linguistic Typology* 11:119–132.

Hedberg, Nancy. 1993. On the subject-predicate structure of pseudoclefts. In *Principles and prediction: The analysis of natural language. Papers in honor of Jerry Sanders*, ed. M. Eid and G. Iverson, 119–119. John Benjamins.

Hedberg, Nancy. 2008. Approaching clefts: Syntax, semantics, pragmatics. Handout from the Workshop on Clefts, Zentrum für Allgemeine Sprachwissenschaft (ZAS). Berlin, Germany.

Heggie, Lorie. 1988. The syntax of copular constructions. Doctoral Dissertation, University of Southern California.

Heggie, Lorie, and Eiichi Iwasaki. 2012. Pseudo-inverse copular structures and copular omission in English comparative correlatives. Manuscript (in preparation).

Heim, Irene. 1979. Concealed questions. In *Semantics from different points of view*, ed. Rainer Bäuerle, Urs Egli, and Arnim von Stechow, 51–60. Springer.

Heim, Irene. 1982. The semantics of definite and indefinite noun phrases. Doctoral Dissertation, University of Massachusetts Amherst.

Heim, Irene. 1991. Notes on the first person. Class notes.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*, volume 13. Blackwell Oxford.

Henry, Alison. 2005. Non-standard dialects and linguistic data. *Lingua* 115:1599–1617.

Heringa, Herman. 2012. A multidominance approach to appositional constructions. *Lingua* 122:554–581.

Heycock, Caroline. 1991. Layers of predication: The non-lexical syntax of clauses. Doctoral Dissertation, University of Pennsylvania.

Heycock, Caroline. 1994. The internal structure of small clauses: New evidence from inversion. In *Proceedings of NELS 25*, 223–238. Amherst, MA: GLSA.

Heycock, Caroline. 1995. Asymmetries in reconstruction. *Linguistic Inquiry* 26:547–570.

Heycock, Caroline. 2006. Embedded root phenomena. In *The Blackwell companion to syntax*, ed. Martin Everaert and Henk C van Riemsdijk, 174–209. Wiley Online Library.

Heycock, Caroline. 2012. Specification, equation, and agreement in copular sentences. *The Canadian Journal of Linguistics/La revue canadienne de linguistique* 57:209–240.

Heycock, Caroline, and Anthony Kroch. 1999. Pseudocleft connectedness: Implications for the LF interface level. *Linguistic inquiry* 30:365–397.

Higgins, Francis. 1979. *The pseudo-cleft construction in English*. Garland, New York.

Hooper, Joan, and Sandra Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 465–497.

Horn, Laurence. 1985. Metalinguistic negation and pragmatic ambiguity. *Language* 121–174.

Horn, Laurence, Raffaella Zanuttini, and Jim Wood. 2015. The syntax of speech acts: Evidence from Southern American English dialects.

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

Huddleston, Rodney. 1971. *The sentence in written English: A syntactic study based on an analysis of scientific texts*. Cambridge University Press.

Iatridou, Sabine. 2000. The grammatical ingredients of counterfactuality. *Linguistic Inquiry* 31:231–270.

Jacobson, Pauline. 1994. Binding connectivity in copular sentences. In *Proceedings of Semantics and Linguistic Theory*, volume 4, 161–178.

Jayaseelan, KA. 2013. Coordination, relativization and finiteness in dravidian. *Natural Language & Linguistic Theory* 32:191–211.

Jehn, Richard. 1979. That's something that i wouldn?t want to account for, is a sentence like this. *Calgary Working Papers in Linguistics* 5:51–62.

Johnson, Kyle. 2009. Gapping is not (VP-) ellipsis. *Linguistic Inquiry* 40:289–328.

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

Kayne, Richard. 1994. *The antisymmetry of syntax*. 25. MIT Press.

Kayne, Richard. 2015. The silence of heads. Manuscript, New York University.

Kayne, Richard S. 1981. ECP extensions. *Linguistic Inquiry* 93–133.

Kayne, Richard S. 2005. *Movement and silence*, volume 36. Oxford University Press.

Keller, Frank. 2000. Gradience in grammar: Experimental and computational aspects of degrees of grammaticality. Doctoral Dissertation, University of Edinburgh.

Khanjian, Hrayr. 2013. (Negative) concord and head directionality in Western Armenian. Doctoral Dissertation, Massachusetts Institute of Technology.

Kim, Kitaek. 2011. Overgenerated *be* from topic marker to verbal inflection. In *Selected proceedings of the 2009 Second Language Research Forum*, ed. Luke Plonsky and Maren Schierloh, 70–81. Somerville, MA: Cascadilla Proceedings Project.

Kissock, Madelyn J. 2013. Evidence for "finiteness" in Telugu. *Natural Language & Linguistic Theory* 32:1–30.

Klein, Wolfgang. 1994. *Time in language*. London: Routledge.

Klein, Wolfgang. 1998. Assertion and finiteness. In *Issues in the theory of language acquisition: Essays in honor of Jürgen Weissenborn*, ed. Norbert Dittmar and Zvi Penner, 225–245. Peter Lang Publishers, Inc.

Klein, Wolfgang. 2014. Is aspect time-relational? *Natural Language and Linguistic Theory* 32. Special issue on Aspect Across Languages, ed. María Arche.

Kluck, Marlies. 2011. *Sentence amalgamation*, volume 285 of LOT Dissertation Series. Utrecht: LOT.

Koppen, Marjo van. 2005. *One probe-two goals: Aspects of agreement in Dutch dialects*. Dissertation, University of Leiden. LOT: Netherlands Graduate School of Linguistics.

Koster, Jan. 1978. Why subject sentences don't exist. In *Recent transformational studies in Euro-*

pean languages, ed. Samuel Jay Keyser, 53–64. Cambridge, MA: MIT Press.

Koster, Jan. 2000. Extraposition as parallel construal. Manuscript, University of Groningen.

Kratzer, Angelika. 1998. More structural analogies between pronouns and tenses. In *Proceedings of SALT*, volume 8, 92–110.

Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40:187–237.

Krifka, Manfred. 2014. Embedding illocutionary acts. In *Recursion: Complexity in cognition*, 59–87. Springer.

Labov, William. 1995. The case of the missing copula: The interpretation of zeroes in African-Amerircan English. *Language* 1:25–54.

Lakoff, George. 1970. *Irregularity in syntax*. Holt, Rinehart and Winston New York.

Lakoff, George. 1974. Syntactic amalgams. In *Papers from the Tenth Regional Meeting of the Chicago Linguistic Society*, 321–344.

Lambrecht, Knud. 2001. A framework for the analysis of cleft constructions. *Linguistics* 39:463–516.

Lambrecht, Knud, and Sebastian Ross-Hagebaum. 2006. Apo koinou and intrusion: Toward a syntactic-pragmatic typology of amalgam constructions in spoken English. In *Proceedings of the 4th workshop on discourse structure. University of Texas, Austin.*

Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language & Linguistic Theory* 22:811–877.

Landau, Idan. 2007. EPP extensions. *Linguistic Inquiry* 38:485–483.

Larson, Richard. 1987. "Missing prepositions" and the analysis of English free relative clauses. *Linguistic Inquiry* 18:239–266.

Lasnik, Howard, and Mamoru Saito. 1991. On the subject of infinitives. In *27th regional meeting of the Chicago Linguistic Society*, volume 27, 324–343.

Lecarme, Jacqueline. 2004. Tense in nominals. In *The syntax of time*, ed. Jacqueline Guéron and Jacqueline Lecarme, 440–475. MIT Press.

Levin, Juliette, and Diane Massam. 1986. Classification of Niuean verbs: notes on case. In *FOCAL I: Papers from the fourth International Conference on Austronesian Linguistics*, 231–244.

Lin, Jo-Wang. 2010. A tenseless analysis of Mandarin Chinese revisited: A response to Sybesma 2007. *Linguistic Inquiry* 41:305–329.

Lohndal, Terje. 2013. Sentential subjects: topics or real subjects? In *Proceedings of the 31st West Coast Conference on Formal Linguistics 31*.

Lohnstein, Horst. 2015. Verum focus. In *The Oxford handbook of information structure*, ed. Caroline Féry and Shinichiro Ishihara. Oxford University Press.

Maienborn, Claudia. 2005. On the limits of the Davidsonian approach: The case of copula sentences. *Theoretical linguistics* 31:275–316.

Massam, Diane. 1999. *Thing is* constructions: the thing is, is what's the right analysis? *English Language and Linguistics* 3:335–352.

Massam, Diane. 2013. Intrusive *be* constructions in (spoken) English: Apposition and beyond. In *Proceedings of the 2012 annual conference of the Canadian Linguistic Association*.

Matthewson, Lisa. 2004. On the methodology of semantic fieldwork 1. *International Journal of American Linguistics* 70:369–415.

Matthewson, Lisa. 2006. Temporal semantics in a superficially tenseless language. *Linguistics and Philosophy* 29:673–713.

McCloskey, James. 1991. *There, it*, and agreement. *Linguistic Inquiry* 22:563–567.

McCloskey, James. 1996. On the scope of verb movement in Irish. *Natural Language & Linguistic Theory* 14:47–104.

McConvell, Patrick. 1988. To be or double be? Current changes in the English copula. *Australian Journal of Linguistics* 8:287–305.

McConvell, Patrick. 2004. Catastrophic change in current English: Emergent double-be's and free-be's. Talk given at the CRLC Seminar at the Australian National University.

McFadden, Thomas. 2013. On subject reference and the cartography of clause types. *Natural Language & Linguistic Theory* 32:115–135.

McFadden, Thomas, and Sandhya Sundaresan. 2013. Finiteness in south asian languages: an introduction. *Natural Language & Linguistic Theory* 32:1–27.

McWhorter, John. 2012. From focus marker to copula in Swahili. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, volume 20.

Meinunger, André. 2004. Verb position, verbal mood and the anchoring (potential) of sentences. In *The syntax and semantics of the left periphery*, ed. Horst Lohnstein and Susanne Trissler, 313–341. Mouton de Gruyter.

Meinunger, André. 2006. The discourse status of subordinate sentences and some implications for syntax and pragmatics. In *Architecture of focus*, ed. Valéria Molnár and Susanne Winkler,

459–487. Mouton de Gruyter.

Mikkelsen, Line. 2004b. Specifying who: On the structure, meaning, use of specificational copula sentences. Doctoral Dissertation, University of Santa Cruz.

Mikkelsen, Line. 2005. *Copular clauses: Specification, predication and equation*, volume 85. John Benjamins Publishing.

Mikkelsen, Line. 2011. Copular clauses. In *Semantics: An international handbook of natural language meaning*, ed. Claudia Maienborn, Klaus von Heusinger, and Paul Portner, volume 2, 1805–1829. Mouton de Gruyter.

Moro, Andrea. 1997. *The raising of predicates: Predicative noun phrases and the theory of clause structure*. Cambridge University Press.

Moro, Andrea. 2000. *Dynamic antisymmetry*. 38. MIT Press.

Moulton, Keir. 2013. Not moving clauses: Connectivity in clausal arguments. *Syntax* 16:250–291.

Nathan, Lance Edward. 2006. On the interpretation of concealed questions. Doctoral Dissertation, Massachusetts Institute of Technology.

Nikolaeva, Irina. 2007. *Finiteness: Theoretical and empirical foundations*. Oxford University Press.

Ogihara, Toshiyuki. 1996. *Tense, attitudes, and scope*. Kluwer Dordrecht.

Ogihara, Toshiyuki, and Yael Sharvit. 2012. Embedded tenses. In *The Oxford handbook of tense and aspect*, ed. Robert Binnick, 638–68. Oxford University Press, Oxford and New York.

O'Neill, Teresa. 2012. Coordination in English copular amalgams. Qualifying Paper, CUNY Graduate Center.

O'Neill, Teresa. to appear. Demystifying double-*is*. In *University of Pennsylvania working papers in linguistics: Proceedings of the 38th Penn Linguistics Colloquium*, volume 21.1.

Ott, Dennis, and Mark de Vries. to appear. Right-dislocation as deletion. *Natural Language & Linguistic Theory* .

Partee, Barbara. 1977. 'John is easy to please'. In *Linguistic structures processing*, ed. Antonio Zampolli, volume 5, 281–312. North-Holland.

Partee, Barbara. 1986. Ambiguous pseudoclefts with unambiguous *be*. In *Proceedings of the North East Linguistic Society*, volume 16, 354–366.

Partee, Barbara. 1987. Noun phrase interpretation and type-shifting principles. In *Studies in discourse representation theory and the theory of generalized quantifiers*, ed. Jeroen Groenendijk,

Dick de Jongh, and Martin Stokhof, volume 8, 115–143. Dordrecht: Foris.

Partee, Barbara. 2000. Copula inversion puzzles in English and Russian. *Issues in Semantics* 23:183–208.

Partee, Barbara Hall. 1973. Some structural analogies between tenses and pronouns in English. *The Journal of Philosophy* 601–609.

Pesetsky, David. 1982. Paths and categories. Doctoral Dissertation, Massachusetts Institute of Technology.

Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. *Current Studies in Linguistics Series* 36:355–426.

Pesetsky, David, and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. *Phrasal and clausal architecture: Syntactic derivation and interpretation* 262–294.

Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 365–424.

Prior, Arthur N. 1957. *Time and modality*. Oxford University Press.

Prior, Arthur N. 1967. *Past, present and future*, volume 154. Clarendon Press Oxford.

Prior, Arthur N. 1969. *Papers on time and tense*. Clarendon Press Oxford.

Pustet, Regina. 2003. *Copulas: Universals in the categorization of the lexicon*. Oxford University Press.

R Core Team. 2013. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL `http://www.R-project.org`.

Rapoport, Tova. 1987. Copular, nominal, and small clauses: A study of Israeli Hebrew. Doctoral Dissertation, Massachusetts Institute of Technology.

Raposo, Eduardo. 1987. Case theory and Infl-to-Comp: The inflected infinitive in European Portuguese. *Linguistic Inquiry* 85–109.

Reichenbach, Hans. 1947. *Elements of symbolic logic*. London: Macmillan.

Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. Croom Helm London.

Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657–720.

Richards, Marc D. 2007. On feature inheritance: An argument from the Phase Impenetrability Condition. *Linguistic Inquiry* 38:563–572.

van Riemsdijk, Henk. 2006. Grafts follow from Merge. In *Phases of interpretation*, ed. Mara

Frascarelli, volume 91, 17–44. The Hague: Mouton de Gruyter.

Ritter, Elizabeth, and Martina Wiltschko. 2005. Anchoring events to utterances without tense. In *Proceedings of WCCFL*, volume 24, 343–351.

Ritter, Elizabeth, and Martina Wiltschko. 2009. Varieties of infl: Tense, location, and person. *Alternatives to cartography* 153–202.

Ritter, Elizabeth, and Martina Wiltschko. 2014. The composition of INFL: An exploration of tense, tenseless languages and tenseless constructions. *Natural language & Linguistic Theory* 32:1331–1386.

Rizzi, Luigi. 1990. *Relativized minimality.*. The MIT Press.

Rizzi, Luigi. 1994. Early null subjects and root null subjects. In *Language acquisition studies in generative grammar*, ed. Teun Hoekstra and Bonnie Schwartz, 151–176. John Benjamins Amsterdam.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. Liliane Haegeman, Kluwer International Handbooks of Linguistics, 281–337. Springer Netherlands.

Rizzi, Luigi, ed. 2004. *The structure of CP and IP: The cartography of syntactic structures*, volume 2. Oxford University Press.

Rizzi, Luigi. 2006. On the form of chains: Criterial positions and ECP effects. In *On* wh-*movement: moving on*, ed. Lisa Lai Shen Cheng and Norbert Corver. MIT Press.

Rizzi, Luigi. 2010. On some properties of criterial freezing. In *The complementizer phase: Subjects and operators*, ed. E. Phoevos Panagiotidis. Oxford Scholarship Online.

Rizzi, Luigi, and Ur Shlonsky. 2007. Strategies of subject extraction. In *Interfaces + recursion= language?: Chomsky's minimalism and the view from syntax-semantics*, ed. Uli Sauerland and Hans-Martin Gärtner, volume 89, 115–160. Walter de Gruyter.

Romero, Maribel. 2004. Tense and intensionality in specificational copular sentences. In *Proceedings of Semantics and Linguistic Theory*, volume 14, 271–288.

Romero, Maribel. 2005. Concealed questions and specificational subjects. *Linguistics and Philosophy* 28:687–737.

Romero, Maribel. 2006. On concealed questions. In *Proceedings of Semantics and Linguistic Theory*, ed. Masayuki Gibson and Jonathan Howell, volume 16, 208–227.

Rooryck, Johan. 2000. *Configurations of sentential complementation: perspectives from Romance languages*. Psychology Press.

Rooth, Mats. 1985. Association with focus. Doctoral Dissertation, University of Massachusetts at

Amherst.

Rosenbaum, Peter. 1967. *The grammar of English predicate complement constructions*, volume Research monograph 47. Cambridge, MA: MIT Press.

Ross, John R. 1967. Constraints on variables in syntax. Doctoral Dissertation, Massachusetts Institute of Technology.

Ross, John R. 1999. That is the question. Talk given at Humboldt University, Berlin.

Ross, John R. 2000. The frozenness of pseudoclefts—towards an inequality-based syntax. Manuscript, University of North Texas.

Ross, John R. (Haj). 2004. The syntax of emphasis–a base camp. In *Lexique, syntaxe et lexique-grammaire (syntax, lexis & lexicon-grammar): Hommage à Maurice Gross*, 547–559. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Ross, John Robert. 1972. Act. In *Semantics of natural language*, ed. Donald Davidson and Gilbert Harman, 70–126. Reidel Dordrecht.

Ross-Hagebaum, Sebastian. 2004. The *that's X is Y* construction as an information-structure amalgam. In *Paper presented at the 30th meeting of the Berkley Linguistics Society*, 403–414.

Rothstein, Susan. 1995. Small clauses and copular constructions. In *Small clauses*, ed. Anna Cardinaletti and Maria Teresa Guasti, 27–48. San Diego, CA: Academic Press.

Rothstein, Susan. 1999. Fine-grained structure in the eventuality domain: The semantics of predicative adjective phrases and *be*. *Natural Language Semantics* 7:347–420.

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

Russell, Bertrand. 1919. *Introduction to mathematical philosophy*. London: Allen and Unwin. [1993 reprint from Courier Corporation].

Safir, Ken. 1983. On small clauses as constituents. *Linguistic Inquiry* 730–735.

Schlenker, Philippe. 1999. Propositional attitudes and indexicality. Doctoral Dissertation, MIT.

Schlenker, Philippe. 2003. Clausal equations (a note on the connectivity problem). *Natural Language & Linguistic Theory* 21:157–214.

Schlenker, Philippe. 2003b. A plea for monsters. *Linguistics and Philosophy* 26:29–120.

Schlenker, Philippe. 2004. Sequence phenomena and double access readings generalized. In *The syntax of time*, ed. Jacqueline Guéron and Jacqueline Lecarme. MIT Press.

Schütze, Carson. 2004. When is a verb not a verb? *Nordlyd* 31.

Schütze, Carson, and Jon Sprouse. 2014. Judgment data. In *Research methods in linguistics*, ed. Devyani Sharma and Robert J. Podesva, 27–50. Cambridge University Press.

Schütze, Carson T. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. University of Chicago Press.

Schütze, Carson T. 2004b. Why nonfinite *be* is not omitted while finite *be* is. In *Proceedings of the 28th Boston University Conference on Language Development*, 506–521.

Schütze, Carson T. 2009. Web searches should supplement judgements, not supplant them. *Zeitschrift für Sprachwissenschaft* 28:151–156.

Schütze, Carson Theodore Robert. 1997. Infl in child and adult language: Agreement, case and licensing. Doctoral Dissertation, Massachusetts Institute of Technology.

Selkirk, Elizabeth. 1995. Sentence prosody: Intonation, stress, and phrasing. In *The handbook of phonological theory*, ed. John Goldsmith, 550–569. Oxford: Blackwell.

Shaer, Benjamin. 2003. Toward the tenseless analysis of a tenseless language. *Proceedings of SULA* 2:139–156.

Sharvit, Yael. 1999. Connectivity in specificational sentences. *Natural Language Semantics* 7:299–339.

Sharvit, Yael. 2003. Tense and identity in copular constructions. *Natural Language Semantics* 11:363–393.

Sharvit, Yael. 2008. The puzzle of free indirect discourse. *Linguistics and Philosophy* 31:353–395.

Sharvit, Yael. 2009. Revisiting connectivity in copular sentences. In *Proceedings of Semantics and Linguistic Theory*, volume 19, 427–445.

Shlonsky, Ur. 2010. The cartographic enterprise in syntax. *Language and Linguistics Compass* 4:417–429.

Sigurðsson, Halldór Ármann. 2004. The syntax of person, tense, and speech features. *Italian Journal of Linguistics* 16:219–251.

Sprouse, Jon. 2011a. A test of the cognitive assumptions of magnitude estimation: Commutativity does not hold for acceptability judgments. *Language* 87:274–288.

Sprouse, Jon. 2011b. A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods* 43:155–167.

Sprouse, Jon, and Diogo Almeida. 2010. A quantitative defense of linguistic methodology. Manuscript.

von Stechow, Arnim. 1995. On the proper treatment of tense. *Proceedings of Semantics and Linguistic Theory V* 362–386.

von Stechow, Arnim. 2002. Binding by verbs: Tense, person and mood under attitudes. In *Proceedings of the North East Linguistic Society*, volume 33, 379–403.

Stowell, Tim. 1981a. Complementizers and the empty category principle. In *Proceedings of the Eleventh Annual Meeting of the North East Linguistic Society*, ed. V. Burke and J. Pustejovsky, 345–353.

Stowell, Tim. 1982. The tense of infinitives. *Linguistic Inquiry* 13:561–570.

Stowell, Tim. 1996. The phrase structure of tense. In *Phrase structure and the lexicon*, 277–291. Springer.

Stowell, Tim. 2007. The syntactic expression of tense. *Lingua* 117:437–463.

Stowell, Timothy. 1981b. Origins of phrase structure. Doctoral Dissertation, Massachusetts Institute of Technology.

Stowell, Timothy. 1991. Empty heads in abbreviated English. Ms., UCLA. (Paper presented at GLOW 1991.).

Suazo, Salvador. 2002. *Conversemos en garífuna: Gramática y manual de conversación*. Editorial Guaymuras.

Szabolcsi, Anna. 1987. Functional categories in the noun phrase. In *Approaches to Hungarian volume 2: Theories and analyses*, ed. Kenesei István. Szeged: Jate Szeged.

Szabolcsi, Anna. 2009. Overt nominative subjects in infinitival complements cross-linguistically: Data, diagnostics, and preliminary analyses. In *New York University Working Papers in Linguistics*, ed. Patricia Irwin and Violeta Vázquez Rojas Maldonado, volume 2, 1–55.

Takahashi, Shoichi. 2010. The hidden side of clausal complements. *Natural Language & Linguistic Theory* 28:343–380.

Thráinsson, Höskuldur. 2007. *The syntax of icelandic*. Cambridge University Press.

Tonhauser, Judith. 2011. Temporal reference in Paraguayan Guaraní, a tenseless language. *Linguistics and Philosophy* 34:257–303.

Tortora, Christina. 2014. *A comparative grammar of Borgomanerese*. Oxford University Press.

Trutkowski, Ewa. 2010. Referential null subjects in German. In *Proceedings of the Sixth Cambridge Postgraduate Conference in Language Research*, ed. Chris Cummins, Chi-Hé Elder, Thomas Godard, Morgan Macleod, Elaine Schmidt, and George Walkden, 206–217. University of Cambridge.

Turk, Amazon Mechanical. 2011. Requester best practices guide. Amazon Web Services.

Verkuyl, Henk. 1993. *A theory of aspect*. Cambridge: Cambridge University Press.

de Vries, Mark. 2006. The syntax of appositive relativization: On specifying coordination, false free relatives, and promotion. *Linguistic Inquiry* 37:229–270.

de Vries, Mark. 2009. Specifying coordination: An investigation into the syntax of dislocation, extraposition and parenthesis. *Language and linguistics: Emerging trends* 37–98.

Webelhuth, Gert. 1992. *Principles and parameters of syntactic saturation*. Oxford University Press.

Widmann, Cintia. 2005. Factors at play in determining the acceptability of sentential subjects in English: The role of constituent relative weight. Ms., University of South Carolina.

Williams, Edwin. 1983. Semantic vs. syntactic categories. *Linguistics and Philosophy* 6:423–446.

Wiltschko, Martina. 2003. On the interpretability of tense on D and its consequences for case theory. *Lingua* 113:659–696.

Wiltschko, Martina. 2011. Nominal licensing via case or deictic anchoring. In *Proceedings of the 2011 annual conference of the Canadian Linguistic Association*, ed. Lisa Armstrong.

Wiltschko, Martina. 2014. *The universal structure of categories: Towards a formal typology*, volume 142. Cambridge University Press.

Zagona, Karen. 1990. Times as temporal argument structure. Manuscript, University of Washington.

Zagona, Karen. 2003. Tense and anaphora: Is there a tense-specific theory of coreference? In *Anaphora: A reference guide*, ed. Andrew Barss, 140–171. Blackwell Publishing Oxford.

Zeijlstra, Hedde. 2004. *Sentential negation and negative concord*. Utrecht: LOT.

Zimmer, Ben. 2011. The elusive triple "is". Language Log post, September 25, 2011. Online: `http://languagelog.ldc.upenn.edu/nll/?p=3452`.

Zimmer, Ben. 2012. Obama's "is is". Language Log post, October 23, 2012. Online: `http://languagelog.ldc.upenn.edu/nll/?p=4269`.

Zwart, Jan-Wouter. 1993. Verb movement and complementizer agreement. In *MIT working papers in linguistics*, volume 18, 281–296. Department of Linguistics and Philosophy, Massachusetts Institute of Technology.