

Abstract
Information Structure and the Licensing of English Subjects
Jennifer Elaine Mack
2010

Most approaches to argument realization in English are grounded in lexical semantic structure. While it is widely acknowledged that there is an intimate relationship between information structure and grammatical relations such as *subject*, there have been few attempts to formalize this observation. This dissertation proposes an *interface model of argument realization* in which information structure and lexical semantics jointly determine argument realization.

The model proposes two mechanisms through which information structure drives argument realization. In *direct licensing*, informational relations such as *topic* underlie the licensing of arguments. Though this is widespread in “topic-prominent” languages such as Mandarin, it is generally taken to be forbidden in “subject-prominent” languages such as English (Li and Thompson 1976). This dissertation demonstrates that direct licensing by information structure underlies a range of subject selection phenomena in English, and thus that languages fall along a continuum with respect to the availability of direct licensing. In the constructions that I investigate in depth, Topical Exclamatives and Copy Raising, the main-clause subject is licensed to function as a topic. This is formalized through a construction that changes the valence and the information structure requirements of the main predicate.

The second mechanism, *resolution*, has not been discussed in previous work. In resolution, information structure selects among two or more candidates for argument realization that satisfy the constraints of the lexical semantic system. This too is evident in English subject selection. I demonstrate that in the Instrument Subject construction, information structure resolves underspecified input from the lexical semantic linking system. In order to be realized as a subject, an instrument must be associated with a proposition that is activated in the discourse. This is implemented by a construction that pairs the linking of an instrument subject with specific informational constraints.

The constructions that underlie Topical Exclamatives, Copy Raising, and Instrument Subjects can be seen as two concrete components of the often-assumed, but sometimes nebulous-seeming link between subjecthood and information structure in English.

Information Structure and the Licensing of English Subjects

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Jennifer Elaine Mack

Dissertation Director: Laurence Robert Horn

May 2010

© 2010 by Jennifer Elaine Mack
All rights reserved.

Table of Contents

Abstract	i
List of Figures and Tables	ix
Acknowledgements	x
Chapter 1. An Interface Model of Argument Realization	1
1. Introduction	1
2. Two Components of an Interface Model	2
2.1. Direct Licensing	3
2.2. Resolution	5
3. Structure of the Dissertation	8
Chapter 2. Subject and Topic, Revisited	9
1. Introduction	9
2. Subjects	10
2.1. <i>Subject</i> as a Generalization	12
2.2. Special Classes of Subjects	15
3. Argument Realization	17
3.1. Semantic Models of Argument Realization	20
3.1.1. Semantic Determinants of Argument Realization	21
3.1.2. Linking Semantics to Syntax	24
3.1.3. Constructions and Lexical Representations	26
3.1.4. The Present Approach	27
3.1.5. Dowty (1991) on Perspective-Dependent Roles	28
3.2. Informational Approaches to Argument Realization	34
3.2.1. Information Structure and the Selection of Subjects	35
3.2.2. Informational Effects on Object Selection	40
3.2.3. Free Datives	41
3.2.4. Null Instantiation	42
3.2.5. Previous Models of the Interface	43
4. Elements of Information Structure	46
4.1. Topic	48
4.1.1. Identifying Topics	50
4.1.2. Topic, Predication, and Frame-Setting	59
4.1.3. Topicality and Truth Conditions	62
4.2. Focus	64
4.2.1. Identifying Focus	66
4.2.2. Truth-Conditional Effects of Focus	69
4.2.3. Contrast and Exhaustivity	71
4.3. Informational Relations and Grammatical Constructions	75
5. A Brief Overview of Sign-Based Construction Grammar	78
5.1. Signs	79
5.2. Constructions	82

5.3.	Subjects in SBCG	84
5.4.	Direct Licensing	85
5.5.	Resolution	86
6.	Conclusion	87
Chapter 3. Direct Licensing I: Topical Exclamatives		88
1.	Introduction	88
2.	Description of TE	91
2.1.	Usage	91
2.2.	TE, Right Dislocation, and Nominal Extraposition	93
2.3.	Constraints on the Predicate	95
2.4.	Properties of the Extraposed Complement	96
3.	Evidence for the Direct Licensing Account	98
3.1.	Pragmatic Type Restrictions	99
3.2.	The Relevance Requirement	104
3.3.	Information Status	106
3.4.	Topicality and Semantic Role	108
3.5.	Topic-Marking Constructions	108
3.6.	Contrastive Topic Construction	110
3.7.	The Exclamativity Requirement	111
4.	Direct Licensing	118
5.	Alternative Semantic Accounts	122
5.1.	Cause and Object of Emotion	123
5.2.	Property Factoring, Qualification, and Facets	126
6.	TE and the Topic-Comment Relation	130
6.1.	Foundations of the Topic-Comment Link	131
6.2.	TE Subjects as Subjective Topics	138
7.	NP Predicates and TE	141
7.1.	“Impression” of Predication vs. Actual Predication	142
7.2.	Syntax and Semantics of the Post-Verbal Complement	142
7.3.	Exclamativity	144
8.	Conclusion	144
Chapter 4. Direct Licensing II: Copy Raising		146
1.	Introduction	146
2.	Previous Accounts of CR	149
2.1.	CR as Raising	150
2.1.1.	Evidence for a Raising Analysis	151
2.1.2.	Problems for a Raising Analysis	152
2.1.3.	The Significance of (Non)-Synonymy	155
2.1.4.	SSR and Information Structure	158
2.2.	CR Subjects as Arguments of Sentential Predicates	160
2.3.	CR Subjects as Perceptual Sources	161
2.4.	Two Routes to Licensing	163

3.	Semantics and Argument Realization in CR	165
3.1.	Epistemic and Perceptual Resemblance Verbs	167
3.2.	Specifying the Evidential Source	168
3.3.	Epistemic Strengthening	172
4.	CR Subjects as Topics	177
4.1.	Pragmatic Type Restrictions	178
4.1.1.	Bare NPs	178
4.1.2.	Indefinites and NPs with Numeric Quantifiers	180
4.1.3.	Copular Clauses	182
4.1.4.	Violations of Pragmatic Type Restrictions	183
4.2.	Intra-Sentential Topic Continuity	185
4.3.	Topic-Marking Constructions	186
4.4.	CR Subjects as Subjective Topics	186
4.5.	Topicality, Evidential Sources, and Epistemic Strengthening	189
5.	The Second Route to Licensing	191
6.	Conclusion	195
Chapter 5. Resolution: Instrument Subjects		196
1.	Introduction	196
2.	Previous Analyses	202
2.1.	Semantics of Instruments	203
2.1.1.	Intermediary and Facilitating Instruments	203
2.1.2.	Semantic Roles Assigned to Instruments	207
2.2.	Semantics of Predicates	213
2.3.	The Effects of Pragmatics	216
3.	Lexical Semantic Constraints on Instrument Subjects	218
3.1.	Delimiting the Class of Predicates	219
3.1.1.	Requirement 1: Two or More Participants	221
3.1.2.	Requirement 2: Affectedness	222
3.1.3.	Requirement 3: Physical Domain	223
3.1.4.	Representing Physical Affectedness	224
3.1.5.	The Lexicon as a Multiple Inheritance Hierarchy	229
3.2.	Action Chains and the Semantics of Instrument Actors	235
3.2.1.	Representing the Action Chain	236
3.2.2.	Constraints on Actors	237
3.2.3.	Instruments, Agents, and Selection	238
3.3.	Linking Actors as Subjects	242
4.	Pragmatic Constraints on Instrument Subjects	244
4.1.	Instrument Subjects and Activated Propositions	245
4.1.1.	Pragmatically Marked Constructions	246
4.1.2.	Overt Questions	249
4.1.3.	Conversational Context and Implicit Questions	249
4.2.	Formalizing the Effects of Pragmatics	252
4.3.	Discourse Context and the Semantics of the Action Chain	254

5.	Conclusion	258
Chapter 6. Conclusions and Future Directions		259
1.	Introduction	259
2.	Linking Information Structure to Argument Realization	261
3.	A Typological Perspective	264
4.	Constraining the Interface Model	268
5.	Extending the Interface Model	269
6.	Conclusion	271
Bibliography		273

List of Figures and Tables

Chapter 2. An Interface Model of Argument Realization

Figure 1. A Nominal Sign in SBCG	81
Figure 2. A Verbal Sign in SBCG	81
Figure 3. The Subject-Predicate Construction	83

Chapter 3. Direct Licensing I: Topical Exclamatives

Figure 1. The Topic Licensing Construction: Preliminary Version	119
Figure 2. <i>It</i> -Extraposition with Sentential Complements	120
Figure 3. The Topic Licensing Construction: Final Version	140

Table 1. Distribution of Lexical and Pronominal Subjects in TE and NE	108
---	-----

Table 2. Semantic/Pragmatic Foundations of the Topic-Comment Link	134
---	-----

Chapter 4. Direct Licensing II: Copy Raising

Figure 1. Topic Licensing: Final Version	187
Figure 2. Subject-to-Subject Copy Raising	192

Chapter 5. Resolution: Instrument Subjects

Figure 1. Partial Hierarchy of Lexemes, Organized by Event Schema	230
Figure 2. Partial Hierarchy of Lexemes, Organized by Event Schema and Ontological	

Domain	233
--------------	-----

Figure 3. ACT-Trans-P-lxm: Inheritance of Semantics	236
---	-----

Figure 4. ACT-Trans-P-lxm: Inheritance of Argument Linking	242
--	-----

Figure 5. The Non-Initial Actor-Subject Construction	252
--	-----

Table 1. Classes of Verbs Entailing Caused Physical Affectedness	220
--	-----

Acknowledgements

I would like to express my gratitude to the people who have supported me personally and intellectually while I was in the process of writing this dissertation. First, I would like to thank my dissertation committee members, Larry Horn, Ashwini Deo, and Laura Michaelis-Cummings. Larry Horn has been a constant source of guidance and inspiration over the last few years. I am especially appreciative of his generosity with his time and energy, as well as his creative and entertaining seminars, starting with the Spring 2006 seminar on information structure where this dissertation was born, and ending with a musical medley of free datives in the Fall 2009 seminar on conventional implicature. Ashwini Deo helped me to shape the direction of the dissertation in its early stages, and provided me with encouragement when I most needed it. I am very grateful that Laura Michaelis-Cummings was willing to come on board as a reader. I always looked forward to her thoughtful feedback, which helped me to see which of my arguments were working and which were not.

Yale University provided a wonderful environment in which to write a dissertation. I appreciate all that the Yale administration and faculty have done to support us professionally. I would like to thank the many faculty members that I have had the opportunity to learn from and work with at Yale, especially Maria Mercedes Piñango. I am lucky to have begun learning about linguistics as an undergrad at Rice University, and would like to express my gratitude to my professors there, especially Robert Englebretson.

I would like to thank my friends for helping me through the difficult times and helping me create and enjoy the far more numerous good ones, especially Raj Dhillon, Eva Wittenberg, Scott McClure, Madeleine Delpha McClure, Erich Round, Mike Proctor, Julia Kuznetsova, Jenny Yang, Argyro Katsika, Mike Freedman, and Matthew Farina.

I am also extremely grateful to have had the support of my family throughout this process, especially my parents, Dave and Kathi Mack, my brother, Steve Mack, my grandparents, Larry and Pat DeSutter, and my parents- and brothers-in-law, Eyal, Dorly, Gilad, Alon, and Yoav Fuerst.

Finally, I want to thank Yael Fuerst for her love, support, friendship, and ideas. Not to mention her shared enthusiasm for topicalizing every sentence about beans.

*Jennifer Mack
New Haven, Connecticut
December 2009*

Chapter 1

An Interface Model of Argument Realization

1. Introduction

As Du Bois (2003: 34) points out, “linguists have always been aware that certain pragmatic factors (for example topicality) played some kind of role relative to aspects of argument structure (for example subject selection), but they have never quite known what to do with it.” Although some researchers have claimed that there is a rich and pervasive relationship between pragmatics and argument realization (Lambrecht 1995, Aissen 1999, Goldberg 2004), most models of argument realization are exclusively semantic (Jackendoff 1990, Dowty 1991, Croft 1994, 1998, Rappaport Hovav and Levin 1998, Levin and Rappaport Hovav 2005, Davis 2001). This dissertation presents a formal model in which both lexical semantics and information structure underlie argument realization.

I define *argument realization* as the system that selects the core grammatical constituents of the clause, such as subjects and objects.¹ Clearly, one of the central determinants of argument realization is the lexical semantic representation of the predicate. I propose that information structure also plays a crucial role.

¹ In the theoretical framework that I assume, Sign-Based Construction Grammar (SBCG; Michaelis 2009, Sag 2010, submitted), this corresponds to the syntactic combinatorial potential of a predicate, represented as its argument structure (ARG-ST) and valence (VAL) lists. These lists are closely related to the core grammatical constituents of the clause; for example, the subject is typically the first member of both lists. I will elaborate upon the distinction between argument structure and valence in Chapter 2.

Information structure, a component of pragmatics, represents the aspects of discourse that influence the form of sentences. It tracks the *activation status* of discourse referents, that is, the degree to which a referent is salient to the speaker and hearer (e.g. Prince 1981a, Ariel 1990). It also contains a set of *pragmatically structured propositions* that specify how propositional content is processed and stored (e.g. Reinhart 1981, Lambrecht 1994). Cross-linguistically, information structure has been shown to affect the form of sentences and phrases at several levels of representation, including prosody, morphology, and syntax.

This dissertation investigates the relationship between information structure and argument realization through a single grammatical phenomenon in one language: subject selection in English. Even in this circumscribed environment, it is clear that information structure has a profound effect on argument realization. This chapter begins to lay the groundwork for the investigation. Section 2 introduces two ways in which information structure can influence argument realization: direct licensing and resolution. This dissertation will demonstrate that both are involved in English subject selection phenomena. Section 3 provides a brief outline of the remainder of the dissertation.

2. Two Components of an Interface Model

This dissertation investigates two ways in which information structure shapes the syntactic combinatorial potential of a predicate. In *direct licensing*, information structure serves as the formal licensor of a phrase, just as lexically encoded semantic roles do. In *resolution*, information structure influences the mapping between lexical semantics and syntax when it is underspecified. Both phenomena are unexpected if one assumes a purely semantic

theory of argument realization. Accordingly, they serve as strong support for an interface model.

2.1 Direct Licensing

In direct licensing, information structure enables the licensing of a syntactic constituent that has no semantic tie to any predicate in the sentence. This runs counter to a key assumption about the syntax-semantics interface that appears in many theoretical frameworks: that every meaningful syntactic constituent must correspond to some element in the semantics (e.g. the θ -criterion in Government and Binding Theory (Chomsky 1981: 36), the Completeness Constraint in Role and Reference Grammar (Van Valin and LaPolla 1997: 325), and Jackendoff's (1990: 22) correspondence constraint linking syntax to conceptual structure). The interface model presented here assumes that phrases can be licensed on the basis of their functions within the informational component. This is done via *constructions*, grammatical schemata that specify how morphosyntactic representations can be constructed. In the case of direct licensing, a construction pairs informational content with a change in the predicate's valence.

To the best of my knowledge, no previous research has argued for direct licensing in English or typologically similar languages. This dissertation proposes that direct licensing is required in order to account for subject selection in two English constructions: Topical Exclamatives (TE), discussed in Chapter 3, and Copy Raising (CR), discussed in Chapter 4. CR has been explored in previous research (Rogers 1971, 1972, 1973, 1974, Horn 1981a, Lappin 1984, Heycock 1994, Potsdam and Runner 2001, Asudeh 2004, Asudeh and

Toivonen 2005, 2006, 2009) while TE has not previously been discussed in the literature.

The following is a brief overview of the analyses I propose.

Both constructions are characterized by the presence of a referential subject that alternates with an expletive. TE and its expletive subject counterpart alternant appear in (1-2), while CR and its expletive counterpart are illustrated in (3-4).²

- (1) a. [People are amazing when you give them the information and the tools they need.] They're amazing the responsibility they'll accept.³
b. It's amazing the responsibility they'll accept.
- (2) a. [The vote for Republican Governor though was simply a vote against Spitzer. I think] he's horrible the way he goes after business.⁴
b. It's horrible the way he goes after business.
- (3) a. He seems like he'd be a good candidate for the job.
b. It seems like he'd be a good candidate for the job.
- (4) a. [His style is wooden, old-fashioned, and artificial] ... he feels to me like he belongs in another era.⁵
b. It feels to me like he belongs in another era.

Both are extraposition constructions, in which the most prominent argument of the main predicate is displaced to a post-verbal position; in (1), this is the noun phrase *the responsibility they'll accept*. For syntactic reasons, extraposition is obligatory for the class of verbs that head CR, namely epistemic verbs (*seem, appear*) and perceptual resemblance verbs (*look, sound, feel, taste, smell*), but optional for the evaluative predicates that head TE (*amazing, horrible, unusual, sucks*). For example, (2) has a “canonical” counterpart in which

² Throughout the dissertation, I provide the sources of naturally-occurring examples in footnotes. If the example is not footnoted, the reader can assume that it was constructed. Portions of examples that appear in brackets represent supporting material that is not part of the target structure.

³ www.sbnonline.com/Local/Article/9794/77/0/Squeezing_the_tube.aspx?Category=92, accessed 11/1/2009

⁴ www.abovetopsecret.com/forum/thread232895/pg3, accessed 11/1/2009

⁵ www.amazon.co.uk/review/R3RD409NYAYBYY, accessed 11/1/2009

the NP complement of *horrible* appears in subject position: *The way he goes after business is horrible.*

In Chapters 3 and 4, I argue that referential subjects in TE and CR are not selected at the semantic level, but rather are licensed to serve as *topics*, defined as an aboutness relation that links an individual to a proposition (Reinhart 1981, Lambrecht 1994). There are two main sources of evidence for this analysis. First, subjects in TE and CR do not stand in a systematic semantic relationship to any predicate in the sentence. Second, their interpretation is pragmatically constrained due to the fact that they conventionally express topicality. Subjects in TE and CR are constrained by *pragmatic type restrictions*, which I take to be the defining feature of topic expressions: most importantly, they must refer to a specific entity or set. NP types that do not refer, such as predicative nominals, cannot serve as topic expressions, and thus cannot appear as subjects in TE or CR. Referentially ambiguous NPs, such as bare plurals, indefinites, and numerically quantified NPs, are obligatorily interpreted as specific. Further support for the analysis comes from the observation that the denotata of TE and CR subjects must stand in an aboutness relationship to the proposition formed by the application of the main predicate to its arguments. I will demonstrate that subject licensing in TE and CR can be captured through a single construction that takes a predicate that licenses extraposition as input and returns a predicate that licenses a topical subject.

2.2 Resolution

In the second component of the model, information structure determines aspects of the *linking* between lexical semantics and syntax: the system that determines which

semantically-selected arguments are realized as subjects and objects. Resolution requires a theory in which linking is not entirely determined by the semantics. When lexical semantic structure does not fully determine argument linking, information structure can play a decisive role.

I am not aware of any previous analyses that invoke resolution to account for argument realization phenomena. In Chapter 5 of this dissertation, I argue that resolution underlies constraints on the Instrument Subject (IS) alternation in English. This alternation centers on the syntactic realization of instruments, which can appear as adjuncts, as in the (a) sentences below, or as subjects, as in (5-7b). Examples (7) and (8) demonstrate that the alternation is tightly constrained: (7b) is acceptable only with significant contextual support (e.g. when there is contrastive stress on the subject) while (8b) is impossible.

- (5) a. John broke the vase with the club.
b. The club broke the vase.
- (6) a. John opened the door with the key.
b. The key opened the door.
- (7) a. Melville wrote *Moby Dick* with this pen.
b. ? This pen wrote *Moby Dick*.
- (8) a. John ate the spaghetti with the fork.
b. * The fork ate the spaghetti.

In some previous research, it has been argued that the observed constraints emerge entirely from the semantic properties of instruments and the events in which they participate (Reinhart 2002, Grimm to appear). Other researchers have noted that information structure is somehow involved in determining whether a particular instrument makes an acceptable subject (DeLancey 1984, 1991, Schlesinger 1989, Brousseau 1998, Alexiadou and Schäfer 2006).

This dissertation provides the first account that makes the contributions of semantics and information structure explicit. Starting with the semantics, I show that instrument subjects only occur with polyvalent verbs entailing physical affectedness. This is because the event structure of these verbs includes an *action chain*, representing subevents in which *actors* transmit force (cf. Jackendoff 1990, Croft 1991, 1994, 1998). Both the instrument and the implicit agent of instrument-subject sentences like (5-7b) may appear as actors. The semantic component of the linking system contains a constraint that requires the subject of a verb entailing affectedness to be an actor. This constraint delimits the class of instruments that can be realized as subjects: instruments such as *the fork* in (8b) are ruled out because they are not actors.⁶ However, it underdetermines subject selection, as both agents and instrument actors are potential subjects.

In a pragmatically “neutral” context (to the extent such a thing exists), the initiator of the action chain, typically the agent, will be linked as the subject. A non-initiating actor, such as the instruments in (5-7b), can serve as subject only when it is connected to a salient (open) proposition. Evidence for this analysis comes from the observation that instrument subjects improve in acceptability when they are contained in grammatical structures that conventionally express open propositions, such as clefts, relative clauses, and bare focus. They also become more acceptable in discourse contexts that make instrumentality salient, such as advertisements and product reviews. The effects of information structure are encoded in a construction that permits a non-initiating actor to be linked as subject only when it satisfies the appropriate informational constraints.

⁶ The action chain implements the event of physical affectedness denoted by the verb. The verb *eat* denotes an event of consumption, but *the fork* in (8b) contributes directly only to the transport of the food, not to its consumption.

3. Structure of the Dissertation

The remainder of the dissertation is structured as follows. Chapter 2 presents the theoretical foundations of the interface model, starting with the approach I take to defining grammatical functions such as *subject* and then turning to the assumptions that I make about the informational component, particularly the relations *topic* and *focus*. The chapter concludes with a brief overview of Sign-Based Construction Grammar, emphasizing the aspects of the syntactic, semantic, and pragmatic components that are essential to the interface model. Chapters 3 and 4 turn to the phenomenon of direct licensing, demonstrating that information structure underlies the licensing of matrix subjects in Topical Exclamatives and Copy Raising. Chapter 5 focuses on resolution, which I argue is the source of constraints on the distribution of instrument subjects. Finally, Chapter 6 offers concluding remarks on the implications of the interface model and its possible extensions in English and across languages.

Chapter 2

Subject and Topic, Revisited

1. Introduction

Chapter 1 sketched a preliminary picture of the context in which the interface model is situated. On the one hand, it has long been known there is a close correlation between information structure and grammatical functions such as *subject* and *object* (Firbas 1964, 1966, Halliday 1967, the papers in Li 1976, Reinhart 1981, Van Oosten 1986, Lambrecht 1994, 2000, Michaelis and Francis 2007). On the other hand, many highly influential theories of argument realization are based entirely on lexical semantics (Jackendoff 1990, Dowty 1991, Croft 1991, 1994, 1998, Rappaport Hovav and Levin 1998, Levin and Rappaport Hovav 2005). The main goal of this dissertation is to make progress towards bridging the gap: to come to a better understanding of the relationship between information structure and argument realization (cf. Tomlin 1985, Lambrecht 1995, Aissen 1999, Goldberg 2004, Ruppenhofer 2004).

This chapter presents further background information that is helpful in contextualizing the assumptions and predictions of the interface model. Section 2 introduces the assumptions that I make about grammatical functions, with special focus on subjects. Section 3 turns to previous models of argument realization, starting with models that are grounded in semantic structure and then surveying previous research on the informational

component of argument realization. In Section 4 I discuss the components of information structure that underlie the interface model, namely propositions that are pragmatically structured by relations such as *topic* and *focus*. Section 5 presents the basic principles of the theoretical framework that I assume in this dissertation, Sign-Based Construction Grammar. Section 6 provides concluding remarks.

2. Subjects

Subjects are many things to many theories. The term *subject* has been used for constructs across several distinct levels of linguistic structure. This is evident in the traditional distinction made between *grammatical*, *logical*, and *psychological* subjects, which pick out prominent constituents at the grammatical, semantic, and pragmatic/informational levels, respectively. In this dissertation, the term *subject* refers to grammatical subjecthood, in line with the dominant usage of the term in contemporary linguistics. Even within the domain of grammatical subjecthood, however, there is little consensus across frameworks as to what constitutes a subject. This section surveys several of the key questions that emerge in attempts to define subjecthood as well as some of the influential approaches to these issues that have been proposed.

What the many distinct approaches to grammatical subjecthood share is the idea that a subject is a constituent that is formally marked as prominent. Formal prominence manifests itself in a variety of ways: marking of the constituent itself through word order or morphological case, control of other grammatical aspects of the sentence, e.g. verb agreement, and behavior in particular syntactic constructions (Keenan 1976, Van Valin and LaPolla 1997). In the typical case, the subject's formal prominence reflects the

prominence of its denotatum at the level of interpretation, broadly defined. As Keenan (1976) points out, subjects are prototypically agentive and topical, i.e. prominent (in some sense) with respect to lexical semantics and information structure. However, grammatical subjecthood is independent from its semantic and pragmatic correlates.

Beyond this, approaches to subjecthood vary considerably, reflecting a larger debate on the status of grammatical relations within the language system. In some frameworks, such as Relational Grammar (Perlmutter 1980), Lexical Functional Grammar (Bresnan 2001), and Simpler Syntax (Culicover and Jackendoff 2005), grammatical functions such as *subject* and *object* play a direct and central role in morphosyntactic representations. In other frameworks, for example the Standard Theory of transformational grammar (Chomsky 1965), Government and Binding (Chomsky 1981), Minimalism (Chomsky 1995), and Sign-Based Construction Grammar (Michaelis 2009, Sag 2010, submitted), subjects have no theoretical status and serve only as shorthand for distinguished elements in syntactic representations (cf. the discussion in McCloskey 1997). For the frameworks in which grammatical relations have the status of formal objects, there is an additional distinction as to whether they are primitives of the theory (as in Relational Grammar) or derived notions (as in LFG, in which they are defined in terms of features).

In this dissertation, I adopt the SBCG view of grammatical functions. *Subject*, for example, is not a formal element of the theory but a convenient way of referring to a prominent syntactic constituent. In SBCG, the subject is typically the first member of a predicate's argument structure (ARG-ST) and valence (VAL) lists, as well as the external argument (XARG), which is accessible to non-local grammatical constraints. These

components of the theory are discussed in Section 5 of this chapter. A question this raises is what is gained by continuing to refer to “subjects” as such in a framework that lacks them. The view of grammatical functions proposed by Van Valin and LaPolla (1997) provides one possible answer: “subjecthood” reflects a generalization over the multiple distinct aspects of syntactic prominence, an observation that they tend to coincide.

2.1 *Subject as a Generalization*

Van Valin and LaPolla argue that a grammatical function is a “restricted neutralization of semantic roles for syntactic purposes” (1997: 253), i.e. a syntactic pattern that crosscuts distinctions at other levels of representation. They illustrate the notion of restricted neutralization with data from English subject-to-subject raising (SSR; p. 252).

- (1) a. Jack_i seems _____i to be eating a hamburger.
- b. Jack_i seems _____i to be taller.
- c. Jack_i seems _____i to have been arrested by the police.
- d. * Jack_i seems the police to have arrested _____i.

In SSR, an argument of the embedded clause is realized as the subject of the matrix clause. Example (1) demonstrates that syntactic rather than semantic constraints determine which argument appears as the main-clause subject. There is a neutralization of the semantic role of the distinguished argument: in (1a) it is an agent, in (1b) the bearer of a property, and in (1c) a theme. The matrix subject must, however, be the “missing” subject of the embedded clause, as demonstrated by the distinction between (1c) and (1d). This is preliminary evidence that subjecthood is a construct relevant to the grammatical description of English.

Grammatical relations such as subject and object are not themselves components of Van Valin and LaPolla’s theory. In their place, they propose two notions: *controller*, a

constituent whose features determine some aspect of formal structure, and *pivot*, the privileged constituent in a syntactic construction. For example, in English the subject is generally the controller of verb agreement. The pivot of SSR – the element that is realized as the matrix subject– is the embedded subject. An important aspect of their theory is that controllers and pivots “exist only with reference to specific morphosyntactic phenomena, and each grammatical phenomenon may define one controller and/or one pivot” (p. 275). According to Van Valin and LaPolla, what we take to be grammatical relations are in fact generalizations over construction-specific designations of controller or pivot. When the same constituent consistently serves as controller and pivot across constructions within a language, we take it to be a subject. I will adopt this general take on subjecthood, replacing the notions of *controller* and *pivot* with the syntactically prominent constituents of SBCG: e.g. the XARG and the first member of the ARG-ST and VAL lists.

One advantage of this model is that it accounts for variability with respect to grammatical relations that exists across languages and across constructions within a language. First, in many languages grammatical relations are not as well motivated as they are in English. For example, Van Valin and LaPolla (1997: 260-263) argue that the category of subject is not motivated in Mandarin because there exist no restricted syntactic neutralizations of the type that are pervasive in English (cf. Li and Thompson’s (1976) claim that Mandarin is topic-prominent rather than subject-prominent). Second, even within the class of languages with well-motivated grammatical relations there is variation with respect to the ground covered by each individual relation. A well-known example of this is the distinction between nominative-accusative and ergative-absolutive grammatical

systems, which differ in how the formal marking of grammatical relations is distributed across sentence types. Finally, Van Valin and LaPolla demonstrate that variability in controller/pivot choice exists even in languages like English that have well-motivated subject constructs. Control of co-reference is a typical property of subjects across languages (Keenan 1976, Van Valin and LaPolla 1997). However, while English subjects typically assume control of co-reference across clauses, they do not do so in the Control construction. Example (2) illustrates that the matrix subject sometimes (2a-b) but not always (2c) co-refers with the “missing” embedded subject.

- (2) a. Mary_i wants ____i to leave.
b. Mary_i promised Bill_j ____{i/*j} to leave.
c. Mary_i asked Bill_j ____{*i/j} to leave.

Co-reference is instead determined by the lexical semantics of the matrix verb (Comrie 1986, Sag and Pollard 1991, Jackendoff and Culicover 2003).

I will continue to refer to NPs bearing the formal properties characteristic of subjects as such, with the understanding that subjecthood is a generalization over different aspects of syntactic prominence. Keenan’s (1976) notion of subjecthood as a prototype is helpful here. English subject NPs are typically coded in the following ways: they appear in pre-verbal position (further defined in different ways by different syntactic frameworks) and receive nominative case when coded as pronouns. In terms of syntactic behavior, subjects are typically the controllers of verb agreement and co-reference across clauses. A prototypical subject in English exhibits all of these properties; less typical subjects may exhibit only a proper subset.

2.2 Special Classes of Subjects

On that note, there are two special classes of subjects, each of which merits a brief discussion. The first is expletive subjects, meaningless elements that exhibit (some of) the formal behavior of prototypical subjects. The standard take on expletives within formal models of grammar is that they are licensed in order to satisfy purely syntactic constraints.¹ In constructions with expletive subjects, the NP that plays the most prominent semantic role in the sentence – in traditional terminology, the logical subject – appears in a syntactic position not typically associated with subjecthood. This is a type of *extraposition*, in which a constituent appears in a less prominent syntactic position than it would be expected to on the basis of semantic prominence alone.² Two examples of extraposition appear below: existential constructions with the expletive subject *there* (3) and nominal extraposition with expletive *it* (4).

- (3) There are two dogs chasing a cat over there.
cf. Two dogs are chasing a cat over there.
- (4) *It are / It is amazing the people she met during the project.
cf. The people she met during the project are amazing.

Both expletives appear in preverbal position, as is typical of subjects. However, the two constructions differ with respect to verbal agreement, which is typically controlled by the subject. In existential constructions, the semantic features of the extraposed NP typically

¹ In contrast, Langacker (1995), working within the framework of Cognitive Grammar (Langacker 1987, 1991) argues that so-called expletives in fact do have conceptual content. Bolinger (1973) makes a similar claim.

² This definition of extraposition is intended to be neutral with respect to whether extraposition constructions are derived from their “canonical” counterparts or whether the two constructions have distinct representations.

determine verb agreement;³ this does not occur in nominal extraposition. This demonstrates that atypical subjects may exhibit only a proper subset of the formal properties associated with subjecthood, and also that subject properties can be divided between multiple NPs.

This observation brings us to the second special case, the Multiple Subject Construction (MSC). In a MSC, multiple phrases have coding and/or behavioral properties that are associated with prototypical subjects in that language. There are two major types of MSC: constructions in which one of the subjects is an expletive and constructions in which both subjects are meaningful. The former case is illustrated by Icelandic (5), in which both an expletive and a referential NP appear in syntactic positions associated with subjecthood (McCloskey 1997: 215, cf. Bobaljik and Jonas 1996). A well-known case of the latter comes from Japanese. In the Japanese MSC (6), both of the sentence-initial NPs exhibit some aspects of subject-like syntactic behavior, such as the ability to bind a reflexive within the clause (example from Landau 2009: 94).

- (5) pað dansaði maður í garðinum.
 there danced a-man in the-garden
 ‘There danced a man in the garden.’
- (6) Taro-ga_i usagi-ga_j zibun-zisini-no_{i/j} heya de sinda.
 Taro-NOM rabbit-NOM self-GEN room in died
 Reading 1: ‘Taro’s rabbit died in his (Taro’s) room.’
 Reading 2: ‘Taro’s rabbit died in his (own) room.’

The upshot of this discussion is that a monolithic view of subjecthood cannot hold across languages, or even across constructions within the same language. Instead, subjecthood is best seen, as in Van Valin and LaPolla (1997), as a generalization over the

³ Some speakers also accept sentences such as (i), in which the auxiliary does not agree in number with the extraposed NP:

(i) % There’s two dogs chasing a cat over there.

patterns of morphosyntactic coding and the syntactic behavior of prominent constituents in a particular language. It is only in a language such as English in which these generalizations are quite strong that we can meaningfully speak of subjecthood as independent of the formal properties that mark it and the semantic/pragmatic properties that typically underlie it.

3. Argument Realization

This section reviews previous approaches to argument realization, the system that selects the core grammatical functions of a clause. Because this dissertation deals with subject selection phenomena, the discussion will focus on the realization of subjects.

A good starting point for the discussion is Aissen's (1999) Optimality Theoretic analysis, which provides a broad overview of the factors that influence subject selection across languages (see also Legendre et al. 1993). She identifies three relevant interpretive parameters: person, semantic prominence and informational or discourse prominence. Her notion of semantic prominence is adopted from Dowty's (1991) theory of argument selection, in which the argument that is closest to the prototypical agent is realized as subject of an active sentence (see discussion in Section 3.1). On her account, informational prominence consists of attention and activation status: discourse-prominent referents are those that are the focus of attention and/or are linked to the preceding discourse.⁴

Universal constraint rankings indicate the degree to which a mapping between a particular component of meaning and a grammatical function is marked. Specifically, third person NPs are more marked as Subjects (Su) than first and second person pronouns,

⁴ She refers to such discourse-prominent constituents as topics, but they are not equivalent to aboutness topics as defined in Section 4.1.

which Aissen calls “local persons” ((7a); p.681). The notation indicates that the constraint penalizing third-person subjects (*Su/3) is ranked more highly than the constraint penalizing local persons as subjects (*Su/Local). Likewise, patients are more marked as subjects than agents ((7b), p. 683), and non-discourse-prominent referents (x) are more marked than prominent participants (X; (7c)).⁵

- (7) a. *Su/3 >> *Su/Local
 b. *Su/Pat >> *Su/Ag
 c. *Su/x >> *Su/X

The ranking of these constraints varies across languages, accounting for several aspects of cross-linguistic variation in subject selection. One example is the availability of the passive construction, in which the subject is discourse-prominent but not semantically prominent. On Aissen’s account, the relative ranking of constraints involving discourse prominence and semantic prominence determines whether a language has a passive construction. In English, non-discourse-prominent subjects are more highly marked than non-semantically-prominent subjects, and for this reason English has the passive voice (p. 688). In contrast, the Algonquian language Fox and the Sino-Tibetan language Nocte, which lack passive constructions, rank semantic prominence above all other constraints (p.686). The constraint rankings that derive this result are illustrated below; *GR/Pers and *GR/X refer collectively to the body of constraints linking grammatical relations to person and informational prominence, respectively.

- (8) a. English: *Su/x >> *Su/Pat >> *GR/Pers
 b. Fox/Nocte: *Su/Pat >> ... >> {*GR/Pers, *GR/X}

⁵ Aissen does not explicitly specify a constraint ranking for discourse prominence parallel to person and semantic prominence. However, it can reasonably be inferred from the analyses she presents.

Aissen's constraint ranking for English most strongly penalizes non-discourse-prominent subjects, followed by patients linked as subjects. Constraints related to person are ranked even lower and do not directly influence English subject selection. The consequences of this are that (1) the passive construction emerges as optimal in sentences in which the patient is more informationally prominent than the agent and (2) active sentences are optimal otherwise. This analysis seems to capture broad generalizations about English subject selection. Previous research has demonstrated that discourse constraints govern the choice between the active and passive voice (Tomlin 1985, Thompson 1987) and subject selection within active clauses is widely taken to be determined by semantic prominence (Jackendoff 1990, Dowty 1991, Rappaport Hovav and Levin 1998, Levin and Rappaport Hovav 2005).

However, Aissen's account misses subtler aspects of subject selection within active clauses. In some respects, the analysis overgenerates. For example, it fails to predict the existence of active clauses in which the object is more informationally prominent than the subject. However, the response in (B1) below illustrates that sentences of this sort can be entirely acceptable.⁶ (The passive response in B2, which Aissen's account predicts to be optimal, is also felicitous.)

- (9) A: What happened to the tree?
B1: A truck hit it.
B2: It was hit by a truck.

In other respects, the account is limited. The effects of information structure are restricted to selecting between members of the candidate set, which consists of various argument

⁶ It would be possible to generate both forms in (9B) with a slight modification to Aissen's account: free ranking of *Su/x and *Su/Pat.

realization patterns. There appears to be no way to capture direct licensing phenomena, in which information structure influences the set of arguments itself.

The remainder of this section gives an overview of the two major approaches to argument realization that have previously been proposed: purely semantic approaches and models with an informational component. This serves as background for the interface model proposed in this dissertation. A note on terminology: I will use the terms (*argument*) *realization*, *selection*, *mapping* and *linking* interchangeably, as is common in the literature.

3.1 Semantic Models of Argument Realization

The dominant view across frameworks dating back to the work of Gruber (1965, 1976) and Fillmore (1968) has been that semantic prominence is the main factor governing argument realization in active clauses. The numerous proposals for deriving argument realization from semantic prominence have differed with respect to (1) how dominance is encoded in the semantics, (2) how the relevant components of the semantics are mapped to syntax, and (3) whether the set of licensors includes constructions as well as predicates. The literature on semantic theories of argument realization is vast, and I will not attempt to do justice to it here. For a comprehensive recent review, the reader is referred to Levin and Rappaport Hovav (2005). I will briefly discuss each of the issues mentioned above, and then present the approach to the semantic component of argument realization that I assume in this dissertation. This subsection will conclude by touching on the claims that Dowty (1991) makes – from the perspective of a purely semantic model – about the relationship between information structure and argument realization.

3.1.1 Semantic Determinants of Argument Realization

One central issue that has guided research on argument realization is how to characterize the semantic content that influences the selection of subjects and objects. Though many models make reference to familiar semantic role labels such as *agent*, *patient*, and *goal*, they differ as to whether these roles are primitives of the theory or, increasingly in contemporary research, defined in terms of their own semantic content. The latter approach underlies several highly influential theories of argument realization. In the models of Jackendoff (1990), Croft (1991, 1994, 1998) and Levin and Rappaport Hovav (1995, 2005, Rappaport Hovav and Levin 1998), semantic roles are defined as distinguished elements within highly articulated event structure representations. In the theories proposed by Dowty (1991), Wechsler (1995), and Davis (2001), argument realization proceeds on the basis of lexical entailments that hold of event participants.

Starting with the event structure approach, the models of Jackendoff (1990) and Levin and Rappaport Hovav (1995, 2005, Rappaport Hovav and Levin 1998) have several fundamental similarities. In both models, event structures are organized into subevents headed by a small inventory of primitive predicates. Though the specific inventories are not identical, they overlap to a significant degree: for example, both include predicates that represent states and dynamic events, as well as causation. The two models also differ in several key respects. Jackendoff's event structures have two levels of representation that have no counterpart in the work of Levin and Rappaport Hovav: the *action tier*, which encodes the ways in which entities act upon each other, and the *thematic tier*, which deals with location and motion. Levin and Rappaport Hovav also divide event structures into

two components, but on an entirely different basis. They propose that the event representation of a verb has two components: an *event schema* that relates the obligatory participants through a small vocabulary of primitive predicates, and a *root* representing the aspects of the verb's meaning that distinguish it from others with the same event schema. Example (10) is Jackendoff's (1990: 142) representation of *Phil opened the door*; (11a) represents the same sentence in the framework of Rappaport Hovav and Levin (1998).⁷

- (10) $\left(\begin{array}{l} \text{CS ([PHIL], [INCH [BE ([DOOR], [OPEN])]])} \\ \text{AFF ([PHIL], [DOOR])} \end{array} \right)$
- (11) a. [Phil ACT [CAUSE [BECOME [DOOR <OPEN>]]]]
 b. [X ACT [CAUSE [BECOME [Y <STATE>]]]]

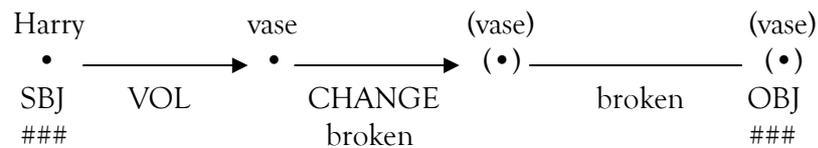
Starting with Jackendoff's representation, CS (abbreviating *conceptual structure*) encodes the thematic tier, while AFF encodes the action tier. The first argument of the thematic tier, *Phil*, initiates an dynamic event (INCH) that results in a state (BE) in which the door is open. The action tier indicates that Phil acts upon the door. In Rappaport and Hovav's representation, Phil acts in a way that causes the door to become open. The general event schema associated with (11a) appears in (11b); the root specifies the nature of the result state, in this case the state of being open.

Croft (1991, 1994, 1998) argues that a fundamentally different type of event structure representation underlies argument realization. In his model, events are represented as causal chains composed of subevents denoting asymmetric force-dynamic relationships between participants (cf. Talmy 1976, 1988, Langacker 1987, 1991). Representations are linear, with force passing left-to-right from participant to participant. The following is

⁷ Example (10) is a simplified version of Jackendoff's representation of the sentence *Phil opened the door with a key*. Example (11) is constructed on the basis of the general event schema that Rappaport and Levin (1998: 108) provide for causative verbs.

Croft's (1994: 38) representation of the causal structure of *Harry broke the vase*.⁸ In this scenario, Harry acts on the vase, and thus precedes it in the representation of the causal chain.

(12) Harry broke the vase.



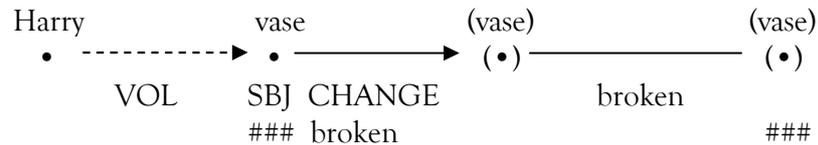
In Croft's model, argument realization depends not only on the causal chain, but also on the *verbal profile*. The verbal profile indicates the portion of the event structure that is emphasized (or "asserted"), in contrast with the portion that is "presupposed," or in the background. Subjects and objects are the initial and final endpoints, respectively, of the profiled portion of the event representation. Verbs can be associated with more than one profile, resulting in argument structure alternations. For example, verbs that undergo the causative-inchoative alternation, such as *break*, have two profiles: a causative profile that begins with the initiator of the event (example (12) above) and an inchoative profile that begins with the patient (13).⁹ Example (13) depicts the same scenario as (12) but with a change in profile that excludes the causing subevent. This is reflected by the pound signs that delimit the profiled portion of the event, as well as the dashed arrow that links the

⁸ Dots represent participants; a dot in parentheses indicates that the participant is identical to the preceding one in the representation. Arrows represent dynamic subevents, lines states. The marker underneath each line indicates something about the nature of the event or state. For example, VOL indicates that Harry acts on the vase volitionally, CHANGE and *broken* indicate that the vase undergoes change that results in a state of being broken. The pound signs delimit the profiled portion of the event, and thus coincide with the choice of subject (SBJ), and object (OBJ), if applicable. See further discussion below.

⁹ Croft does not provide the representation for (13). The representation provided is inferred from Croft's representation of (12) as well as the representations he provides for other inchoative sentences (e.g. 1994: 40).

agent to the patient. Because the vase is the first participant in the profiled portion of the event, it is realized as subject.

(13) The vase broke.



Dowty's (1991) model is perhaps the most influential of the entailment-based approaches. He proposes that certain lexical entailments are associated with the prototypical agent (*proto-agent*) and others with the *proto-patient*. The argument that bears the most proto-agent entailments will be realized as the subject of a (transitive) active clause, while the argument that is the closest match to the proto-patient will be the object. The lexical entailments associated with the proto-agent and proto-patient are the following (p. 572):

- (14) Dowty's Proto-Agent Entailments
 - A. Volitional involvement in the event or state
 - B. Sentience (and/or perception)
 - C. Causing an event or change of state in another participant
 - D. Movement (relative to the position of another participant)
 - E. Exists independently of the event named by the verb
- (15) Dowty's Proto-Patient Entailments
 - A. Undergoes change of state
 - B. Incremental Theme
 - C. Stationary relative to movement of another participant
 - D. Does not exist independently of the event, or not at all.

3.1.2 Linking Semantics to Syntax

Turning to the relationship between lexical semantics and syntax, Levin and Rappaport Hovav (2005: 145-152) point out that there are two main approaches: absolute and relative

mapping. In absolute mapping, an argument is linked to a grammatical constituent without reference to other arguments in the clause. In relative mapping systems, the syntactic realization of an argument depends on its semantic prominence relative to its co-arguments.

A particularly restrictive version of absolute mapping is the Uniformity of Theta Assignment Hypothesis (UTAH), in which “identical thematic relationships between items are related by identical structural relationships between those items at the level of d-structure” (Baker 1988: 46, Baker 1997).¹⁰ According to UTAH, all agents are linked to a specific position in underlying syntactic structure, both across constructions and across languages; the same goes for other semantic roles. Absolute mapping principles also play a central role in the models of Jackendoff (1990), Levin and Rappaport Hovav (1995, 2005, Rappaport Hovav and Levin 1998), and Davis (2001). One example is Levin and Rappaport Hovav’s (1995: 135) Immediate Cause Linking Rule, which states that “the argument of a verb that denotes the immediate cause of the eventuality described by that verb is its external argument.”

An influential early relative mapping approach is that of Fillmore (1968), in which argument realization is determined by the ranking of arguments on a semantic hierarchy. Semantic hierarchies figure into many other models of argument selection, including those of Dik (1978), Givón (1984), Bresnan and Kanerva (1989), Grimshaw (1990), Jackendoff (1990) and Baker (1997). Levin and Rappaport Hovav (2005: 158-162) note that the main

¹⁰ The precursor to UTAH was the Universal Alignment Hypothesis (UAH, which claims that “there exist principles of UG which predict the initial [grammatical] relation borne by each nominal in a given clause from the meaning of the clause” (Perlmutter and Postal 1984: 97). As Levin and Rappaport Hovav (2005: 131) point out, the UAH does not make explicit what the relevant principles of UG are, and thus constitutes a less restrictive approach to mapping than UTAH.

motivation for semantic hierarchies is context-dependence in argument realization, both within and across languages. For example, Fillmore (1968) observes that although both agents and instruments are possible subjects in English (16a-b), if both are realized within a single clause the agent must be the subject (16a vs. 16c).

- (16) a. John opened the door (with the chisel).
b. The chisel opened the door.
c. * The chisel opened the door by John.

He accounts for this through the proposal that agents outrank instruments on the semantic hierarchy. As the most highly ranked argument must be realized as subject, an instrument can be realized as subject only if it is not outranked by an agent. (In Chapter 5 of this dissertation, I will argue that constraints on the distribution of instrument subjects can be captured via an absolute mapping approach.)

Dowty's (1991) model assumes a different type of relative mapping approach. In his analysis, argument realization is determined through the closeness of the match between the lexical entailments of a predicate's arguments and those associated with the proto-agent and proto-patient. This constitutes a relative mapping approach because the way an argument is mapped is determined by comparing its lexical entailments with those of its co-arguments.

3.1.3 Constructions and Lexical Representations

The third key issue is whether arguments are licensed on the basis of constructions, in addition to lexical representations. Adele Goldberg (1995, 2002, 2006, cf. Kay 2005) has argued in favor of an inventory of *argument structure constructions* that pair frame-semantic representations with syntactic representations. The frames associated with the construction

may be a superset of those of the verb that enters into it, with the result that the construction licenses constituents that are not selected by the verb. One example is the Caused Motion construction, illustrated by the examples below (Goldberg 1995: 152). Goldberg proposes that the construction selects for a *cause* that causes a *theme* to move along a *path*. The argument structure of the construction combines with that of the verb, with the consequence that verbs like *laugh* and *sneeze*, which typically select only one argument, can license the complex argument realization patterns below.

- (17) a. They laughed the poor guy out of the room.
b. Frank sneezed the tissue off the table.
c. Mary urged Bill into the room.

Rappaport and Levin (1998) propose an alternative lexical account of phenomena of this sort. They argue that verbs such as *laugh* and *sneeze* allow augmentation of their basic (univalent) event schema, resulting in syntactic patterns such as (17). It is a continuing topic of debate whether template augmentation can provide the same level of empirical coverage as argument structure constructions (see Goldberg 2005 for a discussion of points of departure between the two approaches).

3.1.4 The Present Approach

My approach to the semantic component of argument realization will be presented in detail in Chapter 5; I will provide only a brief overview here. I adopt elements from a range of models, including those of Levin and Rappaport Hovav (1995, 2005, Rappaport Hovav and Levin 1998), Jackendoff (1990), Croft (1991, 1994, 1998), and Davis (2001). I assume that the semantic underpinnings of argument realization come in the form of event

structure representations, both lexical and constructional. Most aspects of the event structures that I propose are adapted from those of Levin and Rappaport Hovav. However, I propose one significant innovation that is inspired by Croft's model. This is that the first argument of a particular class of verbs can be broken down into an *action chain*.¹¹ The action chain is conceptually equivalent to the causal chains that comprise Croft's model; both represent chains of subevents in which force is transmitted from one event participant to another. Following Levin and Rappaport Hovav as well as Davis (2001), I adopt an absolute mapping approach to the linking between lexical semantics and syntax.

As in other research in Sign-Based Construction Grammar (Michaelis 2009, Sag 2010, submitted) and its precursor, Head-Driven Phrase Structure Grammar (e.g. Davis 2001), I assume that the lexicon has a hierarchical structure. In the present model, this amounts to a claim that verb classes inherit both event structure representations and linking constraints from the broader-grained classes that subsume them. I will elaborate on the significance of the hierarchical model in Section 5 of this chapter.

3.1.5 Dowty (1991) on Perspective-Dependent Roles

Perhaps unsurprisingly, much of the literature on semantic approaches to argument realization has little to say about the possible role of information structure. Dowty's (1991) article is an exception, in that it directly tackles the question of whether to include information structure in the system (p. 562-566). Dowty discusses what he calls "perspective-dependent" predicates, in which variability in argument realization is

¹¹ The set of verbs in question is the class of polyvalent physical affectedness verbs, which permit instrument subjects. See Chapter 5 for details.

associated with different perspectives, rather than different event representations. This includes the class of symmetric predicates, such as similarity predicates (*similar, same, different, match*), reciprocal physical events (*touch, collide*) and social relationships (*be married to, meet*), spatial relations (*near, far*), and identity statements with copular *be*. These predicates, in the relevant uses,¹² have been called symmetric on the basis of the observation that if the relation holds between A and B, it will also hold between B and A. It follows from this analysis that A and B have identical semantic roles. This view predicts that the sentence pairs below are truth-conditionally equivalent.

- (18) a. John is similar to Sue.
b. Sue is similar to John.
- (19) a. The porch touches the grass.
b. The grass touches the porch.
- (20) a. John is married to Sue.
b. Sue is married to John.
- (21) a. The porch is near the grass.
b. The grass is near the porch.
- (22) a. John is Dr. Smith.
b. Dr. Smith is John.

Nevertheless, they are associated with subtle interpretive differences. In a series of studies, Tversky and Gati show that reversal of the arguments of a symmetric statement affects the degree to which subjects agree with it (Tversky 1977, Tversky and Gati 1978, Gati and Tversky 1984). For example, their subjects more strongly agreed with (23a) than with (23b). Gleitman et al. (1996) provide further support for this finding, demonstrating that subjects find pairs with reversed argument realization (23) to be *more* different than

¹² Many of these predicates can also be used in a way that is clearly non-symmetric, e.g. (i) does not entail (ii).
(i) The truck collided with the tree.
(ii) # The tree collided with the truck.

pairs in which both arguments are conjoined in subject position, but in different linear orders (24).

- (23) a. North Korea is similar to Red China.
b. Red China is similar to North Korea.
- (24) a. North Korea and Red China are similar.
b. Red China and North Korea are similar.

This suggests that argument realization, and subject selection in particular, is correlated with significant differences in the interpretation of symmetric predicates. These results raise two questions for an account of argument realization. The first is whether the interpretive differences associated with subject selection are semantic or pragmatic/informational in nature. The second is whether these differences in interpretation affect argument realization directly or whether they emerge from the conventional properties of subjects and objects.

Talmy (1978, 1983) argues that the two arguments of “symmetric” predications are in fact assigned two distinct roles: the *figure*, defined as “a moving or conceptually movable entity whose site, path, or orientation is conceived as a variable the particular value of which is the relevant issue,” and the *ground*, “a reference entity, one that has a stationary setting relative to a reference frame, which respect to which the Figure’s site, path, or orientation is characterized” (Talmy 2000: 184). In his model, argument realization in pairs such as (18)-(22) is directly motivated by this semantic distinction: the figure is linked as subject, while the ground is realized as object. Gruber (1976: 45-50), Jackendoff (1978: 94-98), and DeLancey (1991: 341) also argue that subject and object constituents are assigned distinct semantic roles, though their characterizations of the roles involved vary.

Dowty (1991) offers a different type of solution. He argues that the subject and object of symmetric predicates are in fact assigned identical semantic roles, and that the observed interpretive distinctions are due to a conventional association between subjecthood and aspects of information structure. Gleitman et al. (1996) and Bowdle and Gentner (1997) reach similar conclusions. Gleitman et al. (1996) offer support for Talmy's claim that the subject of a symmetric predicate is interpreted as a figure and the object as a ground. However, they deny that the figure/ground distinction is encoded in the lexical semantic representations of similarity predicates. Instead, they claim, it is conventionally associated with the grammatical distinction between subject and object. As for the question of why subjecthood would be conventionally associated with the role of figure, they offer the suggestion that "topics tend to surface foregrounded as the subjects of predication" (p. 360, fn. 23). They suggest that the North Korea/China examples may be interpreted differently because sentences in which *North Korea* is subject are taken to be judgments about North Korea, and vice versa.¹³ Bowdle and Gentner (1997) also argue for an informational account. They suggest that because given information usually appears before new information, subjects of similarity judgments tend to be given and objects tend to be new. Due to the close link between given status and topicality, subjects are typically topical, and therefore similarity judgments tend to be interpreted as contributing to our knowledge of the subject denotatum.

¹³ They also note that the interpretation of similarity predicates involves a domain of comparison, which is usually implicit. For example, John's hair may be similar to Mary's with respect to style but not color. They suggest that one reason the North Korea/China examples may differ in interpretation is due to differences in the domains of comparison that are typically associated with each country.

Dowty situates his analysis within a general theory of the aspects of interpretation that influence argument realization. He distinguishes *event-dependent* thematic roles, which encode the objective properties of an event or state, from *perspective-dependent* roles, which reflect the perspective from which the event is viewed or conceptualized. According to Dowty, event-dependent roles are an element of truth-conditional semantics while perspective-dependent roles are part of the pragmatics or information structure. *Agent* and *patient* are examples of event-dependent roles, which he claims remain constant across different perspectives on a given event. In a situation in which Sue murders John, Sue will always be the agent and John the patient. Examples of perspective-dependent roles are *figure* and *ground*. We may say *Sue resembles John* if our attention is primarily on Sue, or *John resembles Sue* if we are focusing on John, but the state of affairs that we are describing is essentially the same.

Dowty claims that event-dependent roles fully determine argument selection, with no contribution from perspective-dependent roles. His arguments appear below (p. 564).¹⁴

- (i) *In an adequate linguistic description, greater relative degrees of connectedness to previous discourse, givenness, etc., must be explicitly specified as a semantic correlate of grammatical subject denotations (in English-like languages).*
- (ii) *All putative instances of perspective-dependent thematic roles and other ‘perspective-indicating’ lexical entailments of words can be shown to be instances of (i), when properly analyzed.*
- (iii) *Therefore, by Ockham’s razor, perspective-dependent thematic roles are unnecessary, and all roles are event-dependent in meaning.*

¹⁴ In Dowty’s model, thematic roles are defined as the aspects of lexical semantic representations that determine argument selection. His denial that perspective-dependent thematic roles exist (iii) amounts to a claim that perspective-dependent elements of meaning reside outside of this system.

These assumptions enable Dowty to account for the connection between information structure and argument realization in symmetric predicates without explicitly admitting information structure to the system.

There are two significant problems with this account. First, Dowty proposes to “explicitly specify” the connection between information structure and subjecthood somewhere in the grammatical system. However, it is unclear exactly where this information would be represented, given that Dowty wishes to exclude it from the system through which subjects are selected. One possibility is to conventionally associate subjecthood with information structure at the sentence level, i.e. through a phrasal construction. Still, this idea is problematic because subjecthood in English, seen as a generalization over constructions, is not *conventionally* associated with givenness or topicality, in the technical sense; as Dowty himself argues (p. 565), there are subjects that are neither given nor topical. What Dowty seems to want to encode is a probabilistic relation between subjecthood and information structure: for example, that if something is a subject, it is likely to be given.

More importantly, Dowty’s model makes overly conservative predictions about the relationship between information structure and argument realization. It claims that all apparent effects of information structure in languages like English emerge from the informational correlates of subject- and objecthood. It thus predicts that phenomena such as direct licensing and resolution should not exist. If thematic role assignment were the basis of argument realization and if there were no “perspective-dependent” (i.e. pragmatic or informational) roles, then it would be impossible for an argument to be licensed on the

basis of its informational function, as in Topical Exclamatives and Copy Raising. Likewise, if semantic structure alone underlies the selection of subjects, it should be impossible for information structure to advance a less-prominent event participant to subject position, as I argue occurs in the Instrument Subject alternation. For these reasons, purely semantic models such as Dowty's must be replaced (or at least augmented) by an interface model.

3.2 Informational Approaches to Argument Realization

This section reviews the body of previous research which argues that information structure has an effect on argument realization. Goldberg (2004) provides an overview of this topic, and this discussion will overlap with hers in several respects. Perhaps because research on the interface between information structure and argument realization is at a relatively early stage, nearly all studies to date have focused on specific phenomena, rather than the general principles of the system. Among the phenomena that have been addressed are subject selection (discussed in Section 3.2.1), object selection (Section 3.2.2), the licensing of so-called "free" datives (Section 3.2.3), and the null instantiation of semantically selected arguments (3.2.4). At the end of this section, we turn to the few models of the system as a whole that have been proposed to date (Section 3.2.5).

We begin with some preliminary comments that will serve to better situate the discussion. It has often been claimed that discourse prominence is the diachronic source of patterns of argument realization. For example, Givón (1983) refers to subjects as "clause topics" and argues that the historical origin of subjecthood is to mark topicality. Mithun (1991: 160) shares this view, claiming that "the function of subjects is clear: they are essentially grammaticalized clause topics." I will have nothing to say about the diachronic

basis of grammatical functions here. Instead, I will focus on the widespread synchronic effects of information structure on argument realization.

3.2.1 Information Structure and the Selection of Subjects

Previous research has suggested that information structure influences subject selection in two distinct ways across languages. First, subjects tend to denote activated referents, i.e. referents that are salient to the speaker and hearer at utterance time. As we will see, the connection between subjecthood and activation status has been discussed from many perspectives. Second, information structure has been claimed to license subjects directly in topic-prominent languages such as Mandarin and Japanese.

Subjecthood and Activation Status

It is a robust cross-linguistic generalization that given information tends to precede new information within sentences. In early work within the Prague School (e.g. Firbas 1964, 1966), this was captured through the notion of *communicative dynamism*, the idea that sentences tend to become increasingly dynamic, i.e. informative, as they progress. This connects to the observation that there exists a “conspiracy of syntactic constructions” that place activated referents in sentence-initial position (Prince 1981a: 247, cf. Horn 1986, Birner and Ward 1998). It also accounts for the fact that subjects tend to have activated denotata in languages like English (Prince 1992, Michaelis and Francis 2007), where subjects typically appear near the beginning of the sentence.

The literature on Preferred Argument Structure (PAS; DuBois 1987, DuBois et al. 2003) addresses the relationship between activation status and subjecthood from a

different perspective. Like research in the Prague School, PAS offers generalizations about the distribution of given and new information within sentences. However, PAS makes different assumptions about how sentence structure should be broken down. It posits three core grammatical constituents: the subject (A) and object (O) of transitive sentences and the subject (S) of intransitive sentences. DuBois (1987) proposes two constraints that govern the activation status of A, S, and O: (1) the Quantity Generalization, “avoid more than one new argument per clause” (p. 819) and (2) the Given A Generalization, “avoid new A’s” (p. 827). Support for this proposal has come from a range of typologically diverse languages. PAS is thus an additional contribution to the generalization that subjects, and particularly the subjects of transitive clauses, tend to be activated.

An independent line of research has tested whether this generalization accounts for constraints on raising alternations, in which a single constituent can be realized “lower” or “higher” in a clause. The most well-known case in English is Subject-to-Subject Raising (SSR), shown below; (25a) is the “unraised” alternant and (25b) is the “raised” alternant.

- (25) a. It seems that John is in Alaska.
b. John seems to be in Alaska.

Though sentence pairs that are related via raising are usually taken to be truth-conditionally equivalent, they often exhibit subtle interpretive differences that have proved difficult to capture. Langacker (1995) suggests that raised matrix subjects have a distinct pragmatic function that is roughly equivalent to topicality. Achard (2000) and Ruppenhofer (2004, Chapter 3) test the predictions of Langacker’s hypothesis through corpus-based investigations of raising constructions in French and English, respectively. Their results fail to converge: Achard reports that raised subjects in French are more

frequently discourse-old than their unraised counterparts, while Ruppenhofer finds no effect of activation status in English. Further research is needed to determine whether this reflects a genuine distinction between the two languages or whether it is simply an artifact of discrepancies between the two researchers' methodologies. In Chapter 4, I will return to the relationship between information structure and raising.

At the present time, it is an open question whether activation status plays a direct role in how the arguments of a predicate are mapped to the syntax. This is because the correlation is likely the result of several types of choices that speakers make in presenting information, including verb choice and construction choice.

Extra-Thematic Subjects

In contrast, researchers have argued that information structure plays a direct role in the licensing of “extra-thematic” subjects, which are not selected by any predicate in the sentence. Two distinct phenomena of this sort have been identified: Multiple Subject Constructions (MSCs), introduced in Section 2.2 above, and indirect passives, in which an extra-thematic event participant appears as the subject of a passive construction. According to Li and Thompson (1976), extra-thematic subjects are a hallmark of topic-prominent languages. They are widely taken to be licensed on the basis of their pragmatic function.

MSCs have been investigated in a wide range of languages, including Chinese (Li and Thompson 1976, Xu and Langendoen 1985, Shi 2000, Xu 2001, Pan and Hu 2008),¹⁵ Japanese (Kuno 1973, Heycock 1994, Portner and Yabushita 1998), and Hebrew (Heycock

¹⁵ It is arguable whether “MSC” is the most appropriate label for the Chinese construction under consideration, given Van Valin and LaPolla’s (1997) claim that grammatical functions are not well motivated in Chinese. However, I will continue to use the label for convenience.

and Doron 2003).¹⁶ The following examples are from the Tibeto-Burman language Lahu ((26); Li and Thompson 1976: 468), Hebrew ((27); Heycock and Doron 2003: 5), and Japanese ((28); Kuno 1973: 64).

- (26) hɔ̃ ̄ na-qhô yì ve yò.
 elephant TOP nose long PRT DECL
 ‘Elephants, noses are long.’
- (27) im be’emet dani ha-xavera ʃelo aba ʃela mi-tsarfaf,
 if really Dani the-girlfriend his father her from-France
 ex ze ʃe hu af pa’am lo haya ʃam?
 how it that he never not was there
 ‘If indeed Dani’s girlfriend’s father is from France, how come he was never there?’
 (Lit ‘If indeed Dani his girlfriend her father is from France, ...’)
- (28) kono class-wa dansei-ga yoku dekuru.
 this class-TOP male-NOM well are-able
 ‘This class, the boys do well.’

Many researchers have argued that the extra-thematic subject is licensed to function as a topic (Li and Thompson 1976, Xu and Langendoen 1985, Heycock 1994, Portner and Yabushita 1998, Shi 2000, Xu 2001, Pan and Hu 2008) or as the subject of a categorical judgment, a closely related notion (Heycock and Doron 2003; see Section 4 for discussion of the relationship between topicality and categorical judgments). The mechanism through which licensing occurs varies across accounts. Focusing on the case of Chinese, Xu (2001) proposes that the topical subject is licensed by a functional projection, TOP. Shi (2000) argues for a movement-based account in which all extra-thematic subjects are base-generated within the comment, then moved to a position in the left periphery. Heycock (1994) suggests that licensing involves a syntactic relation of predication that holds between

¹⁶ Landau (2009) argues that the relevant Hebrew construction is actually a form of left dislocation rather than a true MSC. There appear to be compelling arguments on both sides of the debate, but I will set the issue aside here.

the topical subject and the comment. Several of the remaining accounts do not explicitly specify a licensing mechanism.

The indirect passive is best known from studies of Japanese, Chinese, and Korean (Shibatani 1994, Oshima 2006, Otsuka 2006). The following examples come from Shibatani (1994: 464); (29) is Japanese, (30) Chinese, and (31) Korean.

- (29) Taro_o-wa Hanako_{ni} piano_o hik_{are}-ta.
Taro-TOP Hanako-DAT piano-ACC play-PASS-PST
'Taro was adversely affected by Hanako's playing the piano.'
- (30) Wo bei Zhang San toukanle ri_{ji}.
I PASS Zhang San steal.look.ASP diary
'I had my diary read by Zhang San.'
- (31) Na-nun sensayngim-eykey ilum-ul cek_{hi}-ess-ta.
I-TOP teacher-DAT name-ACC write-down-PASS-PAST-IND
'I had my name written down by the teacher.'

These examples illustrate that the subjects of indirect passives are not arguments of the passive verbs. For example, in (29), the arguments of *hik-are-ta* "played" are Hanako and the piano. Taro is linked to the event not through semantic selection, but through the implication that the event affected him adversely. The adversative reading is a common feature of indirect passives across languages, as illustrated by all three examples above. Shibatani (1994) argues that the subjects of indirect passives are licensed on the basis of a relation of *relevance* that holds between the subject denotatum and the event. The implication of adversity is simply one way in which individuals can be pragmatically linked to events in which they do not directly participate; others include possession and benefactive affectedness. Shibatani's relevance requirement is closely related to the definition of topicality that I assume in this dissertation, in which a relation of aboutness links an individual to a proposition.

In Chapters 3 and 4, I will build on this research by arguing that topicality underlies the licensing of extra-thematic subjects in English extraposition constructions, specifically Topical Exclamatives (TE) and Copy Raising (CR). I will propose a novel licensing mechanism in which a construction pairs the function of topic-marking with a change in valence. In Chapter 3, I will return to Shibatani's observation that the relevance requirement can be satisfied through a variety of relationships between the topic and the comment. I will propose a typology of these relationships and show where on the typology TE and CR fall.

3.2.2 Informational Effects on Object Selection

We now turn to the effects of information structure on the selection of objects. There is a considerable body of literature claiming that information structure plays a role in the English dative alternation, illustrated below.

- (32) a. John gave the book to Mary.
b. John gave Mary the book.

In the dative alternation, the verb typically selects for an *agent*, *theme*, and *recipient*, with the latter two arguments participating in the alternation: the theme may be realized as direct object and the recipient as an oblique, as in (32a), or both may be linked as objects, as in (32b). Several researchers have proposed that information structure or closely related pragmatic notions condition the alternation (Green 1974, Erteschik-Shir 1979, Thompson 1990, Goldberg 1995, 2004, Basilico 1998, Polinsky 1998, Arnold et al. 2000, Wasow 2002, Ruppenhofer 2004). A common observation is that recipients in the ditransitive construction (32b) tend to have a wide range of semantic and pragmatic properties typically

associated with topicality. The reader is referred to Ruppenhofer (2004, Chapter 2) for a detailed overview of these claims. Though I will not address the dative alternation in this dissertation, it has certainly proven itself to be an important test case for any interface model of argument realization.

3.2.3 Free Datives

Another phenomenon that must eventually be addressed by the interface model is the family of so-called “free dative” constructions, in which a (typically) dative-marked constituent is not selected by the verb.¹⁷ Free datives are attested in a range of languages, including French (Authier and Reed 1992, Shibatani 1994), German (Shibatani 1994, Lambert 2008), Hebrew (Berman 1982, Borer and Grodzinsky 1986, Landau 1999, Linzen 2008), and some varieties of English (Webelhuth and Dannenberg 2006, Horn 2008). The French example in (33) below is from Authier and Reed (1992: 295); the Hebrew example in (34) is from Linzen (2008: 7). (35) is a line from Bob Dylan’s song “Up to Me,” cited in Horn (2008: 170).

- (33) Le gosse lui a démoli son pull.
 the kid her-DAT has ruined her sweater
 ‘The kid ruined her sweater on her.’
- (34) cavat le-ima sheli et ha-xeder.
 I-painted to-mother my ACC the-room
 ‘I painted my mother’s room (for her).’
- (35) I’ve only got me one shirt left and it smells of stale perfume.

For the present purposes, the key issue is how the dative constituent is licensed, if not via the semantic requirements of the verb. Several approaches have been proposed. Borer and Grodzinsky (1986) and Authier and Reed (1992) claim that free datives are assigned

¹⁷ Free datives have also been called “ethical datives” (as in Shibatani 1994) and “personal datives” (as in Webelhuth and Dannenberg 2006).

thematic roles via special mechanisms. Landau (1999) argues that Hebrew free datives arrive in their surface positions via movement. Webelhuth and Dannenberg (2006) propose that free datives in Southern American English are licensed by a construction that highlights the involvement of the dative-marked denotatum in the event. Finally, Shibatani (1994) argues that the relevance-based account of extra-thematic subjects discussed above should be extended to free datives.

3.2.4 Null Instantiation

It is also widely acknowledged that information structure plays a role in determining whether lexically selected arguments can be omitted. Languages vary with respect to the availability of null instantiation. It is common across grammatical environments in *topic-drop* languages such as Mandarin and German, but in English it is largely restricted to particular syntactic environments, such as the subject position of infinitival clauses, and registers, such as “diary drop” (Huang 1984, Haegeman 1990). Huang (1984) argues that this can be accounted for by a typology in which languages are classified as “hot” or “cool,” the terminology having been proposed by J.R. Ross (1982). In “hot” languages such as English, arguments tend to be overtly realized, and accordingly the comprehender must do relatively little work in computing predicate-argument relations. As arguments are often null in “cool” languages like Mandarin and German, determining their identity requires more work – specifically, closer attention to the activation status of discourse referents and their roles within the wider discourse.¹⁸ Erteschik-Shir (2007: 24) suggests an alternative

¹⁸ Huang’s typology also includes a class of “medium” languages, such as Spanish and Italian, which permit null subjects across clause types but do not generally allow null topics outside of subject position. He argues that the distinction between “hot” and “medium” languages emerges from differences in the richness of

way of viewing the hot/cool distinction: cool languages spell out information structure more fully than hot languages, because in cool languages recoverability can be coded by null realization.

Even in “hot” languages such as English, null instantiation is possible in specific grammatical environments, with appropriate contextual support (Thrasher 1977, Haegeman 1990, Haegeman and Ihsane 2001, Reiman 1994, Goldberg 2000, 2004, 2005, 2006, Ruppenhofer 2004, Scott 2004, Michaelis 2010). Goldberg (2004) provides the following examples of existential null complementation, i.e. unspecified object deletion.

- (36) a. The chef-in-training chopped and diced all afternoon.
b. Tigers only kill at night.
c. Pat gave and gave, but Chris just took and took.

She argues that null instantiation is possible when the covert argument is of low discourse prominence. Though I will not elaborate on the discourse properties of null instantiation here, it is clearly part of a full account of the interface between information structure and argument realization.

3.2.5 Previous Models of the Interface

We now turn to a brief overview of the models of the interface that have previously been proposed. One model, that of Aissen (1999), was discussed at the beginning of this section. To review, Aissen’s account of subject selection in English predicts that the most informationally-prominent argument of a transitive clause should always appear as the

agreement morphology, with “medium” languages having relatively rich agreement. Like the hot/cool distinction, it deals with the extent to which arguments must be recoverable from overt sentence structure. The null subjects of “medium” languages, which co-occur with rich agreement morphology, are intermediate with respect to recoverability between the overt arguments of “hot” languages and the null arguments of “cool” languages, which are not necessarily supported by agreement morphology.

subject. I argued that this constraint is too strong, as we do find active clauses with topical objects and non-topical subjects. Similar objections apply to Tomlin's (1985) model, which advocates a purely pragmatic process of subject selection. Tomlin (1985: 61) claims that "subject encodes thematic information over agent ... when several NPs compete for the subject relation, the one representing the most thematic information will win." He supports this view with data from play-by-play in hockey games that indicates that subjects have a strong tendency to be thematic, on his idiosyncratic definition. For Tomlin, thematic information is whatever is most germane to the communicative goals of the discourse. Thematicity is determined on the basis of the general aims of the discourse, "essentially independently of any actual text" (p. 69). He proposes that in hockey play-by-plays, reference to players with the puck is more thematic than reference to the puck itself, which in turn outranks reference to players without the puck. This notion of thematicity differs significantly from standard informational perspectives on topicality, which address how information is evaluated and stored within a particular context. Regardless, he acknowledges that there exist sentences with subjects that are not thematic (p. 72), which demonstrates that the basis of subject selection cannot be entirely pragmatic.

According to Lambrecht's (1995) account, both semantic and informational prominence influence subject selection in English. He proposes that alongside a semantic hierarchy, there also exists a pragmatic hierarchy that links informational relations to grammatical functions. In his pragmatic hierarchy, topic outranks focus for realization as subject. The model differs from the purely pragmatic approaches in that both hierarchies can be violated. Just as a patient can be realized as subject, so too can a focus expression.

However, his account strongly penalizes doubly-marked associations between the three levels of representation. For example, he claims that semantically prominent subjects cannot be focal. This constitutes a doubly-marked association because focal expressions are highly marked as subjects (per Lambrecht's pragmatic hierarchy) and semantically prominent arguments are marked as focus expressions. Support for this claim comes from the distribution of arguments associated with the verb *hurt*. This is illustrated below (adapted from Lambrecht 1995: 161).¹⁹

- (37) a. That slap HURT. (*instrument?*)²⁰
 b. My foot still HURTS. (*location*)
- (38) a. # That SLAP hurt.
 b. My FOOT hurts.
- (39) a. That slap hurt my foot.
 b. # My foot hurt that slap.

Lambrecht takes the subjects in (37) to be topical, due to the prosodic pattern in which primary stress falls on the predicate. In (38), the subjects are prosodically marked as focal. (37) and (38b) demonstrate that subjects may be either topical or focal. However, there is a three-way interaction between syntax, semantics, and information structure: so-called “instruments” are acceptable as topical subjects (37a) but not as focal subjects (38a). According to Lambrecht, this is due to the relative semantic prominence of *instrument* with respect to *location*, as shown by the fact that (39a) but not (39b) is a well-formed argument realization pattern in English.

¹⁹ In examples involving focus I follow the convention of using capital letters to indicate the focus expression, which is prosodically prominent.

²⁰ The semantic roles assigned to the subjects come directly from Lambrecht (1995), including the uncertainty about the appropriateness of the role of *instrument* in the (a) sentences.

There are several aspects of Lambrecht's account that I find to be problematic. The first is that it is not clear how the semantic and pragmatic hierarchies interact, e.g. which level determines subject selection when the two hierarchies are in conflict. Another problem is that the source of the penalty on doubly-marked associations is not made explicit in the model, and in fact depends on implicit associations between pragmatic and semantic roles. This dissertation builds upon Lambrecht's observation that information structure influences argument realization within active sentences, providing a model that specifies exactly how it takes place.

4. Elements of Information Structure

The discussion thus far has been vague as to which aspects of information structure influence argument realization. As indicated in Chapter 1, I assume that there are two major components of information structure. One tracks the activation status of discourse referents: the degree to which they are salient in the consciousness of the speaker and hearer at the time of a particular utterance (e.g. Prince 1981a, Ariel 1990, Birner and Ward 1998). The other component is a set of pragmatically structured propositions, in which propositional content is partitioned according to the discourse representations of the speaker and hearer (Reinhart 1981, Prince 1986, Lambrecht 1994, Birner and Ward 1998, Endriss 2009). In this dissertation, I propose that the content of pragmatically structured propositions has the capacity to drive argument realization. I have yet to find similar effects for information status. For this reason, I will set activation status aside and focus on the properties of pragmatically structured propositions.

Pragmatically structured propositions are partitioned on the basis of *topic* and *focus*. Both terms have a long history in the literature on information structure and have been used in a variety of ways. I adopt Lambrecht's (1994) proposal that both topic and focus are informational relations. Topic relates a proposition to a relevant individual or set (the *topic expression*) while focus relates a salient open proposition to a denotatum that combines with it (the *focus expression*). Neither topic nor focus is defined in terms of any aspect of sentence structure, though they are often marked formally in specific constructions. In English, for example, focus tends to be marked prosodically while topic is often expressed by constructions with detached initial NPs. The classification of an expression as a topic or focus does not hinge on its information status. Topic expressions do not need to be previously activated, and focus expressions are not required to be new referents, or – as we will see – referents at all. Furthermore, topic and focus are essentially independent. Crucially, they are *not* in complementary distribution: a particular constituent may be topical, focal, neither, or even both.

Although I adopt a pragmatic/informational approach to topic and focus here, it must be acknowledged that both relations have reflections in the semantics. Partee (1992) argues that the choice of topic and focus influences quantificational structure, with topics being mapped into the restrictive clause and focus being represented in the nuclear scope. This is one possible explanation for the clear truth-conditional effects of focus; for a review, see Beaver and Clark (2008). Whether topic choice influences truth conditions is a more controversial question, one which Endriss (2009) answers in the affirmative. In the discussion that follows, I will briefly review the evidence that each relation has truth-

conditional effects. I will continue to treat topic and focus as informational phenomena, with the understanding that an account that captures their semantic aspects is ultimately required.

This section surveys several influential approaches to topic and focus that have been proposed in the literature. It expands on the assumptions that I make in this dissertation and details the ways in which topic and focus relations can be identified in linguistic data. It then moves to the question of how the relations are encoded in formal sentence structure, focusing on the ways in which they are marked syntactically in English. Finally, we turn to the relationship between topic, focus, and grammatical subjecthood, touching not only on the long tradition of research linking subjecthood to topicality but also the less investigated relationship between subjecthood and focus. Throughout this section I follow the convention of using the terms *topic* and *focus* to refer not only to the relations themselves, but also to the distinguished elements of topic- and focus-marking structures, the topic/focus expressions.

4.1 Topic

The view of topicality as an aboutness relation between an individual and a proposition originated with Strawson (1964) and was developed further by Reinhart (1981) and Lambrecht (1994), among others. Lambrecht gives the following definition (p. 131): “a referent is construed as being the topic of a proposition if in a given situation the proposition is construed as being about this referent, i.e. as expressing information that is relevant to and which increases the addressee’s knowledge of this referent.” Topic relations

shape the developing discourse representations of the speaker and hearer. Strawson (1964) proposes two ways in which this occurs. First, topical referents provide a point of entry for evaluating the truth of the associated proposition. For example, he claims that the truth of *all crows are black* is evaluated via the topic expression (*all*) *crows*, as we tend to verify the statement by inspecting the set of crows to determine if all are black, rather than inspecting the set of non-black entities to check whether it includes any crows. Second, topic relations serve to organize the set of propositions contained within the discourse representation. Reinhart (1981:62) illustrates this with the pair of sentences in (40), which express identical propositions that she claims are stored differently in the interlocutors' discourse representations.

- (40) a. Rosa goes out with Felix.
 b. Felix goes out with Rosa.
- (41) I'm not sure which facet of this pairing blows my mind more: that Billy Corgan is dating Tila Tequila or that Tila Tequila is going out with Billy Corgan.²¹

Due to the tendency to interpret subjects as topical, the speaker's utterance of (40a) serves as an instruction to the hearer to store the proposition alongside other information that is known about Rosa; (40b) functions as a statement about Felix. The naturally-occurring sentence (41) is a more colorful illustration of this point, in which the speaker reports on two ways of evaluating the same proposition: once on the basis of what she knows of Billy Corgan, and once on the basis of her knowledge of Tila Tequila.

Reinhart also proposes a means of formally representing the topic relation. Her proposal builds on Stalnaker's (1978) notion of the *context set*, the set of possible worlds that is compatible with the set of propositions taken for granted by the speaker and hearer

²¹ <http://gofugyourself.celebuzz.com/>, accessed 4/7/2009

(the *common ground*) at the point of utterance. In her model, topics are a means of organizing the common ground. To utter a sentence that marks a topic relation is to make a *pragmatic assertion* linking a proposition to a topical referent. A pragmatic assertion leads to the storage of the proposition under the “file card” associated with the topic expression (cf. Heim 1982). She formalizes the notion of pragmatic assertion in the following way (p. 80-81).

(42) $PPA(S) = \emptyset$ together with $\{ \langle \alpha, \emptyset \rangle : \alpha \text{ is the interpretation of an NP expression in } S \}$

(42) says that the set of *possible pragmatic assertions* (PPA) associated with a sentence S expressing the proposition \emptyset consists of (1) the simple proposition \emptyset , which is not associated with a topic, and (2) a set of pragmatically structured propositions of the form $\langle \alpha, \emptyset \rangle$, where α is a topic expression. Portner and Yabushita (1998) and Endriss (2009) adopt a similar means of representing topicality, but argue that the relation should be situated within the semantic component rather than in the pragmatics.

4.1.1 Identifying Topics

Though many researchers have shared the intuition that topicality is defined by aboutness, there has been considerable debate as to how to identify topics on this basis. Aboutness itself cannot be verified simply by investigating a particular utterance in context, as this would require direct access to the discourse representations of the interlocutors. For this reason, researchers have looked for more concrete methods of identifying topics. Three major approaches have emerged: investigating the distribution of the hypothesized topic expression in the discourse, evaluating whether it is compatible with specific

morphosyntactic forms believed to mark topicality, and determining whether the grammatical environment in which it appears is topic-marking.

Activation Status and Persistence

The first approach hinges on the assumption that aboutness at the sentence level is correlated with the role that the topic expression plays in the wider discourse. One discourse correlate that has been proposed is the activation status of the topical referent at the time of utterance. Another is its persistence into the subsequent discourse. The simplest hypothesis, that topics must be previously activated or discourse-old, has been disconfirmed (Reinhart 1981, Lambrecht 1994, Michaelis and Francis 2007; though see Gundel 1985 and Portner and Yabushita 1998 for an opposing view). In the following example, based on Reinhart (1981: 66), the topical subject has a discourse-new denotatum.²²

(43) A woman I know considers herself a good Christian.²³

However, some researchers have argued that there should be a significant connection between aboutness and activation status, on the grounds that highly activated referents serve as discourse topics and therefore are likely to be the topics of individual utterances. Lambrecht and Michaelis (1998) and Michaelis and Francis (2007) frame this connection in terms of *predictability*, arguing that topics are predictable arguments of predication. Psycholinguistic research by Hornby (1972), discussed in Gundel (1985: 89), supports this

²² On Prince's (1981a) approach to classifying expressions with respect to activation status, the subject of (43) is *brand-new anchored*, indicating that the referent itself is discourse-new, but bears a relationship to an established discourse referent, namely the speaker.

²³ <http://crooksandliars.com/node/22248>, accessed 11/24/2009

view. Hornby's study demonstrates that people are predisposed to interpret activated referents as the topics of subsequent utterances. As we have seen, correlations between activation status and topicality have been used to argue that particular constructions are (or are not) topic-marking. For example, Achard (1990) reports that subject NPs in the French subject-to-subject raising construction are disproportionately discourse-old and concludes that the construction marks a topical subject. However, because the correlation is imperfect, the activation status of a referent cannot determine conclusively whether it is topic of a particular utterance.

The flip side of a referent's activation in the preceding discourse is its persistence into the subsequent discourse. We assume that sentence topics are relatively likely to persist as topics of the discourse that follows. Michaelis and Francis (2007) use this reasoning to argue that lexical NP subjects in English, whose referents systematically persist in subsequent discourse, are a means of introducing new topics. Again, this diagnostic applies to constructions rather than to individual utterances, as there is no guarantee that a particular aboutness topic will persist. Furthermore, Michaelis and Francis's research illustrates that the connection between topicality and activation status is complex even at the constructional level. They argue that lexical NP subjects denote topics that are discourse-new, suggesting that topic-marking structures can diverge from the expected correlations with activation status.

Topic-Marking Constructions

The second major approach to identifying topics consists of checking the compatibility of the putative topic expression with formal structures believed on independent grounds to be

topic-marking (Reinhart 1981). Reinhart (1981) proposes two diagnostic constructions for English: *As for X* constructions, in which X marks a topic, and *prolepsis*, in which the topic is contained within an adjunct prepositional phrase. Her approach is to take a hypothesized topic relation, place its constituents within the topic-marking construction, and judge whether the result is felicitous. An example is given below (p. 64-65).

- (44) a. Kracauer's book is probably one of the most famous ever written on the subject of the cinema. Of course, many more people are familiar with the book's catchy title than are acquainted with its turgid text.
b. *As for* this book, many more people are familiar with its catchy title than are acquainted with its turgid text.
c. He said about this book that many more people are familiar with its catchy title than are acquainted with its turgid text.

Reinhart's hypothesis is that the topic of the second sentence in (44a) is the book, rather than the subject NP *many more people*. It is supported by the fact that the topic relation is compatible with the *As for X* construction in (44b) and with *prolepsis* in (44c). Thus (in)compatibility with specific formal structures can support or fail to support hypotheses about topicality.

Reinhart notes, however, that this technique has limitations. Ward (1985: 23) concurs, demonstrating that the *As for X* construction fails to accommodate a significant portion of the NPs that can be fronted by topicalization and allows NPs that cannot be topicalized. There are two ways in which constructional diagnostics may go wrong. First, the diagnostic constructions may have additional use conditions beyond topicality and as a result may be incompatible with perfectly good topics. Second, the diagnostic construction may not be topic-marking at all. Clearly, it is essential that we be able to show that diagnostic

constructions in fact do mark topics.²⁴ This essentially brings us full circle. How do we know whether a given construction marks topicality?

Again, Reinhart (1981) provides a possible answer. She states that “the choice of a noun phrase as the topic of a given sentence is sensitive to the semantic properties of the NP itself” (p. 65). NPs that are grammatically marked as topics must be interpreted in a way that is consistent with the pragmatics of topicality. In other words, topics are constrained by *pragmatic type restrictions*.

Pragmatic Type Restrictions

Topic expressions must refer to a specific individual or set of individuals (Reinhart 1981, Davison 1984, Cresti 1995, Portner and Yabushita 1998, 2001, Endriss 2009).²⁵ Reference to an individual or set is an essential component of the topic relation as I have defined it: an aboutness relation that links an individual or set to a proposition. Expressions that do not refer to an individual or set are barred from the distinguished position of topic-marking constructions. This can be seen through the distribution of copular clauses in prolepsis, which we have hypothesized is a topic-marking construction. There are two types of copular clauses: *predicational* clauses, which ascribe a property to the subject NP (45a) and *specificational* clauses, which specify the value of a predicate that appears in subject position (45b) (Akmajian 1979, Heycock and Kroch 1999, 2002, Mikkelsen 2004).

- (45) a. Obama is the most popular candidate.
b. The most popular candidate is Obama.

²⁴ The problems associated with diagnostic constructions lead Vallduví (1992: 33) to reject the idea that aboutness topics can appear outside of sentence-initial position.

²⁵ This generalization extends to topical indefinites, which are interpreted as specific (Cresti 1995, Portner and Yabushita 2001).

According to Mikkelsen (2004), subjects of predicational clauses denote individuals or sets, while subjects of specificational clauses denote properties. As subjecthood is not marked as topical in simple subject-predicate sentences in English, both individuals (45a) and properties (45b) are possible subjects. However, the topic expression in prolepsis must denote an individual or set, as in (46a), rather than a property, as in the infelicitous (46b).

- (46) a. He said about Obama that he is the most popular candidate.
b. # He said about the most popular candidate that he's/it's Obama.

The requirement that topic expressions refer to a *specific* individual or set is a consequence of the function that topicality serves in organizing the discourse. The speaker's choice of a topic-marking construction serves as an instruction to the hearer to evaluate and store the proposition under the organizational heading provided by the topic expression. It follows from this that the speaker must have the referent of the topic expression in mind at speech time, and thus that she must be referring to it specifically. This corresponds to Kripke's (1977) notion of *speaker reference*, as opposed to semantic reference. It also connects to Ladusaw's (1994: 3) claim that that specific interpretations are presuppositional in that the speaker's mind "must be directed first to an individual, before the predicate can be connected to it." However, they need not be presuppositional in the sense of previously activated in the discourse (contra Gundel 1985, Portner and Yabushita 1998). When expressions that are interpreted specifically are not previously activated, they are often understood as a proper subset of a previously activated set. This results in a *partitive* reading of quantified expressions. In the partitive reading of (47), the three ships

of the second sentence are taken to be a proper subset of the five sets of the first (example from Frazier et al. 2005: 202).²⁶

(47) Five ships appeared on the horizon. Three ships sank.

This perspective on the specificity requirement has parallels in the research on *categorical* and *thetic judgments*, which originated with the philosophers Brentano and Marty and was developed within the linguistics literature by Kuroda (1972, 1992), Babby (1980), Sasse (1987), Horn (1989) Ladusaw (1994), Rosengren (1997), Basilico (1998), and Heycock and Doron (2003). Categorical and thetic judgments are characterized by a difference in conceptualization that is reflected in their propositional structures. Thetic judgments have one component: affirmation or denial of judgment material, i.e. that an entity exists or that a state of affairs holds. Categorical judgments have two components: recognition of an entity and assertion that a property holds of that entity. The categorical judgment is essentially equivalent to a proposition that is pragmatically structured by topicality, as the “recognition of an entity” is comparable to the act of making specific reference to the topic expression. The major difference between the two constructs is that categorical judgments are represented as structured propositions at the semantic level, whereas the topic relation has often been represented at the pragmatic level (though see e.g. Portner and Yabushita 1998 and Endriss 2009 for approaches in which topicality is situated within the semantics).

As Kuroda (1972) and Ladusaw (1994) demonstrate, structures that mark categorical judgments place restrictions on the interpretation of the topic of the judgment. Kuroda

²⁶ An additional property of specific expressions is that they are typically associated with wide scope. Endriss (2009) provides an overview of current perspectives on the connection between specificity and scope.

(1972) argues that the suffix *-wa* marks a categorical judgment in Japanese and demonstrates that the *wa*-marked expression must be interpreted as specific. Ladusaw (1994) shows that categorical judgments are compatible only with the strong readings of quantifiers, in which specific reference is made to an individual or set. This is illustrated by the following examples, adapted from Ladusaw (1994).

- (48) a. The cows are intelligent
b. * *Sm* cows are intelligent.
c. Cows are intelligent.
d. Cows are available.

Ladusaw's account additionally incorporates Carlson's (1977) distinction between *individual-level* predicates, roughly defined as those that hold of individuals across space and time (e.g. *intelligent*), and *stage-level* predicates, which hold of individuals in particular situations (e.g. *available*). He proposes that individual-level predicates are compatible only with categorical judgments. Therefore, because the predicate *intelligent* in (48a-c) is individual-level, the sentences are taken to be categorical judgments, involving specific reference to the subject NPs. This is felicitous in (48a) because definite NPs receive strong (specific) readings. In contrast, (48b) is ruled out because the phonologically reduced quantifier *sm* ("some") allows only a weak reading, which involves existential commitment but no specific reference.

Examples (48c-d) show that the interpretation of bare plural NPs, which are ambiguous between a weak existential reading and a strong reading making reference to kinds, is influenced by judgment type. The categorical judgment in (48c) requires a strong reading of the subject NP, while thethetic judgment in (48d) involves a weak reading (cf. Laca 1990, Cohen and Erteschik-Shir 2002, Heycock and Doron 2003). I assume, following

Carlson (1977), that kinds are a type of individual, and thus that the strong reading of bare plurals involves reference to a specific individual. The class of bare plurals is one of several referentially ambiguous NP types that receive specific interpretations when they appear as topic expressions; others include mass nouns, indefinites and numerically quantified NPs (Ladusaw 1994, Cresti 1995, Portner and Yabushita 2001, Krifka et al. 2006, Endriss 2009).

The contrast between (48c) and (48d) shows that the syntactic form of simple subject-predicate sentences is neutral with respect to judgment type. It is the main predicate that determines whether the sentence is taken to express athetic or categorical judgment. However, there are syntactic constructions, such as prolepsis, that impose precisely the same type of interpretive restrictions as individual-level predicates do. The distinguished argument in prolepsis must be interpreted specifically as in (49a), not existentially as in (49b).

- (49) a. He said about cows that they are intelligent.
b. # He said about cows that they are available.

Heycock and Doron (2003) make similar observation about what they claim is a MSC in Hebrew.²⁷ In this construction, the first subject must be interpreted as the topic of a categorical judgment and as such is sensitive to pragmatic type restrictions. They show that bare plurals, which can be interpreted either specifically or existentially in simple subject-predicate structures (50a), must receive the specific reading in the MSC (50b).

²⁷ Landau (2009) debates this claim, arguing that the construction in (50b) is a form of dislocation. This is not crucial for the present point.

- (50) a. hacagot tovot 'olot midey pa'am.
 plays good perform every now and then
 Reading 1: 'Good plays are such that they are performed every now and then.'
 Reading 2: 'There are good plays performed every now and then.'
- b. hacagot tovot ma'alim otan midey pa'am.
 plays good they-perform ACC-them every now and then
 Reading 1 only: 'Good plays are such that they are performed every now and then.'

Because pragmatic type restrictions cut to the core of topicality, rather than tapping into its many correlates, I consider them the gold standard for identifying topic-marking constructions. In this dissertation, I assume that a construction is topic-marking if and only if it requires that specific reference be made to the topic expression. The activation status of the topical referent and its persistence in the discourse are simply correlates of topicality. Therefore, they can lend support to a claim that a particular construction is topic-marking but can neither prove nor disprove it. Relying on pragmatic type restrictions provides us with a clear means of identifying topic-marking constructions, but it admittedly does not shed much light on the issue of how to determine whether a given NP in a non-topic-marking construction functions as a topic. I will set this issue aside.

4.1.2 Topics, Predication, and Frame-Setting

The definition of topic-marking that I adopt here diverges considerably from Jacobs's (2001) proposal to account for the pragmatic diversity of so-called "topic-marking" constructions by splitting topicality into four properties: *informational separation*, *addressation*, *predication* and *frame-setting*. He defines these properties in the following way. Informational separation is delimitation at the informational level between the topic and the relevant proposition – essentially, a pragmatically structured proposition. Addressation corresponds to an aboutness relationship, a claim that propositions are stored and

evaluated with respect to their topics. Predication requires that the topic referent be an element of the associated proposition, the comment. Frame-setting is a type of domain restriction in which one constituent provides a domain in which an associated proposition holds. The following example, from Krifka (2006: 36), illustrates frame-setting: B's response indicates that the proposition *John is fine* applies in the domain of his health.

- (51) A. How is John?
B: Healthwise, he is fine.

In Jacobs's model, topicality is a prototype, with typical topic-marking constructions exhibiting all four properties and less typical constructions exhibiting fewer. In contrast, I assume that true topic-marking constructions always exhibit informational separation and aboutness, which are fundamental components of topicality. They do not necessarily require that the topic expression be an object of predication. This is illustrated by prolepsis. In the typical case (52a), the topic denotatum is an argument of the comment: the topic expression, *the class*, is assigned the property of being a waste of time. However, examples like (52b) show that this is optional: the topic expression in this sentence is not an argument of the comment.

- (52) a. He said about the class that it was a waste of time.
b. He said about the class that three people managed to finish the problem.

Instead, it stands in a set-subset relationship with an argument, *three people* (at least on one metonymic interpretation of *the class*). While predication is not required, there is generally an inferable semantic connection between the topic and some element of the proposition. This is due to the requirement that topics be relevant. Some link between the topic and the proposition is necessary in order for the two to be associated pragmatically (cf. Shibatani

1994). Chapter 3 will further explore the ways in which semantics can ground the topic-comment link.

I assume that Jacobs's final property, frame-setting, is fundamentally distinct from topicality, as I have defined it. He associates the two notions on the basis of the observation that several constructions that have been hypothesized to mark topicality, such as the suffix *-wa* in Japanese, include frame-setters among the class of possible "topics." Yet some such frame-setters, such as *professionally* in (53a), are clearly not aboutness topics – to begin with, they do not make specific reference. The essential distinction, though, is that frame-setters obligatorily contribute to the content of the associated proposition, while topic expressions sometimes do not.

- (53) a. Professionally, John is pleased with Mary.
b. As for John, he is pleased with Mary.

In (53a), *professionally* is an element of the proposition, which is that John is pleased with Mary's professional performance. In contrast, the initial topic expression in (53b) does not make any independent contribution to the proposition. Thus frame-setters influence the interpretation of propositions in a way that topics do not, a point made by Jacobs (2001: 656). On the basis of (53a), it does not follow that John is pleased with Mary *simpliciter*; on the basis of (53b), it does.

In spite of the distinction I am drawing here, frame-setters and topics do have a common function, namely delimiting a conceptual domain in which a proposition is interpreted. It is likely for this reason that the two constructs are conflated by certain morphosyntactic structures, such as *wa*-marking in Japanese. The nature of the delimitation, however, is fundamentally different.

4.1.3 Topicality and Truth Conditions

This leads us back to the question of whether topicality can have truth-conditional effects. Some researchers, such as Endriss (2009), have claimed that it can, although it is clear that many of the apparent truth-conditional effects of topicality are of a different nature than the more straightforward effects associated with focus. As is well-known, in the environment of focus-sensitive particles such as *only* the choice of focus affects truth conditions: (54a) is necessarily false if Bill introduced John to Mary, while (54b) is not. In contrast, topic choice does not usually influence truth-conditions; (55a-b) are semantically equivalent.

- (54) a. Bill only introduced John to SUE.
b. Bill only introduced JOHN to Sue.
- (55) a. He said about Felix that he goes out with Rosa.
b. He said about Rosa that she goes out with Felix.

The apparent truth-conditional effects of topicality that have been observed relate not to topic choice but to the way in which topic expressions must be interpreted. Strawson (1964: 95) argues that topicality influences truth conditions in the event of reference failure. He judges that when an expression such as *the king of France* fails to refer, (56a) results in a truth-value gap (a proposition that is neither true nor false), while (56b) is simply false.²⁸

- (56) a. The king of France visited the exhibition.
b. The exhibition was visited by the king of France.

He attributes this to the interpretive effects of topicality. According to Strawson, topic expressions must successfully refer in order for the truth of the associated proposition to be evaluated. Because subjects are preferentially interpreted as topics, *the king of France* is the

²⁸ See Horn (1986) for an alternative account of Strawson's contrast.

default topic of (56a) while *the exhibition* is the topic of (56b), hence the differences in interpretation.

In addition, as we have seen, topic-marking constructions require that specific reference be made to the topic expression. For this reason, non-referential expressions such as predicate nominals cannot be topics and referentially ambiguous NPs must receive a specific interpretation. Furthermore, Endriss (2009) claims that topic expressions obligatorily receive wide scope. She provides evidence (p. 213-215) that the set of quantifiers that are compatible with topicality is identical to the set of quantifiers that permit exceptional wide scope (i.e. scope out of islands). For example, the strong readings of *ein* 'some' and *drei* 'three' are possible in the topic position of left dislocation in German (57a-b); they can also scope outside of *if*-clauses, which are known to be scope islands (58a-b). However, *jedes* 'every' and *kein* 'no' are incompatible with topicality (57c-d) as well as exceptional wide scope (58c-d).

- (57) a. EIN Pferd, das hat die Bananen gefressen.
some horse it has the bananas eaten
'Some (specific) horse has eaten the bananas.'
- b. DREI Pferde, die haben die Bananen gefressen.
three horses they have the bananas eaten
'Three (specific) horses have eaten the bananas.'
- c. * Jedes Pferd, das hat die Bananen gefressen.
every horse it has the bananas eaten
'Every horse has eaten the bananas.'
- d. * Kein Pferd, das hat die Bananen gefressen.
no horse it has the bananas eaten.
'No horse has eaten the bananas.'
- (58) a. Clarissa wundert sich, wenn ein Pferd krank wird.
Clarissa wonders herself if some horse ill become
'Clarissa will be surprised if some horse falls ill.'
[if > some, some > if]

- b. Clarissa wundert sich, wenn drei Pferde krank werden.
Clarissa wonders herself if three horses ill become
'Clarissa will be surprised if three horses fall ill.'
[if > three, three > if]
- c. Clarissa wundert sich, wenn jedes Pferd krank wird.
Clarissa wonders herself if every horse ill become
'Clarissa will be surprised if every horse falls ill.'
[if > every, *every > if]
- d. Clarissa wundert sich, wenn kein Pferd krank wird.
Clarissa wonders herself if no horse ill become
'Clarissa will be surprised if no horse falls ill.'
[if > no, *no > if]

As I have argued, the requirement that topic expressions be referential and specific is a direct consequence of the functions that topics perform in organizing the discourse. Thus, it can be given a pragmatic explanation. An intention on the part of the speaker to make specific reference to the topic expression can be seen as part of the use conditions for topic-marking constructions. In contrast, the scope effects that Endriss observes genuinely intrude into truth-conditional semantics, and appear to resist a pragmatic account: Endriss argues convincingly that exceptional wide scope cannot be reduced to a (potentially) pragmatic notion such as reference or specificity. It is clear that any fully elaborated account of topicality requires an understanding of its truth-conditional effects.

4.2 Focus

We now turn to focus. I adopt a slightly modified version of Lambrecht's definition of focus (1994: 213): "the semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition." The *assertion* is the proposition that the speaker wishes to enter into the common ground, while what Lambrecht calls the (pragmatic) presupposition is the portion of that proposition that is already familiar to the

speaker and hearer. The following example illustrates the components of the focus relation (Lambrecht 1994: 209).

- (59) John: Where did you go last night?
Mary: I went to the MOVIES.

In Lambrecht's terminology, John's question introduces a presupposition that the addressee, Mary, went somewhere last night. Her response asserts that she went to the movies. Focus is the relation that holds between the assertion and the presupposition. The privileged component of this relation is the *focus expression* or simply the *focus*: the component of the assertion that is not contained within the presupposition. The focus expression in Mary's response is *the movies*, because this is precisely the information that is not presupposed.

The slight terminological modification that I make to Lambrecht's definition is as follows. What he refers to as a presupposition I will call an *open proposition*, following a tradition of research in the Penn school of information structure (e.g. Prince 1986, Vallduví 1992, Birner and Ward 1998, Birner et al. 2007). An open proposition has one or more variables that have yet to be resolved. For example, the open proposition associated with the question in (59) is *you went to X last night*, where X is a variable. This is of course quite similar to the presupposition *you went somewhere last night*, but the two constructs are different types of semantic objects. Presuppositions are complete propositions and thus can be true or false, believed or disbelieved. Open propositions, in contrast, are incomplete semantic objects, lacking truth values and inappropriate as objects of belief. There is reason to believe that salient open propositions, rather than presuppositions, underlie the focus relation. Dryer (1996: 487) discusses examples similar to the following.

- (60) A: Who came to the lecture?
B: NOBODY came.

Presuppositions are usually taken to be part of the common ground, i.e. believed by speaker and hearer.²⁹ If this is the case, and if John's question presupposes that somebody came to the lecture, Mary's response should be infelicitous. In fact, it is not. On the basis of examples of this sort, Dryer (1996) argues that the "background" material in the focus relation is not presupposed but *activated*, that is, a salient part of the interlocutors' discourse representations at the time of utterance. For these reasons, I will assume that background material has the form of an activated open proposition.

4.2.1 Identifying Focus

We now turn to the question of how to identify focus expressions. I assume that focus is defined by the way it structures propositional content, and not by any formal property. Informational focus has a certain degree of independence from its prosodic correlates. This can be seen first through phenomenon of *focus projection*, in which prosodic stress on one constituent reflects informational focus of a larger constituent (e.g. Selkirk 1984, Büring 2006). Consider the following examples. Mary's response, with stress on *movies*, is acceptable both as an expression of narrow focus, as in (61), and broad focus, as in (62). If Mary's response were uttered in isolation, it would be difficult to determine which was intended.

- (61) John: Where did you go last night?
Mary: I went to the MOVIES.

²⁹ Lambrecht (1994: 52) adopts an idiosyncratic definition of presupposition in which presupposed content may be mentally represented but not believed. For this reason, the modification I propose is more terminological than substantial.

- (62) John: What did you do last night?
Mary: I went to the MOVIES.

Information status also fails to provide the answer. Just as topics do not always denote previously activated referents, focus expressions do not necessarily introduce new referents into the discourse. Lambrecht (1994: 211) gives the following example.

- (63) John: Where did you go last night, to the movies or to the restaurant?
Mary: We went to the RESTAURANT.

John's question raises both *the movies* and *the restaurant* to a high level of activation. Nevertheless, *the restaurant* is the focus of Mary's response because it is the part of her assertion that differs from the background. As Lambrecht emphasizes, what is "new" in an utterance expressing focus is not the focus expression itself, but its relation to the background information.

Furthermore, in contrast with topic, focus expressions need not be referential at all. Focus expressions generally permit existential readings of referentially ambiguous NPs, as in (64). In addition, the set of possible denotata for focus expressions is not restricted to individuals or sets. For example, focus expressions may denote predicates, as shown in (65).

- (64) John: What did you see at the zoo?
Mary: We saw IGUANAS.
(65) John: Did you pet the iguanas?
Mary: No, but we FED them.

As Krifka (2006: 8) demonstrates, it is even possible for focus expressions to be formal linguistic objects, rather than denotata. In examples like (66) and (67), the focus does not contribute new content to the discourse, but rather conveys the speaker's preference for an alternative lexical or phonological form.

- (66) Grandpa didn't kick the BUCKET, he passed AWAY.

- (67) A: They live in BERlin.
B: They live in BerLIN!

Thus there appears to be no diagnostic for focus expressions comparable to the pragmatic type restrictions associated with topic.

Instead, focus expressions are identified through their distribution in the discourse. Krifka (2006) makes the point that focus relates to the prior discourse in a more direct way than topic. Utterances involving focus respond to the immediately preceding discourse by resolving a variable of a salient open proposition. In the typical case, the open proposition is overtly introduced into the discourse, as in the question-answer pairs that we have seen so far, or else can be inferred from the prior discourse, as in (68). When a focus-marked structure occurs in the absence of a salient open proposition, one must be inferred. It is for this reason that (69) is acceptable as a discourse-initial utterance only to the extent to which the open proposition *x saw iguanas at the zoo today* can be accommodated. (Generally, it cannot and for this reason (69) is usually infelicitous.)

- (68) John: I'd love to hear more about the zoo trip.
Mary: We saw IGUANAS.
(69) WE saw iguanas at the zoo today.

In contrast, topic is less intimately linked to the prior context. It is perfectly acceptable to start a conversation with (70), where prolepsis indicates that Rosa is to be interpreted as the topic.

- (70) So, I learned about Rosa that she goes out with Felix.

For this reason, the diagnostic that I use to identify focus in this dissertation is context-sensitive. An expression is focal if it felicitously follows a corresponding open proposition, or if uttering the sentence that contains it triggers such an open proposition. This very

general diagnostic will be adapted in order to accommodate the additional constraints associated with the constructions under investigation.

In contrast with topic, an assertion must have a focus in order to be informative. It is also possible for an utterance to have more than one focus expression. Example (71), in which the response resolves two variables of the open proposition introduced by the question, illustrates this.

- (71) A: Who did John give what?
B: He gave a BALL to MARY.

4.2.2 Truth-Conditional Effects of Focus

We now return to the semantic effects associated with focus. As we have seen, focus choice influences truth conditions in the environment of focus-sensitive particles such as *only*, as well as in many other environments (see Beaver and Clark 2008 for a review). Lambrecht's definition of focus, however, is designed to account for the informational effects of focus and is not directly concerned with its intrusions into propositional structure. There does exist a formal semantic model of focus that is in many ways comparable to Lambrecht's, namely the theory of Structured Meanings (von Stechow 1990, Krifka 2001). In this model, focus is a structured proposition that pairs a property formed by lambda extraction with a denotatum. A simplified representation of Sue's response in (72a) appears in (72b).

- (72) a. John: What did Mary see at the zoo?
Sue: She saw IGUANAS.
b. $\langle \lambda x.see(mary, x), iguanas \rangle$
c. John: What did Mary see at the zoo?
Sue: She only saw IGUANAS.
d. $\langle \forall x.see(mary, x) \rightarrow x = iguana \rangle$

The property in a structured meaning roughly corresponds to the open proposition in Lambrecht's model; the paired denotatum corresponds to the focus expression. The crucial difference is that structured meanings are semantic objects rather than pragmatic ones, and are thus sensitive to semantic operations such as quantification. Structured meanings account for truth-conditional effects of focus in the environment of *only* in the following way. *Only* induces universal quantification over the set of structured meanings: the interpretation of (72c) is that all *x such that Mary saw x at the zoo* are *iguanas* (72d). In this dissertation, I adopt Lambrecht's informational model of focus while acknowledging that something like Structured Meanings is necessary in order to account for the interaction between focus and quantification.

We now move briefly to another influential formal semantic model of focus, Alternative Semantics (Rooth 1986, Beaver and Clark 2008). In this model focus-marked utterances have two levels of meaning: the normal propositional level and the level of focus interpretation, consisting of pairings of the background material with alternative denotata. For example, Sue's response in (72b) expresses the proposition *we saw iguanas* alongside a set of alternatives triggered by the focus expression, e.g. [*we saw iguanas, we saw chimpanzees, we saw penguins, ...*]. Focus-sensitive particles like *only* involve quantification over the alternative set. The focus-level interpretation of (72d) is that all true members of the alternative set are *we saw iguanas*. There is considerable debate in the literature as to whether Structured Meanings or Alternative Semantics better captures the interpretation of focus. Much of it is to the effect that Structured Meanings is too powerful, i.e. overgenerates the set of semantic objects that are potentially sensitive to focus, or that

Alternative Semantics is not powerful enough to capture the full range of focus-sensitive phenomena (again, see Beaver and Clark 2008 for review). I have nothing to contribute to this debate, but choose to assume a model similar in spirit to Structured Meanings because it so closely parallels the division between focus and background in Lambrecht's model.

4.2.3 Contrast and Exhaustivity

Lambrecht's definition of focus captures the relation's core meaning, the aspect of interpretation that is common to all instances of focus. Other interpretive effects triggered by focus in certain environments, such as contrastiveness and exhaustivity, can be seen as augmentations of the basic meaning. In *contrastive focus* the value of the focus expression is presented as being unpredictable or surprising, "contrary to some predicted or stated alternative" (Halliday 1967: 206, cf. Zimmerman 2006). The prototypical context for contrastive focus is correction, as in (73). In cases of *exhaustive focus* the focus expression is presented as the only value that satisfies the salient open proposition. English *it*-clefts are interpreted exhaustively (Horn 1981b). For example (74) suggests that Sue is the only person that John kissed (in a relevant set of situations).

- (73) A: John kissed Mary.
 B: No, he kissed SUE.
(74) It is SUE that John kissed.

Contrastiveness and exhaustivity augment the core focus meaning in the following ways. Contrast involves an inference that there is a salient alternative to the focus expression to which the open proposition does not apply. Exhaustivity invokes universal quantification over the background-focus pair, indicating that only the focus expression can resolve the open proposition. Over the years, there has been a continuing debate in the

literature as to whether exhaustivity involves a semantic operator (e.g. Szabolcsi 1981, É. Kiss 1998) or whether it is a pragmatic inference (Horn 1981b). It is of course possible that the answer to this question is language- or construction-specific.

É. Kiss (1998) argues that contrast and exhaustivity are independent in terms of both semantics and formal expression. There are environments that permit contrast without exhaustivity, such as bare (prosodically-marked) focus in English (75, cf. 76).

- (75) A: John kissed Mary.
B: No, he kissed SUE. And LUCY, too.
(76) It is SUE that John kissed. # And LUCY, too.

The converse is also true. Vilkuna (1994) argues that in Finnish, movement to a particular position in the left periphery is associated with contrast but not necessarily with exhaustivity ((77); cited in É. Kiss 1998: 271).

- (77) A: Where do Anna, Kati, and Mikko live?
B: Anna asuu täällä.
'Anna lives here (and others may as well).'

The Finnish example illustrates that specific focus-marking structures may demand contrast and/or exhaustivity. As we have seen, in English *it*-clefts are interpreted exhaustively while bare focus is not. According to É. Kiss (1998), there are two focus positions in Hungarian, one that is associated with exhaustivity (78a) and one that is not (78b).

- (78) a. Tegnap este Marinak mutattam be Pétert.
last night Mary-DAT introduced.I PERF Peter-ACC
'It was to Mary that I introduced Peter.'
b. Tegnap este be mutattam Pétert MARINAK.
last night PERF introduced.I Peter-ACC Mary-DAT
'I introduced Peter to MARY.'

Augmentations to the basic focus meaning can also be marked by specific morphemes, such as the exhaustive particle *yé* in the West Chadic language Bole ((79); Zimmerman 2006: 149) or by prosodic patterns, for example the fall-rise “B” accent in English, which Büring (2003: 511) claims marks a contrastive topic (80).³⁰

- (79) Léngi kàpp-ák (yé) mòró.
 plant.PERF.F.AGR millet
 - *yé*: ‘Lengi planted MILLET.’
 + *yé*: ‘It is MILLET that Lengi planted.’
- (80) Discourse-organizing question: *Who ate what?*
 A: Well, what about FRED? What did HE eat?
 B: FRED (B-accent) ate the BEANS.

The discussion so far has collapsed the notions of *contrastive focus* and *contrastive topic*. The precise content of these two notions continues to be a source of debate; questions currently on the table include whether the two constructs are fundamentally distinct, and conversely whether there is a unified notion of “contrastiveness” that underlies both (see e.g. Lambrecht 1994: 291-295, Lambrecht and Michaelis 1998, Vallduví and Vilkuna 1998, Lee 1999, 2004, 2006 and Hedberg 2003 for discussion). Many languages mark the two relations differently. In Japanese and Korean, contrastive topics and contrastive foci receive different morphological markings (see Heycock 2007 for Japanese, Lee 1999, 2004, 2006 for Korean), while in conversational French the two constructs have different morphosyntax, as illustrated by the examples below (Lambrecht 1994: 292-293).

- (81) a. Mary, TOI tu fais les NETTOYAGE, MOI je fais la CUISINE.
 ‘Mary, YOU do the CLEANING, I’ll do the COOKING.’

³⁰ Lambrecht and Michaelis (1998) demonstrate that this accent pattern can serve the function of *topic ratification*, i.e. the establishment of a referent as a predictable argument of predication. They show that it is not necessarily contrastive. Consider a question that linguists are often asked: *What LANGUAGES do you speak?* Lambrecht and Michaelis (1998: 532) argue that accent on *languages* ratifies it as a topic; it certainly does not contrast languages with other things the linguist may speak. Still, the contrast account seems plausible for B’s response in (80), where *Fred* is already a ratified topic.

- b. Non, c'est MOI qui fais la cuisine, TOI tu peux faire autre CHOSE.
'No, I'M doing the cooking, YOU can do something ELSE.'

Example (81a) contains contrastive topics. The sentence expresses two parallel relations that hold between the topic referents and the comments, indicating the tasks to which they are assigned. This is marked by paired accents on the left-dislocated topical pronouns (*toi* and *moi*) and the nouns denoting housework, *nettoyage* and *cuisine*. Example (81b) expresses contrastive focus, in which the intended link between the speaker and the cooking is exhaustive. This is formally marked by a cleft and a distinct prosodic pattern.

I assume that contrastive topics are simply a subtype of contrastive focus. Contrastive topic constructions resolve open propositions in which one variable is a topical argument and the other is (part of) the predicate. Example (80) illustrates the contrastive topic interpretation: what is in focus is the relation between a topical individual, *Fred*, and part of the predicate, *the beans*. It differs from the more general case of contrastive focus in that in the latter, the focus expression can be any single variable, as in (79), or set of variables, as in (82). That the bare plural foci in (82) are not topics is evident from the fact that they are interpreted existentially; as we have seen, topical bare plurals require a specific interpretation.³¹

- (82) A: What did John give to who?
B: He gave TOYS to NEEDY CHILDREN.

Drubig and Schaffer (2001), building on the work of Dik et al. (1981), propose a more elaborated typology of focus interpretation, in which contrastive focus is divided into corrective and non-corrective subtypes and corrective focus is categorized further

³¹ For Büring (2003), (82) counts as a contrastive topic construction if it is marked by a B-accent, even though neither of the focus expressions is an aboutness topic.

depending on the nature of the correction. Here I assume that pragmatic effects such as correction are further augmentations of the focus meaning. Like contrast and exhaustivity, these components may be associated with specific constructions.

4.3 Informational Relations and Grammatical Constructions

This section turns to the relationship between topic, focus and the formal structures through which they are expressed. The fact that I consider topic and focus to be pragmatic objects is crucial and diverges somewhat from Reinhart's (1981) and Lambrecht's (1994) assumption that topics are properties of sentences. It seems clear to me that Reinhart and Lambrecht made this claim in order to distinguish topics *qua* properties of pragmatically structured propositions from discourse topics, a distinction I endorse. However, I think that it is incorrect to claim that topics are exclusively elements of sentence grammar. Consider example (83).

- (83) A: What's up with John?
B: Well, the paper's still not in.

In order for the dialogue to be coherent, B's response must be taken as providing information about John. Hence John is the topic of (83B) even though there is no topic expression that refers to him. For this reason, I assume that topicality may emerge from the interpretation of utterances in context, as well as from topic-marking constructions.

The relationship between informational relations and formal marking is complex. In the most straightforward cases, formal structures mark pragmatic relations and the relationship between form and pragmatic function is one-to-one. An example of this is prolepsis, which at this point we can conclude to be a topic-marking construction. In other

cases, there is a correlative relationship between topic or focus and a particular form. A clear case of this is grammatical subjecthood in English. All else being equal, subjects tend to be interpreted as topics, but there are subjects that violate the pragmatic requirements for topicality.

Topic and focus are marked in a wide variety of ways across languages. They are marked by prosodic contours, special affixes or particles, syntactic operations such as left- or rightward displacement, as well as specific syntactic constructions. Languages vary with respect to the set of mechanisms that they have available for topic- and focus-marking, as well with respect to the frequency with which they draw on the resources they do have. For example, Lambrecht (1994) and Ladd (1996) show that English depends more heavily on prosody to mark focus than languages such as French, Italian, and Catalan, whereas the Romance languages make more extensive use of displacement and other syntactic constructions. As this dissertation is concerned with the influence of information structure on English subject selection, I will conclude the present discussion with a few comments on what is known about the relationship between topic, focus, and subjecthood.

The relationship between topicality and English subjects has been touched on repeatedly throughout this chapter. We have seen that subjects are typically but not obligatorily topical. Topicality stands in a more direct relationship, though, with a set of constructions containing a detached initial NP, namely *topicalization*, *left dislocation*, and the *unlinked topic* construction. In topicalization (84), an argument NP appears clause-initially with a gap in its canonical position (Prince 1981b, Gregory and Michaelis 2001; example from Chafe 1976: 49, discussed in Lambrecht 1994: 295). In left dislocation (85), the

clause-initial NP has a co-referring pronoun (Prince 1981b, Lambrecht 1994, Gregory and Michaelis 2001). In the unlinked topic construction (86), the denotatum of the detached NP is not an argument at all, and thus is not associated either with a gap or with a co-referring pronoun (Rodman 1974, example from Lambrecht 1994: 193).

(84) The PLAY John saw YESTERDAY.

(85) Lions, they have long manes.

(86) Other languages, you don't just have straight tones like that.

For each construction, it has been argued that the detached NP is topical. This bears on the relationship between subjecthood and topicality because detached initial NPs share a major formal property with subjects, namely pre-verbal position near the beginning of the clause. They thus pair a prototypical subject function with a formal component of subjecthood. This dissertation will provide further evidence for the close relationship between subjecthood and topicality through the phenomenon of direct licensing in Topical Exclamatives and Copy Raising, discussed in Chapters 3 and 4.

Focus has not generally been considered to stand in any privileged relationship with subjecthood in English. Lambrecht (1995) ranks focus below topic for linking to subject position and provides evidence that focal subjects are permitted only in a restricted set of environments. Furthermore, as focus expressions often denote new referents, they are generally predicted to surface near the end of the clause on the basis of Prince's (1981a: 247) "conspiracy of syntactic constructions" that displace old information leftward and new information rightward (cf. Horn 1986, Birner and Ward 1998). Nevertheless, there

exists a construction in which a fronted NP expresses focus (Prince 1981b). *Focus-movement* is illustrated in example (87), from Lambrecht (1994: 295).³²

(87) FIFTY SIX HUNDRED DOLLARS we raised yesterday.

If, as I assume, the pragmatics of fronting constructions bears some relationship to the pragmatics of subjects, then perhaps focus is not as incompatible with subjecthood as has been previously assumed. Indeed, in Chapter 5 I will show that one of the ways in which an instrument can satisfy the informational requirements for realization as subject is through serving as a focus.

5. A Brief Overview of Sign-Based Construction Grammar

This section provides a summary of the basic components of Sign-Based Construction Grammar, focusing on the aspects that are relevant for formalizing the interface model of argument realization. SBCG is a monostratal, constraint-based theory of grammar that synthesizes many components of HPSG (Pollard and Sag 1994, Sag et al. 2003) with the insights of earlier research in Construction Grammar (e.g. Fillmore et al. 1988, Goldberg 1995). For detailed introductions to SBCG, see Michaelis (2009) and Sag (2010). Perhaps the most fundamental principle of SBCG is the Sign Principle (Sag 2010: 23), which states that every *sign* (lexeme, word, or phrase) must be licensed either by a lexical entry or by a construction. This discussion begins with an overview of the properties of signs and constructions.

³² As Lambrecht (1994: 295) points out, focus-movement is characterized by stress on the displaced constituent alone, while topicalization constructions contain stress on the topicalized constituent as well as another element of the clause.

5.1 Signs

Signs are actual units of linguistic structure: lexemes, words, and phrases. The information contained within a sign is organized by levels of linguistic structure: PHON represents phonology and prosody, FORM represents morphology, argument structure (ARG-ST) gives the list of lexically selected arguments of a predicate, SYN and SEM represent syntax and semantics, respectively, and context (CNTXT) represents certain aspects of pragmatics and information structure. Signs are represented as feature structures, with each level of linguistic structure containing an appropriate set of features. For the present purposes, I will focus on the feature structures associated with the syntactic, semantic, and pragmatic components.

SYN contains two components that are of interest to us here: category (CAT), which encodes the syntactic category of a linguistic object, and valence (VAL), which represents its potential to combine with other signs. Along with simple syntactic category descriptions such as *noun*, *verb*, etc., CAT encodes several other aspects of syntax, including case for nouns, verbal form and auxiliary status for verbs, and the external argument (XARG) of an argument-taking word or phrase, which is crucial for accounting for certain types of non-local syntactic phenomena.

VAL differs from ARG-ST in that the latter encodes a predicate's lexically-specified capacity to take arguments, whereas the former represents a predicate's combinatorial potential in a particular syntactic context.³³ Only lexical items, i.e. lexemes and words, have ARG-ST features, while lexemes, words, and phrases all encode VAL. Informally, VAL

³³ Manning and Sag (1999) show that having distinct ARG-ST and VAL lists is an advantage in accounting for cross-linguistic patterns in the binding of pronouns and reflexives. They argue that some binding phenomena are controlled by ARG-ST, and others by VAL.

works by maintaining an updated list of the arguments that a predicate has yet to combine with. For example, the lexical entry of a transitive verb takes two arguments as part of both its ARG-ST and its VAL features. When the verb combines with its complement, the resulting phrase has a VAL feature with only the subject NP remaining, indicating that at that point, the verb has yet to combine with its subject. Constructions can induce mismatches between the ARG-ST and VAL lists associated with particular predicates. Two well-known examples are filler-gap constructions and the null instantiation of arguments, both of which remove elements from the VAL list of a predicate (Sag 2010, submitted, Michaelis 2010).

The SEM component of a sign has two elements. The first, INDEX, gives the referential index associated with an entity or event; I will set this feature aside here for the sake of simplicity. The second encodes the sign's semantic content. Sag (2010, submitted) and Michaelis (2009) do this through FRAMES, a set of semantic roles associated with the "scene" that the sign evokes. Sag points out, however, that SBCG is compatible with a wide range of approaches to semantic representation (2010: 16). I take a different approach here, using event structure representations, rather than frames, to encode the semantic content of verbs and other eventive predicates (cf. Jackendoff 1990, Croft 1991, 1994, 1998, Levin and Rappaport Hovav 1995, 2005, Rappaport Hovav and Levin 1998). This will be described further in Chapter 5.

CNTXT contains linguistically significant aspects of pragmatics, including the referential indices of indexicals and deictic expressions as well as the felicity conditions associated with particular linguistic objects, contained in a set of background propositions

(BCKGRND). Sag (2010), following work of Engdahl and Vallduví (1994, 1996) within HPSG, suggests that CNTXT can be expanded to include an informational component (INFO-STRUCTURE) containing relations such as *topic* and *focus*, a proposal that I will be adopting here.

The following examples illustrate the components of signs that we have discussed. Figure 1 represents the noun *dog*; Figure 2 gives a representation of the verbal lexeme *run*. Both examples are adapted from representations provided by Sag (2010: 18-19).

Figure 1. A Nominal Sign in SBCG

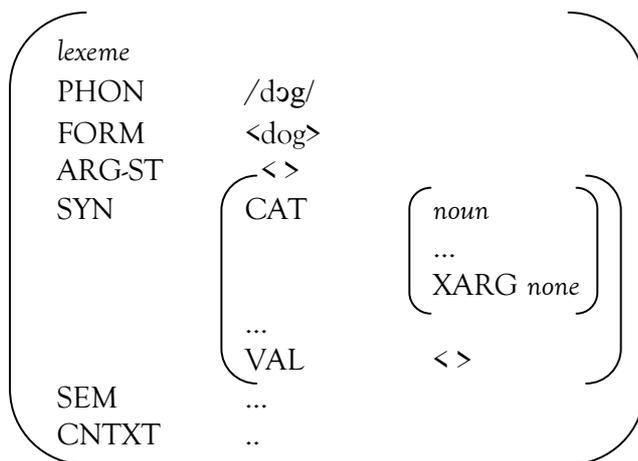
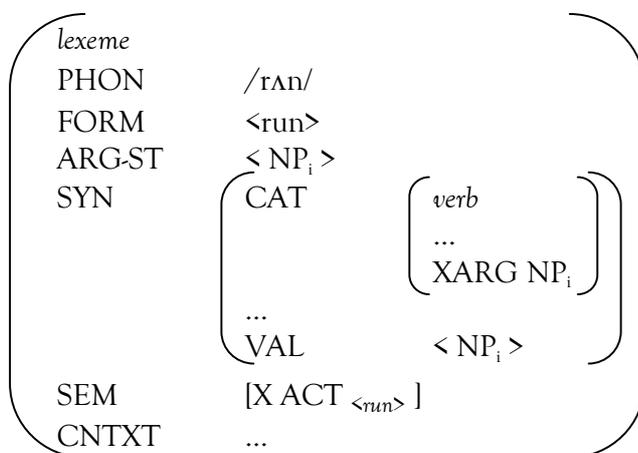


Figure 2. A Verbal Sign in SBCG



In SBCG, model objects are situated within a *type hierarchy*, which plays a central role in the organization of the grammar. Feature structures are classified into types according to the (in)applicability of a set of features. For example, the syntactic category of the lexeme *run* is of type *verb*, which makes it compatible with features such as valence; in contrast, nouns like *dog* are incompatible with valence. Types inherit aspects of their feature structures from more basic types. For example, causative verbs inherit the features associated with the basic type *verb*. The type hierarchy is based on *multiple inheritance*, meaning that types can inherit structure from more than one type. For instance, causative change-of-state verbs inherit features from both the causative verb type and the change-of-state verb type. The type hierarchy is located in the grammar's *signature*, which also lists the set of features contained in the grammar as well as a set of appropriateness conditions for assigning features to types. The type hierarchy operates throughout the grammar, capturing generalizations not only across signs but also across constructs, to which we will turn now.

5.2 Constructions

Sag (2010: 23) discusses the fact that the role of constructions in SBCG differs significantly from earlier formulations of Construction Grammar. He points out that constructions in SBCG are not “any conventionalized pairing between form and meaning” (Goldberg 1995: 4). Simple lexical signs such as the lexeme *run* contain conventionalized pairings between form and meaning but do not involve constructions. Instead, constructions are the means by which complex signs are constructed – words from lexemes, and phrases from words. They consist of a *mother* (MTR), which is a sign, and *daughters* (DTRS), a non-empty list of

signs. Constructs are licensed by *constructions*, grammatical schemata that represent well-formed pairings of mothers and daughters. In this way, constructions constrain the means through which morphosyntactic combination occurs in a language. Constructions underlie inflectional and derivational morphology as well as syntax. They also have varying levels of specificity, licensing processes as general as the combination of subject and predicate and as specific as subject-auxiliary inversion in negative adverb preposing, e.g. *Never will I harm you* (Fillmore 1999). Generalizations across constructions are captured by the type hierarchy. For example, the construction licensing subject-auxiliary inversion in negative adverb preposing inherits its features from a more general construction licensing inversion (Fillmore 1999, as discussed in Michaelis 2009 and Sag 2010). The following illustrates a simplified version of the subject-predicate construction, taken from Sag (2010: 41).³⁴

Figure 3. The Subject-Predicate Construction

$$\textit{subjpred-cxt} \Rightarrow \left[\begin{array}{l} \text{MTR} \left[\text{SYN} \left[\begin{array}{l} \text{MRKG } M \\ \text{VAL } \langle \rangle \end{array} \right] \right] \\ \text{DTRS} \left\langle X, \left[\text{SYN} \left[\begin{array}{l} \text{CAT} [\text{VF } \textit{fn}] \\ \text{MRKG } M:\textit{unmk} \\ \text{VAL } \langle X \rangle \end{array} \right] \right] \right\rangle \end{array} \right]$$

Constructions are implicational statements about types of constructs. Figure 3 says that if a construct consists of a subject and a predicate (i.e. is of the type *subjpred-cxt*), it must have the general feature structure that appears to the right of the arrow. Specifically, the mother must have an empty valence list (VAL < >) that is formed through the combination

³⁴ As Sag notes, a full representation of the construction would indicate how the semantics of the subject and predicate are integrated.

of the two daughters. The head daughter, the predicate, has a valence list (VAL <X>) that selects for the subject (X). The CAT feature of the head daughter indicates that it must be headed by a finite verb (VF *fin*).³⁵ Actual subject-predicate phrases will of course have fully fleshed-out feature structures, but they must conform to the above schema. Composition is achieved via *unification*, the combination of non-conflicting feature structures.

5.3 Subjects in SBCG

We are now in a position to elaborate on the nature of subjecthood in SBCG. As discussed earlier, grammatical relations are not formal objects in this framework. This is illustrated by the fact that the Subject-Predicate construction above does not explicitly refer to a subject. Rather, subjecthood can be seen as a composite of other formal features, along the lines discussed in Section 2 of this chapter. Subjects typically appear as the first member of a verbal sign's VAL list. When the VAL and ARG-ST lists are co-extensive (i.e. barring non-local argument realization, null instantiation, and other valence-altering constructions), the subject is also the first member of the ARG-ST list. In addition, the subject serves as the external argument (XARG), which is sensitive to the effects of non-local constraints (Sag 2010). In the standard case, the XARG coincides with the first member of the VAL and ARG-ST lists.

The subject selection phenomena addressed in this dissertation target different aspects of the formal features associated with subjecthood. Direct licensing alters valence but leaves argument structure untouched, while resolution influences the linking between ARG-ST

³⁵ The MRKG features are irrelevant for the present purposes.

and the semantic component. This section will conclude with a few comments on how each phenomenon is formalized in SBCG.

5.4 Direct Licensing

Direct licensing in Topical Exclamatives and Copy Raising can be accounted for within SBCG in a straightforward way. I propose a construction with a daughter sign that has undergone extraposition, as reflected by the fact that its first VAL member is expletive *it*. In the valence list of the mother sign, expletive *it* is replaced by a referential subject. The mother sign further specifies that the referential subject must be a topic. This is done through co-indexation of the first member of the valence list with the distinguished element of the topic relation, represented in the CNTXT component.

Though I do not pursue this here, it may ultimately be possible to capture the relationship between direct licensing in TE and CR and other topic-licensing phenomena in English through the type hierarchy. For example, it is possible that a more general construction underlies both the direct licensing phenomena investigated here and the licensing of “detached” sentence-initial topics, such as left dislocation and unlinked topic constructions. The constructional approach also provides a means of representing the ways in which languages differ with respect to the environments in which direct licensing can occur. In English, direct licensing of subjects occurs only in the environment of semantically “empty” subject positions formed by extraposition. This is represented through constraints on the valence of the daughter sign in the direct licensing construction. In Mandarin, direct licensing occurs across syntactic environments.

Therefore, the parallel construction in Mandarin presumably does not constrain the valence of the daughters.

As we have seen, the idea that constructions can underlie argument realization phenomena has a long history within Construction Grammar. Section 3.1.2 discussed *argument structure constructions*, form-meaning pairings that license syntactic constituents that are not selected by the verb (Goldberg 1995, 2002, 2006). The central claim of Goldberg's work on argument structure constructions is that crucial aspects of argument realization come from outside the representation of the verb, namely from the semantic representations of constructions. The interface model of licensing makes a related but distinct claim: that argument realization is sensitive to elements of interpretation that reside outside the semantic system.

5.5 Resolution

The SBCG account of resolution that I propose has two components: (1) an argument realization system in which semantics does not fully determine the linking between SEM and ARG-ST, and (2) a set of constructions that resolves underdetermination through conventional associations between ARG-ST and CNTXT. In the Instrument Subject construction, underdetermination results from the presence of multiple actors (the agent and the instrument) combined with a constraint requiring that *some* actor be the first ARG-ST member. Resolution takes place via a construction that links a non-initiating actor (the instrument) as the first ARG-ST member when it is associated with an element of an activated proposition (represented as part of the background, a component of CNTXT).

In addition to the empirical coverage that this type of account provides, it has the additional advantage of making explicit the nature of the connection between subjecthood and topicality. I claim that this connection consists (in part) of a body of constructions that pair elements of information structure with the formal correlates of subjecthood, e.g. initial position on the ARG-ST and VAL lists. The direct licensing and resolution constructions I propose are two components of the link.

6. Conclusion

This chapter has provided a context for the interface model of argument realization advocated here, addressing several key questions that have guided previous research on argument realization. It has also elaborated on the assumptions that I make about information structure and the nature of grammatical functions, subjecthood in particular. The following chapters focus on the effects of information structure on specific subject selection phenomena. I begin in Chapter 3 by presenting a direct licensing account of Topical Exclamatives.

Chapter 3

Direct Licensing I: Topical Exclamatives

1. Introduction

This chapter proposes an informational account of the licensing of the matrix subject in Topical Exclamatives (TE), illustrated in (1-4a). TE differs minimally from *it*-extraposition (1-4b) in that the matrix subject is referential, rather than an expletive. Both constructions involve extraposition, in which the most prominent semantic argument of the main predicate appears in post-verbal position. Supporting this claim is the existence of a “canonical” third alternant in which the post-verbal complement of the (a) and (b) sentences is realized as subject (1-4c).

- (1) a. [People are amazing when you give them the information and the tools they need.] They're amazing the responsibility they'll accept.¹
b. It's amazing the responsibility they'll accept.
c. The responsibility they'll accept is amazing.
- (2) a. [The vote for Republican Governor though was simply a vote against Spitzer. I think] he's horrible the way he goes after business.²
b. It's horrible the way he goes after business.
c. The way he goes after business is horrible.

¹ http://www.sbnonline.com/Local/Article/9794/77/0/Squeezing_the_tube.aspx?Category=92, accessed 11/9/2009

² <http://www.abovetopsecret.com/forum/thread232895/pg3>, accessed 11/9/2009

- (3) a. [Jade is a remarkable girl, she loves to play with Layla & isn't scared of other dogs. She likes to chase the cats when we're not looking & has always loved her food! She already knows how to come, sit & down, especially for cheese! We love our lopsided little girl], she is amazing what she has put up with so far.³
b. It is amazing what she has put up with so far.
c. What she has put up with so far is amazing.
- (4) a. [I watched him; he's having a hard time jumping off of one leg.] He's amazing that he could be as effective as he was without playing.⁴
b. It's amazing that he could be as effective as he was without playing.
c. That he could be as effective as he was without playing is amazing.

This chapter focuses on the alternation between TE and its counterpart with the expletive subject *it*, proposing an account of how the referential subject in TE is licensed. I argue that the subject is not assigned a semantic role by any predicate, or by any construction, and therefore is not licensed at the semantic level. Instead, the subject is licensed by the informational component directly in order to serve as a topic. In other words, TE is a topic-marking construction in which the topic expression is the subject. As a topic-marking construction, TE serves as an instruction from the speaker to the hearer to evaluate the sentence's propositional content from the perspective of its relevance to the topic expression.

Because TE has received little attention in the literature to date (an exception being Mack (to appear)), this chapter begins with a description of the construction's basic properties (Section 2). In Section 3, I present several strands of evidence that converge to support the claim that subjects in TE function as topics. They are sensitive to what I have claimed are the defining properties of topicality: pragmatic type restrictions and the requirement of relevance. TE subjects must be interpreted as referential and specific, and

³ <http://www.dogster.com/dogs/80763>, accessed 11/9/2009

⁴ <http://www.nytimes.com/2005/05/24/sports/24iht-NBA.html>, accessed 11/9/2009

furthermore they must be understood as being related in some way to the sentence's propositional content. They also possess other pragmatic/informational properties that are consistent with topicality. First, TE subjects are disproportionately coded as pronominal, indicating a high level of activation; given that the referents of topic expressions tend to be activated, this supports the topicality account. Second, there is a strong preference for TE subjects to co-refer with the topic expressions of adjacent topic-marking constructions, a constraint that presumably emerges from a preference for topic continuity. Finally, TE sentences are obligatorily exclamative, as the name of the construction indicates. I argue that the co-occurrence of topical subjects and exclamativity is not accidental. Exclamative sentences are especially informative, in a way that will be made explicit later, and therefore facilitate the satisfaction of the relevance requirement.

Section 4 presents the formal analysis of TE: a construction whose input (*daughter*, or DTR) has a valence list that starts with expletive *it* and whose output (*mother*, or MTR) contains a referential first valent that is marked as a topic expression. The same construction is responsible for a range of topic-licensing phenomena in English extraposition constructions, one example being Copy Raising, which is discussed in Chapter 4. Section 5 presents possible alternative analyses of TE that account for the licensing of the subject in semantic, rather than informational, terms. I demonstrate that these accounts cannot fully capture the constraints on the interpretation of TE subjects. In Section 6, I argue that the representation of TE specifies the nature of the relationship between the topic and the comment: that the comment must influence the speaker's beliefs about or attitudes toward the topic. I situate this claim within a typology of the ways in

which topic-comment pairs can satisfy the relevance requirement (cf. Shibatani 1994, Chen 1996, Jacobs 2001, Heycock and Doron 2003). Section 7 addresses the relationship between TE and a superficially similar construction that does not involve extraposition (e.g. *He's a genius the way he uses the ice*), demonstrating that such sentences demand an alternative analysis and do not pose a problem for the present account of licensing in TE. Section 8 concludes.

2. Description of TE

This section provides a preliminary description of TE, starting with a few observations on how the construction is used and turning to its core grammatical properties. As the subject of TE is the focus of the rest of this chapter, I concentrate here on the properties of the predicate and extraposed complement.

2.1 Usage

TE is frequently produced and widely accepted by speakers of American English. All of the native speakers that I have consulted find at least some tokens of TE to be felicitous and natural, given an appropriate discourse context. It is typically used in informal registers; nearly all of the naturally-occurring tokens that I have collected come from spontaneous speech and informal written genres such as blog posts and message boards. Despite the construction's widespread acceptance, it is prescriptively dispreferred. Some speakers who both produce TE sentences and judge them to be felicitous nevertheless perceive the construction as "lazy" or "sloppy." One possible source of this perception is that TE appears to be a "young" construction, a relatively recent innovation. As recently as 1992,

the year in which the Switchboard corpus of spoken American English (Godfrey et al. 1992) was released, it appears that TE was relatively rare: I found no instances of the construction in a large sample of environments in which the construction might be expected to occur.² However, a more recent Google search (performed on May 23, 2006) revealed that 20% of the 45 extraposition constructions containing the phrase *annoying the way* were TE. While differences in register and the nature of the sample may account for some of this effect, it nevertheless suggests that TE is increasing in frequency.

The possibility that TE is an emerging sentence type is also supported by the observation that the construction's acceptability is conditioned by the syntactic category of the extraposed complement. Many speakers accept sentences like (4a), in which the extraposed complement is sentential, and sentences of this sort are widespread in spontaneous speech; however, other speakers judge (4a) to be unacceptable. In contrast, AmE speakers consistently accept TE sentences with extraposed NP complements, such as (1a) and (2a). A reasonable hypothesis is that sentences like (4a) are recent innovations, and that the contrast between CP and NP complements in TE is an effect of change in progress, in which the availability of topic-driven licensing is extended across complement types.

The widespread acceptance of TE and its apparent increase in frequency are two pieces of evidence against the view that TE is a production error, a possibility that has been suggested to me on more than one occasion. In addition, TE is frequently found in written

² No tokens of TE were found within a randomly sampled set of 588 of the 1296 instances of the phrase "the way" found in the Switchboard corpus. I also found no instances of TE in 20 corpus searches of the form "ADJ the", where ADJ is an evaluative adjective known to be compatible with TE (e.g. *amazing*, *annoying*, *unbelievable*).

language, which is largely free of the temporal constraints that underlie many production errors in speech.

It has also been suggested to me that TE is a syntactic amalgam, involving the deliberate syntactic juxtaposition of two independent structures (cf. Lambrecht 1988, Brenier and Michaelis 2005).⁵ The source sentences for TE would presumably be extraposition (e.g. *It's annoying the way she always cracks her knuckles*) and a standard subject-predicate construction (e.g. *She's annoying*). Over the course of this chapter, it will become clear that the amalgam analysis does not predict the unique semantic and pragmatic properties of TE. As I will demonstrate, TE does not inherit the semantics of the subject-predicate construction: the subject is not an argument of the main predicate. In addition, TE is subject to pragmatic constraints that are not part of either of the putative source constructions.

2.2 TE, Right Dislocation, and Nominal Extraposition

Some of the essential properties of TE become evident through comparison with two formally and pragmatically distinct constructions: right dislocation (RD; Lambrecht 1994, Ward and Birner 1996, Grosz and Ziv 1998) and nominal extraposition (NE; Michaelis and Lambrecht 1996a). Examples (5-7) illustrate the contrasts between the three constructions, all of which contain a sentence-final NP (*the way she always cracks her knuckles*). (5) illustrates RD, which is characterized by co-reference between the matrix subject and the final NP. In contrast, the main subject in NE (6) is non-referential

⁵ The structures that I refer to as *syntactic amalgams* were known as *syntactic blends* in earlier literature (e.g. Bolinger 1961). However, in recent years the term *syntactic blends* is largely restricted to speech errors that come about through competition between alternative structures (Coppock 2005, 2009). Coppock (2005: 4) gives as an example of a blend the unacceptable utterance *They should lend a little hand*, which comes about through competition between *They should lend a little help* and *They should give a hand*.

(Michaelis and Lambrecht 1996a). TE (7) contains a referential subject that is not co-indexed with the final NP.

- (5) It's annoying, the way she always cracks her knuckles.
- (6) It's annoying the way she always cracks her knuckles.
- (7) She's annoying the way she always cracks her knuckles.

The claim that the initial NP in TE is the subject is supported by morphosyntactic evidence: it appears in pre-verbal position, triggers verbal agreement, receives nominative case, and undergoes inversion in questions, all properties associated with grammatical subjecthood in English. With respect to these diagnostics, TE subjects pattern with referential subjects in RD and expletive subjects in NE, and as such can be assumed to fulfill the same syntactic function. As we have seen, NE and TE involve extraposition, as shown by the existence of “canonical” variants in which the post-verbal complement is realized as the subject (see 1-4 above).⁵ In contrast, RD does not involve extraposition. Prosodically, TE patterns with NE; both are pronounced under a single intonation contour.⁶ RD, on the other hand, is characterized by a clear prosodic break between the matrix predicate and the post-verbal NP.

⁵ In contrast, Michaelis and Lambrecht (1996a) claim that not all NE sentences have canonical counterparts, on the basis of examples in which the extraposed NP receives a metonymic interpretation that is less salient when it is realized as subject. The dominant reading of (i) is that some *aspect* of the people is deemed amazing, while in (ii) the people themselves are judged to be amazing.

- (i) It's amazing the people you see here.
- (ii) The people you see here are amazing.

However, Michaelis and Lambrecht show that even subjects are sometimes interpreted metonymically. I will set aside the complexities associated with the metonymic reading here.

⁶ The use of a comma signals the boundary between two distinct intonation contours, while capital letters indicate prosodic prominence. I will use this notation whenever it is helpful to include prosodic information.

2.3 Constraints on the Predicate

The central constraint on predicates in TE is that they be compatible with an exclamative interpretation. Exclamative sentences express contravention of the expectations of a judge, typically the speaker (Michaelis and Lambrecht 1996a, 1996b, Michaelis 2001). In Zanuttini and Portner's (2003) analysis, this is formalized as the "widening" of a contextually given set of alternatives to accommodate an extreme value.⁶ As I will show in Section 3, TE conventionally expresses exclamativity, and for this reason its main predicate must be capable of expressing expectation contravention.

This accounts for the fact that the main predicate in TE must be *evaluative*; that is, it must express the subjective evaluation of a judge, which in TE is invariably the speaker. There are two types of evaluative predicates that can appear in TE: adjectives that express moral or aesthetic judgments (*good, bad, beautiful, terrible*) and predicates that entail a caused emotional state (*frightening, amazing, annoying*). Predicates of the latter class often have both verbal and adjectival forms; one example is *amazing/amaze*, which appears as an adjective in (1) and (3-4) above and as a verb in the examples below.

- (8) He amazed me the way he used an open tuning to play with just one finger, like at Woodstock.⁷
- (9) [Natasha might only be slight but] she amazed me the way she managed to own the stage, with perfect pitch and spectacular professionalism and huge sound which filled the hall.⁸

⁶ For example, they argue that the interpretation of the nominal exclamative utterance *The things he eats!* involves widening the contextual domain of what is normally eaten in order to accommodate the object of evaluation.

⁷ <http://www.michaelpowers.com/interview.shtml>, accessed 11/12/2009

⁸ <http://www.mudkiss.com/batforlashes.htm>, accessed 11/12/2009

Furthermore, TE predicates must have the capacity to convey that the speaker's judgment contravenes her expectations, or is otherwise extreme in some way. Michaelis (2001) demonstrates that the degree modifier *so* readily appears in exclamative sentences. For this reason, it can serve as a diagnostic to pick out the class of adjectives that can convey extreme judgments. For example, *amazing* and *terrible* are compatible with *so* (*so amazing, so terrible*), while *acceptable* is not, at least in the absence of strong contextual support (*? so acceptable*). The class of adjectives that co-occur with *so* are roughly the class that appears in TE, as the examples below suggest.

- (10) He's amazing the way he writes.
- (11) He's terrible the way he writes.
- (12) # He's acceptable the way he writes.

This supports the claim that the predicate in TE must be compatible with an exclamative interpretation.

2.4 Properties of the Extraposed Complement

The extraposed complement in TE is also subject to interpretive constraints. In particular, it must contain a proposition, as illustrated by the following examples. Example (13) is ungrammatical when produced with the single prosodic contour characteristic of TE.⁹ Although it is true that extraposed clauses tend to be grammatically “heavy” (e.g. Arnold et al. 2000, Wasow 2002), this is not a result of grammatical weight: (14), which contains a syntactically complex NP, is still ungrammatical because it lacks propositional content. It is also impossible to account for this effect through a constraint banning co-reference

⁹ Both (13) and (14) are acceptable when produced with the prosodic break following the main predicate that is characteristic of RD.

between the subject and extraposed complement. The subject of (15) is interpreted as being co-referential with the proposition expressed by the complement clause; nevertheless, it is acceptable.

- (13) * They're disgusting the dogs.
- (14) * They're disgusting the mangy dogs on the corner by the post office.
- (15) That's disgusting that a place like this can stay open.¹⁰

In the case of sentential complements, such as (16), the complement transparently denotes a proposition: *He's so together for all of that*. In NP complements, such as (17-18), the proposition appears within an embedded clause. In (17), it is expressed through a relative clause. There are two possibilities for the propositional content of (18): a relative clause reading along the lines of *He goes after business in a particular way* and a reading in which the nominal head *way* is semantically bleached and functions essentially as a complementizer. In the latter reading, the proposition expressed is simply *He goes after business*. This is comparable to the presuppositional or "committed" reading of the complementizer *how* (Cruse 1986, discussed in Michaelis 2001: 1046), which is also available in TE (19).

- (16) He's amazing that he's so together for all of that.¹¹
- (17) They're amazing the responsibility they'll accept.
- (18) He's terrible the way he goes after business.
- (19) He's terrible how he goes after business.

In fact, the majority of TE sentences with nominal complements are headed by *the way*, and many of these have the bleached reading. This is one way in which TE differs

¹⁰ 35mm.instantfundas.com/2008/11/honest-restaurant.html, accessed 11/12/2009

¹¹ <http://www.pbs.org/wgbh/pages/frontline/handofgod/etc/script.html>, accessed 11/12/2009

significantly from NE, which occurs with a wide range of NPs. The following are naturally occurring examples discussed by Michaelis and Lambrecht (1996a: 215-216).

- (20) It's astonishing the age at which they become skilled liars.
- (21) It's staggering the number of books that can pile up.
- (22) It's unbelievable the people that are verbally abusive to fat people.

Though TE occasionally appears with NPs other than *the way* (e.g. (1) above), this is rare. I believe that the frequency of committed readings of *the way* is due to constraints on the topic-comment link in TE. In Section 6, I demonstrate that one argument of the comment in TE must be propositional. NPs headed by *the way* are common because they provide a means of directly expressing a proposition through nominal syntax, parallel to the way in which sentential *that*-clauses transparently express propositions.

As we have seen, the extraposed complement can be either nominal or sentential, although the acceptability of the latter complement type varies across speakers. In addition, some speakers accept *wh*-complements, as illustrated below.

- (23) She's amazing what she has put up with so far.¹²

3. Evidence for the Direct Licensing Account

This section provides evidence that the subject in TE serves as a topic expression. The evidence for topicality in TE is drawn from diverse sources: pragmatic type restrictions (Section 3.1), the relevance requirement (Section 3.2), the information status of the subject (Section 3.3), the semantic role assigned to the subject referent within the embedded proposition (Section 3.4), the preference for the subject to co-refer with adjacent topic expressions (Section 3.5), the pragmatic function of TE as illuminated by its distribution in

¹² <http://www.dogster.com/dogs/80763>, accessed 11/9/2009

contrastive topic and focus constructions (Section 3.6), and finally, the requirement that TE sentences be exclamative (Section 3.7).

The diversity of diagnostics used here reflects the fact that the effects of topicality are evident in many aspects of discourse structure. However, I will continue to maintain that a distinction should be made between the diagnostics that capture the essence of topicality and those that tap into its many correlates. The diagnostics that directly probe topicality are pragmatic type restrictions and the relevance requirement; one example of a correlative diagnostic is information status. The correlative diagnostics can be seen as further support for claims that are built upon the primary diagnostics.

3.1 Pragmatic Type Restrictions

As TE subjects are topic expressions, they must refer to a specific individual or set (see discussion in Chapter 2, Section 4.1.1). For this reason, they force the specific reading of referentially ambiguous NP types. Two such cases are bare NPs (mass nouns and bare plurals) and indefinites. As we have seen, bare NPs have two possible interpretations: the *kind* reading, which refers to a specific natural kind, a type of individual, and the *existential* reading, which does not refer to a specific individual or set (Carlson 1977, Laca 1990, Diesing 1992, Kratzer 1995, Chierchia 1998, McNally 1998, Cohen and Erteschik-Shir 2002, Heycock and Doron 2003). Examples (24) and (25) demonstrate that TE requires the specific kind interpretation for mass nouns and bare plurals. The examples below are adapted from Lappin (1984), who observed that similar constraints apply to the subject in Copy Raising.

- (24) a. It's amazing the way snow falls.
b. Snow is amazing the way it falls.
c. It's amazing the way snow is all over the floor.
d. # Snow is amazing the way it's all over the floor.
- (25) a. It's amazing the way cows subsist on grass alone.
b. Cows are amazing the way they subsist on grass alone.
c. It's amazing the way cows have subsisted on the grass in our yard.
d. # Cows are amazing the way they've subsisted on the grass in our yard.

There are two variables at play within each of these paradigms: subject type (expletive or referential) and the semantics of the extraposed NP (whether it expresses a property that can potentially hold of a kind). Taking mass nouns (24) as an example, the (a) and (c) sentences have an expletive subject, while the TE sentences in (b) and (d) have a referential subject. In the (a) and (b) sentences, the extraposed NP expresses a property that can potentially hold of snow as a kind; in the (c) and (d) sentences, it expresses a property that is only compatible with a particular instance of snow. The TE sentence with a kind-type complement (24b) is felicitous, while the sentence with an existential-type complement (24d) is not. Example (24b) works because that the kind-type complement is compatible with the obligatorily specific reading of the subject, which comes about due to the requirement that the subject be topical. The topicality requirement also forces a specific reading of the subject in (24d), but it is incompatible with the existential-type complement with which it is combined. Example (25) shows that the same observation extends to bare plurals. The fact that subjects in TE must be specific is consistent with a body of literature which proposes that the interpretation of bare NPs is conditioned by the informational relations topic and focus, alongside other factors such as the distinction between stage-level and individual-level predicates (Laca 1990, Kamp and Reyle 1993, Cohen and Erteschik-

Shir 2002). Cohen and Erteschik-Shir (2002) argue that topical bare plurals are interpreted specifically; this is precisely the constraint that we see in the interpretation of TE subjects.

Indefinite NPs receive a wide range of interpretations, only some of which are compatible with topicality and thus with the topic position in TE. First, it is well-known that indefinites can receive either specific or existential readings; researchers have debated as to whether the distinction is pragmatic (e.g. Ludlow and Neale 1991) or semantic (e.g. Abbott 2003) in nature. To illustrate, (26) can mean either that John is looking for any individual who satisfies the description *dentist* or that he is looking for a specific individual who happens to be a dentist.

(26) John is looking for a dentist.

Because TE subjects are topics, they demand a specific interpretation of indefinite expressions. Consider the NE/TE sentence pair in (27) below. The NE sentence in (27a), adapted from a naturally-occurring sentence,¹³ is compatible with both specific and non-specific readings of *a program*; the speaker may be expressing annoyance either with a particular program or with some unspecified program. The availability of the existential reading is shown by the naturalness of the continuation sentence, which indicates that the speaker does not have a particular program in mind. In contrast, its constructed TE counterpart (27b) requires that the subject be interpreted specifically. For this reason, the continuation sentence is infelicitous.

¹³ www.red4est.com/dearbrock/, accessed 11/12/2009. The original sentence appears in (i):

(i) It's really annoying the way a program will lock up and the only way out of it seems to be to reboot the machine.

For the purposes of the discussion above, I have omitted the second sentence of the complement clause because it disambiguates the reading of the indefinite: the speaker's use of a pronoun to refer to the program indicates that s/he has a particular program in mind.

- (27) a. It's really annoying the way a program will lock up. (And I don't know which one it is.)
b. A program is really annoying the way it will lock up. (# And I don't know which one it is.)

Indefinites also permit a reading in which specific reference is made to a kind, similar to that observed for bare NPs (see Krifka et al. 1995 for discussion). As expected, TE is compatible with the kind reading of indefinites, as shown in (28).

- (28) A wolf is amazing the way it can crossbreed with a dog.¹⁴

One possible complication for the interpretation of the data discussed above is that the predicates that appear in TE are known to place similar interpretive restrictions on their arguments. Evaluative predicates such as *amazing*, *annoying*, and *terrible* are individual-level predicates, denoting properties that tend to hold of individuals over time (Carlson 1977). As was discussed in Chapter 2, individual-level predicates force strong/specific readings of ambiguous subject NPs. The sentences below illustrate that this holds for mass nouns and bare plurals. (29a) can only mean that *amazing* holds of snow as a kind, not a particular instance of snow; (29b) commits the speaker to the claim that cows as a kind are annoying.

- (29) a. Snow is amazing.
b. Cows are annoying.

For this reason, one possible interpretation of the constraints on TE subjects presented thus far is that they emerge from a predicative relation between the main predicate and the subject. However, my claim is that TE subjects are *not* arguments of the main predicate (or of any predicate at all). For this reason, it is important for the present account to demonstrate that the interpretive constraints on the subject follow from their pragmatic

¹⁴ This sentence is constructed on the basis of the naturally-occurring NE sentence in (i).

(i) It's amazing the way a wolf can crossbreed with a dog.
library.thinkquest.org/CR0212280/wolfdog.htm, accessed 11/9/2009

function, rather than from a lexical semantic constraint.

Fortunately, there is evidence revealing that the pragmatic type constraints associated with topicality are distinct from the semantic constraints imposed by individual-level predicates (cf. Endriss 2009), and that TE subjects are bound by the former. First, certain quantifiers such as *every* and *no* are incompatible with topicality, as suggested by the fact that they are unacceptable in the topic position of the *As for X* construction (30).¹⁵ However, they are perfectly acceptable as subjects of individual-level predicates (31).

- (30) a. # *As for every girl, she is/they are sick.*
b. # *As for no boy, he is sick.*
- (31) a. *Every girl is amazing.*
b. *No boy is annoying.*

As (32-33a) show, TE patterns with other topic-marking constructions in disallowing *every* and *no* in the topic position. In contrast, their NE counterparts with quantified embedded subjects (32-33b), adapted from naturally occurring examples, are perfectly acceptable.

- (32) a. # *Every agency was amazing the way it worked so smoothly.*
b. *It was amazing the way every agency worked so smoothly.*¹⁶
- (33) a. # *No one is annoying the way they can/can't stay in character from one writer to the next.*
b. *It's annoying the way no one can stay in character from one writer to the next.*¹⁷

This demonstrates that TE subjects have interpretive constraints that are consistent with topicality – and are above and beyond those associated with individual-level predicates.

¹⁵ Cf. Endriss (2009), who demonstrates that the same constraints are evident in topic-marking constructions in German and proposes a formal account.

¹⁶ <http://www.erie.gov/clarence/pdfs/minutes/TBM2009-02-25.pdf>, accessed 11/12/2009. The original sentence is *It was amazing the way every agency worked so well together and so smoothly.*

¹⁷ dcboards.warnerbros.com/web/message.jspa?messageID=2005245672, accessed 11/12/2009. The original sentence contains the adjective phrase *really annoying*.

Therefore, I conclude that the constraints are a consequence of topicality.

The requirement that topics be referential also accounts for the fact that idiom chunks are generally barred from the subject position in TE, as illustrated by the contrast between (34a) and (34b).

- (34) a. # The shit is pretty unbelievable the way it hit the fan.
b. It's pretty unbelievable the way the shit hit the fan.

Because pragmatic type restrictions are grounded in the defining characteristics of topicality, the interpretive constraints discussed above constitute one of the strongest sources of evidence that TE subjects are licensed to serve as topics.

3.2 The Relevance Requirement

This subsection turns to the other critical piece of evidence: the observation that TE exhibits pragmatic constraints characteristic of topic-comment constructions. In TE, the comment is the proposition formed by the application of the main predicate to its arguments and adjuncts. In the example below, the comment is formed by applying *amazing* to its sentential complement.

- (35) He's amazing that he's so together for all of that.

When a speaker chooses to mark an individual as a topic expression, she indicates that she is construing the comment as relevant to that individual. Lambrecht states the relevance requirement in the following way: "A statement about a topic can count as informative only if it conveys information which is relevant with respect to this topic" (Lambrecht 1994: 191). This, of course, raises the issue of what constitutes relevance. This is a complex issue, guided not only by general principles of cognition but also the speaker's beliefs and

intentions at the time of utterance. Nevertheless, previous researchers have observed that topic-comment structures tend to be grounded through a small inventory of semantic/pragmatic links that hold between the topic expression and part of the comment (e.g. Shibatani 1994, Chen 1996, Pan and Hu 2008). Section 6 will discuss these links at length and propose that TE subjects instantiate a particular type of link.

For the present purposes, it is sufficient to observe that topic expressions in TE are connected to their comments in two distinct ways. In all of the examples we have seen so far, the referent of the topic expression appears within the propositional argument of the comment. For example, the comment of (35) contains the propositional argument *he's so together for all that*, which contains the topic referent. However, there are sentences in which the topic referent does not appear within the comment at all. Two naturally-occurring examples appear below.

(36) The way the guitar works is so fucking annoying the way you have to switch scales every 10 seconds in order to get the note that you want.¹⁸

(37) I think blogs are amazing, the way we stumble into someone's world and see exactly what we need to see at just the right time.¹⁹

These sentences raise the question of how the relevance requirement is satisfied when the topic referent is not part of the semantics of the comment. The answer I will defend in Section 6 is that the proposition conveyed by the comment influences the speaker's beliefs about or impressions of the topic referent. Reinhart (1981: 68) makes a similar observation regarding constraints on the relationship between topic and comment in the *Speaking of X* construction. She suggests that (38a) is felicitous because the comment bears on the

¹⁸ ohgodwhatisthisidonteven.blogspot.com/.../music-simulation-games.html, accessed 11/9/2009

¹⁹ deroranoo.blogspot.com/.../i-just-wanna-be-honest-with-you.html, accessed 11/9/2009

speaker's knowledge of Marilyn Monroe in a way that the comment of (38b) does not. The examples in (39) show that TE has similar interpretive constraints.

- (38) a. Speaking of Marilyn Monroe, I read a book about her.
b. ? Speaking of Marilyn Monroe, I lost a book about her.
- (39) a. Marilyn Monroe's amazing the way people always read books about her.
b. # Marilyn Monroe's amazing the way people always lose books about her.

This suggests that the propositional content must be construed as relevant to the subject referent, and thus that TE is a topic-marking construction.

3.3 Information Status

Following Reinhart (1981), Lambrecht (1994), Lambrecht and Michaelis (1998), and Michaelis and Francis (2007), among others, I assume that topicality cannot be defined in terms of information status (see discussion in Chapter 2). Nevertheless, topicality is correlated with information status, as topics tend to be previously activated in the discourse. The effects of this correlation are evident in TE. In general, the more highly activated a discourse referent is, the better TE subject it makes. Though indefinite NPs, which typically introduce new referents, seem to be possible in TE, they are rare; I have yet to find a naturally occurring example. Constructed sentences with indefinite subjects such as (40a) are somewhat odd in isolation. However, they considerably improve when the indefinite is "anchored" via a relation to an activated referent (Prince 1981a); in (40b) the subject is anchored through its spatial relationship to the speaker. Activated referents make even better subjects; (40c) is fully felicitous if Annie is previously known to the speaker and hearer.

- (40) [Context: *At the conference last weekend ...*]
- a. A woman was so annoying the way she kept cracking her knuckles.
 - b. A woman right in front of me was so annoying the way she kept cracking her knuckles.
 - c. Annie was so annoying the way she kept cracking her knuckles.
 - d. This one woman was so annoying the way she kept cracking her knuckles.

Further evidence for the claim that TE subjects function as topics is the fact that non-activated referents are perfectly acceptable subjects if they are morphosyntactically marked as topics. (40d), in which the subject referent is not previously activated, is rescued by the fact that the form *this (one) X* serves to introduce new topics into a discourse (Prince 1981c).

These observations receive additional support from the results of a small corpus study that I recently performed. I obtained a sample of TE and NE sentence from a large website containing blogs, then marked whether the matrix subjects of the TE sentences and the embedded subjects of the NE sentences were lexical or pronominal.²⁰ It is well-known that pronominal NPs represent previously activated referents, whereas lexical NPs frequently introduce new referents into the discourse (e.g. Michaelis and Francis 2007). I found that TE subjects were more frequently pronominal (54%) than embedded NE subjects (34%). The full results appear in Table 1.

²⁰ This was done through a Google search performed on November 9, 2009, with the domain restricted to www.blogspot.com. The search form was *PRED the way*, where PRED was one of the four predicates that appears in Table 1. I included a random sample of acceptable NE and TE sentences. Sentences with the following properties were excluded: (1) sentences in which *it* was both the matrix and embedded subjects, because it is often unclear whether sentences with this form are TE or NE and (2) sentences with the subject *God*, which for many speakers resists pronominalization even when activated.

Table 1. Distribution of Lexical and Pronominal Subjects in TE and NE

Predicate	NE Matrix Subjects				TE Embedded Subjects			
	# Pron	# Lex	# Total	% Pron	# Pron	# Lex	# Total	% Pron
<i>amazing</i>	12	28	40	30%	12	10	22	55%
<i>annoying</i>	16	22	38	42%	5	4	8	56%
<i>disgusting</i>	6	18	24	25%	5	4	9	56%
<i>incredible</i>	14	25	39	36%	3	3	6	50%
TOTAL	48	93	141	34%	25	21	46	54%

These findings fit well with the claim that TE subjects serve as topic expressions.

3.4 Topicality and Semantic Role

Further support for the topicality analysis comes from the following observation. All else being equal, referents with semantic roles that are closely associated with topicality make better subjects. It has long been observed that agents more readily serve as topics than less semantically prominent participants (Givón 1984: 139, Van Oosten 1986, Lambrecht 1995). The effects of this constraint are evident in TE. In the sentences in (41), *John* is the agent of the event expressed by the extraposed complement NP; therefore, *John* is predicted to make a better matrix subject than *Bill*. This is reflected by the fact that (41b) is much easier to contextualize than (41c).

- (41) a. It is unusual the way John treats Bill so terribly.
 b. John is unusual the way he treats Bill so terribly.
 c. Bill is unusual the way John treats him so terribly.

This is consistent with the proposal that TE subjects conventionally express topicality.

3.5 Topic-Marking Constructions

In addition, TE is compatible with topic-selecting constructions such as *What's up with X?*, *As for X*, and *Speaking of X* (e.g. Reinhart 1981). In these structures, there is a strong

preference for the topic *X* to persist into the following discourse. When another topic-selecting construction immediately precedes TE, the topic expression typically co-refers with the TE subject, lending support to the claim that TE subjects are topics.²¹ Examples (42) and (43) illustrate this.

- (42) a. As for Mary, she amazes me the way Bill keeps beating her at Scrabble.
b. # As for Bill, Mary amazes me the way he keeps beating her at Scrabble.
- (43) A: What's up with Mary?
B1. Well, she amazes me the way Bill keeps beating her at Scrabble.
B2: # Well, Bill amazes me the way he keeps beating her at Scrabble.

In (42a), the topic expression of the *As for X* construction, *Mary*, co-refers with the TE subject, and the sentence is well-formed. However, when the two topic expressions fail to co-refer, as in (42b), the result is infelicitous. (43) extends this observation to a second topic-marking construction. This is the case despite the fact that Bill is the agent of the proposition expressed by the extraposed complement in (42-43), and thus should be a perfectly good topic by the standards of the requirement of relevance. In fact, if the TE sentence were to appear in isolation, Bill would generally be a better topic than Mary; cf. the discussion in Section 3.4 above.

In addition, TE sentences improve in acceptability when the subject co-refers with an established and highly persistent discourse topic. Example (44), which is adapted from a naturally-occurring discourse, has a prominent discourse topic: *What I think of Hussein* explicitly selects Hussein as the topic of conversation. The TE construction in (44a), which

²¹ One counterexample to the generalization that TE subjects must co-refer with nearby topic expressions is illustrated in (i):

(i) As for Bill, Mary amazes him the way she keeps beating him at Scrabble.
In this example, the topic of *As for X* co-refers with the experiencer argument of the main predicate, rather than with the subject. My take on this example is that the TE sentence has two topics: the subject and the experiencer object. Adjacent topic expressions may co-refer with either one..

appeared in the actual discourse, is fully felicitous in this environment. In fact, English speakers consistently find Continuation (1) more natural in the context provided than in isolation. This effect cannot be attributed to the presence of just *any* context, as the given context does not improve the felicity of the NE sentence (44b). Some speakers prefer the TE construction to its NE counterpart in this environment, even though the NE construction is generally judged to be more natural in isolation.

- (44) [Context: *What I think of Hussein: He's a cool guy. People go on about him not speaking English much, but he's so good at French!*]
- a. He's great the way he sort of corrects my French when we're talking without making a big deal of it.
 - b. It's great the way he sort of corrects my French when we're talking without making a big deal of it.

This is an additional reflection of the fact that TE subjects are licensed to function as topics.

3.6 Contrastive Topic Constructions

Further support for the topicality analysis comes from the fact that TE sentences can have *contrastive topic* interpretations but cannot express *contrastive focus* more generally. I assume that the contrastive topic interpretation is a subtype of contrastive focus, in which one focus expression in a multiple focus sentence also serves as a topic (see discussion in Chapter 2). Contrastive topic and focus are marked differently in many languages. In English, they can sometimes be distinguished via prosody: contrastive topics are characterized by “paired accents” on the topic and predicate, while in contrastive focus it is possible for stress to fall on a single focus expression (for discussion, see e.g. Jackendoff 1972, Lambrecht 1994: 291-295, Lambrecht and Michaelis 1998, Büring 2003, Hedberg

2003). The response in (45) is a contrastive topic sentence, while the response in (46) expresses contrastive focus on the subject NP.²²

- (45) A. What are Anna and Mary doing this afternoon?
B. ANNA has a MEETING; MARY'S finishing her PAPER.
(46) A: Are Anna and Mary going to the meeting?
B: ANNA'S going.

The following examples show that the TE can express the contrastive topic interpretation, but not contrastive focus on the subject NP. The felicitous TE responses in (47) have the prosody characteristic of contrastive topic sentences, in which stress falls on both the topic expression (here, the subject) and part of the predicate. Example (48) shows that contrastive focus on the subject, as reflected by the presence of a single accent, is unacceptable in TE.

- (47) A. How did Anna and Mary perform?
B1. ANNA was AMAZING the way she sang! (And MARY was UNBELIEVABLE!)
B2: ANNA was AMAZING the way she sang! (MARY was AWFUL!)
B3: ANNA was AMAZING the way she sang! (I missed MARY'S performance because I had to leave early.)²³
(48) A: How did Anna and Mary perform?
B: # ANNA was amazing the way she sang.

Again, this is consistent with the topicality analysis advocated here.

3.7 The Exclamativity Requirement

This section presents further evidence for the topicality analysis that comes from a perhaps unexpected source: the requirement that TE sentences function as exclamatives, expressing contravention of the speaker's expectations. Section 2.3 presented preliminary evidence

²² Lambrecht and Michaelis (1998) demonstrate that the prosodic pattern associated with sentences such as (45B) does not always mark contrast. See Chapter 2, Section 4.2 for further discussion.

²³ Thanks to Larry Horn for suggesting this example.

that TE is conventionally exclamative: the main predicate must have the capacity to express a judgment that is extreme in some respect. In this section I present further evidence that TE expresses expectation contravention, as well as another key property of exclamatives: factivity. Building on Michaelis's (2001) observation that there is an intimate relationship between topicality and exclamativity, I argue that the exclamativity requirement is a consequence of the informational function that TE subjects serve.

It has long been observed that exclamative sentences are *factive* or *presuppositional* (Elliot 1974, Grimshaw 1979, Michaelis and Lambrecht 1996a, 1996b, Zanuttini and Portner 2003, Portner and Zanuttini 2004). This constraint is evident in TE. In (49A) below, the proposition expressed by the extraposed complement (*He's so together for all of that*) is treated as presupposed. This is shown by the fact that this proposition cannot easily be targeted as a subject of disagreement (49B2).

- (49) A: He's amazing that he's so together for all of that.
B1: Nuh-uh! I don't find it amazing at all.
B2: (#) Nuh-uh! He's falling apart at the seams.

Zanuttini and Portner (2003) suggest that this property distinguishes exclamatives from other illocutionary forces, such as declaratives, interrogatives, and imperatives: presupposed content cannot felicitously be asserted, questioned, or commanded.

Expectation contravention, the other hallmark property of exclamative utterances, is also a necessary element of TE. This can be seen through constraints on the distribution of TE, as illustrated by the paradigm in (50) below, in which each member consists of a NE sentence and its TE counterpart.

- (50) a. It's / She's annoying the way she always cracks her knuckles.
b. # It's / # She's not annoying the way she always cracks her knuckles.

- c. # Is it / # Is she annoying the way she always cracks her knuckles?
- d. Isn't it / Isn't she annoying the way she always cracks her knuckles?

Compare the well-formed sentences in (50a) and (50d) with the pragmatically odd sentences in (50b-c). The latter sentences are unacceptable because they do not express expectation contravention. Negation of the evaluative predicate *amazing* in (50b) results in the denial of expectation contravention, while interrogative syntax in (50c) raises it as a possibility but does not express it (cf. Elliot 1974, Zanuttini and Portner 2003).²⁴ Zanuttini and Portner (2003) show that the combination of main-clause negation and interrogative syntax is compatible with exclamation. This is because negated yes-no questions presume the truth of the proposition in question, and thus can convey expectation contravention. This accounts for the fact that (50d) is felicitous. As the following examples show, the exclamation requirement extends beyond the distribution of TE with respect to negation and interrogativity. Even in affirmative, non-interrogative sentences, TE is compatible with potentially exclamative predicates (e.g. *awesome*) but not with predicates that fail to express expectation contravention (e.g. *acceptable*).

- (51) a. He's amazing the way he writes.
- b. # He's acceptable the way he writes.

The paradigm in (50) shows that TE and NE have the same distribution with respect to exclamation diagnostics. This raises the question of the source of exclamation in TE. One possibility is that TE inherits exclamative sentences from the constructions upon which it is built. However, the distribution of sentences with sentential extraposed

²⁴ Elliot's original observation was that exclamative complements cannot be embedded under *It isn't amazing* (examples from Zanuttini and Portner (2003: 8)):

- (i) It's amazing how very cute he is!
- (ii) * It isn't amazing how very cute he is!

complements demonstrates that this is not the case. (52) provides a paradigm comparable to (50), but with sentential rather than nominal extraposed complements.

- (52) a. It's / He's amazing that he was able to be as effective as he was.
b. It's not / # He's not amazing that he was able to be as effective as he was.
c. Is it / # Is he amazing that he was able to be as effective as he was?
d. Isn't it / Isn't he amazing that he was able to be as effective as he was?

Here the two sentence types diverge. *It*-extraposition structures are acceptable across conditions, indicating that they do not require expectation contravention. In contrast, TE sentences with sentential complements have the same pattern that we saw in (50): they disallow negation (52b) and interrogativity (52c), except in combination (52d). This indicates that the TE sentences are necessarily exclamative, while the sentential extraposition sentences are not.

This demonstrates that the exclamativity requirement is unique to TE, rather than an inheritance from the extraposition constructions on which it is based. I do not wish to claim, however, that exclamativity is part of the construction's formal representation. Rather, I will attempt to derive the exclamativity requirement from the fact that the construction expresses a topic-comment information structure.

Michaelis (2001) observes that there is a relationship between topicality and exclamative force that appears to hold across languages. She demonstrates that exclamatives often appear within topic-marking constructions, as in the following examples (p. 1044).

- (53) No ci posso credere che hai speso così tanto.
Not it can.1SG belief. INF that has.3SG spent that that.much
'I can't believe that she spent that much.'
- (54) Nereye kadar-yüzmüşsün ki gözlerime inanmıyorum.
Where extent swam.2SG EXCL eyes-my believe.NEG.PRES.1SG
'How far you swam! I don't believe my eyes!'

The exclamative sentence in (53) is an example of right dislocation in Italian, in which the topical pronoun *ci* ‘it’ corefers with a sentence-final proposition, *che hai speso così tanto* ‘that she spent that much.’ That this sentence is exclamative is evident from the presence of the degree adverb *così*, which appears in exclamative utterances. The Turkish example in (54) seems to be an unlinked topic construction, in which the sentence-initial topic *nereye kadar yüzmüşsün* ‘how far you swam’ is not incorporated into the propositional content of the main clause. This sentence’s status as an exclamative can be seen through the exclamative particle *ki*. Michaelis claims that in both cases, the detached (open) proposition serves as a topic. She suggests that this is related to the requirement that exclamatives be presuppositional; in many cases, the presupposed proposition is activated in the discourse.

We now turn to the question of whether this observation can be extended to TE subjects, which denote individuals or sets rather than propositions. Michaelis (2001: 1041) proposes a further constraint on the pragmatics of exclamatives that is potentially relevant: the arguments of the presupposed proposition must be identifiable to the speaker and hearer. She claims that this accounts for the degraded status of sentences like the following, where the indefinite NPs *a guy* and *someone* are presumed to be unidentifiable.

- (55) a. ? I can’t believe how much a guy spent!
 b. ?? Someone is so messy.

Given the close relationship between identifiability and specificity, this constraint bears an uncanny resemblance to the pragmatic type constraints on TE subjects that we have observed. Initially, this may appear to raise the possibility of deriving the type constraints from the exclamativity requirement, rather than the other way around. However, this will not work. As we have seen, there exist sentences such as the following in which the subject

referent is not an argument of the extraposed complement clause.

- (56) The way the guitar works is so fucking annoying the way you have to switch scales every 10 seconds in order to get the note that you want.²⁵
- (57) I think blogs are amazing, the way we stumble into someone's world and see exactly what we need to see at just the right time.²⁶

In these cases, the presupposed status of the extraposed complement does not directly bear on whether the subject referent is identifiable.

Instead, I propose that the topicality of TE subjects underlies the exclamation requirement. Simply put, exclamative sentences are highly informative, and accordingly serve to support the establishment of the topic-comment link. This boost in informativity is necessary because the semantic relationship between the topic and comment is either distant or absent. As discussed in Section 3.2, there are two ways in which the topic may relate to the comment in TE. In the typical case, the topic referent is an argument of the proposition expressed by the extraposed complement. In sentences like (55) and (56) above, the topic expression bears no explicit semantic relationship to the comment, and the topic-comment link must be supported entirely by the pragmatics. Even in the former case, the semantic basis for the topic-comment link is relatively thin. To illustrate this, consider the fact that in many topic-marking constructions, the topic referent is an argument of the main predicate in the comment (rather than an argument of an argument, as in TE). This can be seen in topicalization and left dislocation, represented below. In (58), the topic referent, the play, is an argument of the comment, that John saw the play

²⁵ ohgodwhatisthisidonteven.blogspot.com/.../music-simulation-games.html, accessed 11/9/2009

²⁶ deroranooblogspot.com/.../i-just-wanna-be-honest-with-you.html, accessed 11/9/2009

yesterday; in (59), the comment that lions have long manes includes the topic referent, lions.

- (58) The PLAY, John saw YESTERDAY.
(59) Lions, they have long manes.

Because semantic support for the topic-comment link is weak or absent in TE, it must be on strong footing within the pragmatics. This is provided in part by exclamative force. An exclamative utterance provides the hearer with quite a bit of information about the speaker's impression of the object of evaluation. First, it conveys the precise results of evaluation – typically, that it has an extreme value on a contextually given scale. Furthermore, it indicates that the evaluation has defied the speaker's expectations. The “informativity boost” provided by exclamative force improves the acceptability of other topic-marking constructions that accommodate relatively distant topic-comment links. This is illustrated by the prolepsis examples in (60), in which the topic expression *rodents* is linked to an argument of the comment, *guinea pigs*, via a set-subset relationship. This is a relatively weak link compared to the typical case, in which the topic referent is an argument of the comment. The predicate *awesome* in (60a), which is compatible with exclamative force, is at the root of the sentence's acceptability because it is particularly informative with respect to the subject's stance towards guinea pigs, and presumably rodents by extension. In contrast, (60b) is more difficult to contextualize because it is harder to see how the subject's relatively mild views on guinea pigs would bear on her overall impression of rodents.²⁷

²⁷ It is nevertheless possible to come up with a context in which (60b) is acceptable; for example, if the speaker is reporting on the restrictions that her landlord places on pets.

- (60) a. She said about rodents that guinea pigs are awesome.
b. She said about rodents that guinea pigs are fine.

To summarize, I have proposed that exclamation provides the high level of informativity that is required to establish the relationship between the topic and the comment in TE, due to its weak semantic basis. In Section 6, I will revisit this claim and argue that the nature of the topic-comment link underlies further constraints on the construction.

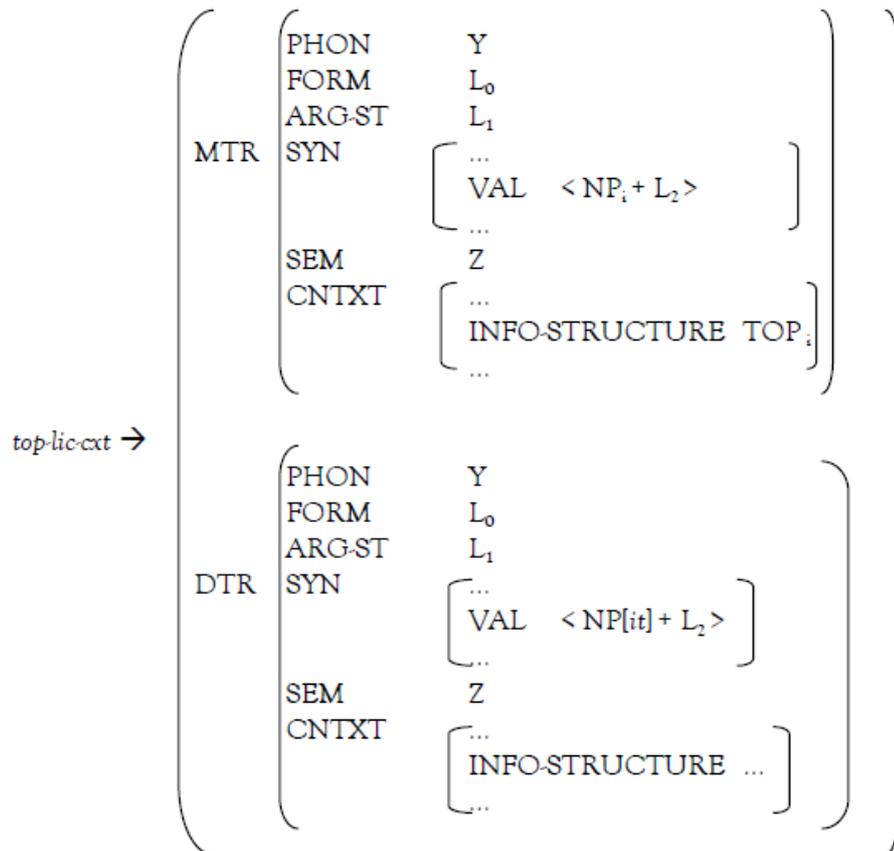
4. Direct Licensing

The previous section demonstrated that several strands of evidence converge to support the claim that TE subjects function as topics. I propose that they are licensed directly by the informational component to fulfill this function. This is done through a valence-altering construction that takes as input a predicate that licenses an *it*-extraposition structure and returns a predicate that licenses a topical subject. A preliminary representation of this construction, which I call Topic Licensing, appears in Figure 1.

Recall that constructions are implicational statements about the form that constructs (MTR-DTR pairs) may take. The Topic Licensing construction in Figure 1 says that if a construct is of the type *top-lic-ext*, it must obey the constraints represented to the right of the arrow. Note that the MTR and DTR must have identical feature specifications except at the levels of valence (VAL) and information structure (INFO-STRUCTURE). The valence list of the daughter sign begins with expletive *it*, which has a unique index in SBCG (Sag 2010: 41). The full valence list of the daughter sign is composed via concatenation (represented by the symbol +) of expletive *it* with the remainder of the valence list (L_2). The

first valent of the mother sign lacks the index associated with expletive *it*, indicating that it can be a non-expletive subject. The remainder of the mother sign's valence list is identical to that of the daughter.

Figure 1. The Topic Licensing Construction: Preliminary Version

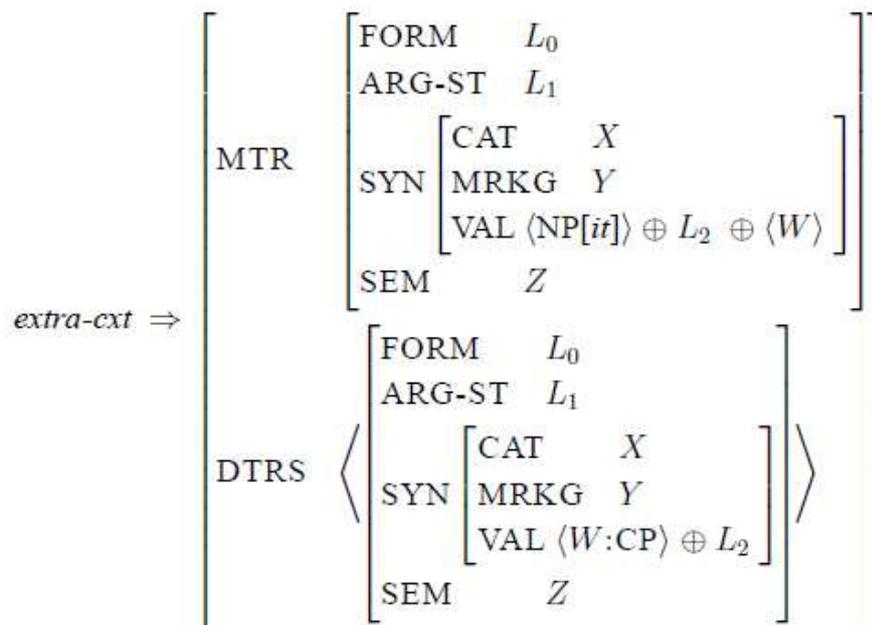


The mother and daughter signs also differ with respect to information structure. There are no constraints on the information structure associated with the daughter sign, whereas the referent of the mother sign's first valent is required to function as a topic, as indicated via co-indexation.

Topic Licensing is a *post-inflectional* construction, meaning that it composes words from other words. Specifically, the construction takes a predicate (verb or adjective) with a

valence list that starts with expletive *it* and returns a predicate that licenses a referential subject in its place. There is a compelling reason to favor a post-inflectional analysis of Topic Licensing over a derivational analysis, in which the input and output of the construction would be lexemes rather than words. Sag (2010, cf. Sag et al. 2003) argues that the construction that licenses *it*-extraposition structures with sentential complements is post-inflectional. The representation that he proposes appears in Figure 2 below (Sag 2010: 41).

Figure 2. *It*-Extraposition with Sentential Complements



It-extraposition licenses constructs in which the first valent of the daughter sign (which is required to be a CP) appears as the final valent of the mother sign; the first valent of the mother sign is expletive *it*. *It*-extraposition feeds Topic Licensing: the input to the latter construction contains a valence list that is created via the former construction. Therefore, it follows that Topic Licensing must also be post-inflectional, as opposed to derivational.

Note that according to Sag's analysis, the extraposition of sentential complements is a purely formal operation which alters valence without any effect on semantics or pragmatics.²⁸ In contrast, Michaelis and Lambrecht (1996a) demonstrate that the extraposition of nominal complements is associated with significant constraints on interpretation. On their account, couched within an earlier formulation of Construction Grammar, Nominal Extraposition is an all-focus construction, with specific constraints on the interpretation of the extraposed NP.²⁹ Though I will not attempt to translate their analysis into SBCG, it is clear that the resulting representation would constrain not only the valence of the mother sign, but also its semantics and pragmatics.

The Topic Licensing construction that licenses TE makes no distinction between nominal and sentential complements. Thus, it is fed both by the sentential extraposition construction in Figure 2 and a nominal extraposition construction comparable to that proposed by Michaelis and Lambrecht. This makes a rather strong prediction regarding the availability of direct licensing to express topicality: it should be available in all extraposition structures that have the capacity to express a topic-comment information structure. However, the predicates that appear in TE are only a small subset of the predicates that permit extraposition in English: others include raising verbs and adjectives (*seem*, *appear*, *likely*) and perceptual resemblance verbs (*look*, *sound*, *taste*). This raises the question of whether topic licensing occurs in these other environments. In Chapter 4, I will argue that raising and perceptual resemblance verbs do in fact permit topic licensing. A small number

²⁸ Miller (2001) argues that there are pragmatic constraints on *it*-extraposition, demonstrating that sentential complements are more frequently extraposed when they are discourse-new (cf. Horn 1986).

²⁹ According to Michaelis and Lambrecht (1996a), the extraposed NP must be accessible but not activated in the discourse, and must receive a metonymic interpretation.

of speakers even produce and accept topic licensing structures headed by raising adjectives.³⁰ This suggests that there are few, if any, constraints on the construction's input, beyond the basic requirement of a valence list that reflects extraposition.

Although I use the name *Topic Licensing* for the construction that licenses topical subjects in extraposition structures, there are certainly other constructions that license topic expressions in English. We have already seen several candidate constructions, for example topicalization, left dislocation, prolepsis, and the unlinked topic construction. Ultimately, it is important to determine how these constructions are formally related to each other. I consider this to be a key goal for future research.

5. Alternative Semantic Accounts

This section turns to possible alternative accounts in which the TE subject is licensed, at least in part, on semantic grounds. Due to the strong evidence that TE subjects function as topics, an alternative account of this sort would presumably augment an informational account rather than replace it. Still, I will argue that there is no need to include semantic constraints in the construction that licenses TE subjects. The construction has a purely informational function.

In this section, I present four possible semantic analyses. According to the first analysis, the subject referent is the target of an emotional state that is caused by the state of affairs expressed by the extraposed complement. On the second analysis, the semantics of TE is comparable to that of a *property factoring construction*, in which an individual and a property

³⁰ This is illustrated by examples like the following.

- (i) How is it wildly popular if not enough people bought it and are unlikely that they will?
www.muniwireless.com/2009/.../nokia-cancels-wimax-tablet, accessed 11/1/2009

that the individual possesses have distinct syntactic realizations. In the third potential account, TE involves *qualification*, or predication that applies to the subject under a particular description. The final account would propose that the extraposed complement is a *facet*, i.e. construal, of the subject denotatum. All of these accounts could presumably be implemented either through the lexical semantics of the main predicate or through argument structure constructions, so I will abstract away from questions about the nature of the formal licenser. However, we will see that these accounts converge on the claim that the subject in TE is an argument of the main predicate. At the end of this section, I will provide evidence that this is not the case, and that the “impression” of predication that emerges in the typical case is a result of the construction’s pragmatics.

5.1 Cause and Object of Emotion

The first semantic account is inspired by Pesetsky’s (1987, 1995) analysis of the argument structure of “object experiencer” verbs, which are so-called because they select for an grammatical object whose denotatum experiences an emotional state. A few examples of this verb class are *frighten*, *amaze*, *annoy*, and *anger* (61). Object experiencer verbs and their adjectival counterparts are among the predicates that can appear in TE. We have already seen several examples in which the main predicate is *amaze/amazing* or *annoy/annoying*; (62) illustrates that TE also appears with *frighten* and *anger*.

- (61) a. The movie frightened her.
b. Guinea pigs amaze John.
c. Her knuckle-cracking annoys people on a daily basis.
d. The article angered me.
(62) a. He frightened me the way he said it.³¹

³¹ thisibelieve.org/essay/59958/, accessed 11/14/2009

- b. [But I didn't really like Bella all that much], she angered me the way she acted sometimes.³²

The argument structure of object experiencer verbs and adjectives has been the topic of much debate in the literature (Pesetsky 1987, 1995, Belletti and Rizzi 1988, Grimshaw 1990, Bouchard 1995). All accounts agree that in addition to the *experiencer*, the individual that undergoes an emotional state, these verbs select for an entity or eventuality that is “responsible” for this emotional state. In some accounts (Belletti and Rizzi 1988, Grimshaw 1990), a single invariant semantic role is assigned to the responsible party. On these accounts, object experiencer verbs are bivalent, selecting for an experiencer and a *stimulus*. However, Pesetsky (1987, 1995) argues that object experiencer verbs are in fact trivalent. He breaks the “stimulus” role into two independent roles: the *cause of emotion*, which triggers an emotional state, and the *object of emotion*, which is the target of the emotional state. He supports this claim with examples similar to (63), where the intended meaning is that John is angry at some aspect of the article (e.g. its contents) but not at the article itself.

(63) The article angered John, but he wasn't angry at the article.

According to Pesetsky, although the verb *anger* and the adjective *angry* select the same set of arguments, they have distinct argument realization patterns: the subject of *anger* is a cause of emotion, while the object of *angry* is an object of emotion. Pesetsky claims that this allows sentences like (63) to escape contradiction, a result which would be unexpected if *the article* were assigned the same role in both clauses.

If one were to adopt Pesetsky's trivalent account of object experiencer verbs, one could

³² www.fanfiction.net/r/2800923/0/60/, accessed 11/14/2009

argue that the subject and the extraposed complement in TE are assigned two distinct semantic roles, one being the cause of emotion, and the other being the object of emotion. This analysis immediately faces two potential stumbling blocks. First, not all predicates that appear in TE are object experiencer predicates. Evaluative adjectives like *great*, *terrible*, and *awful* do not entail an experienced emotional state, a necessary precondition for the distinction between cause and object of emotion. However, one could potentially overcome this problem by proposing that TE is an argument structure construction that directly licenses the cause and object of emotion roles. The second stumbling block is more substantial. As Pesetsky (1987) acknowledges, it is impossible for both roles to be realized within a single clause, hence the ungrammaticality of (64).

(64) * The article angered John at the war.

Of course, this observation might lead us to doubt that cause and object of emotion are distinct roles after all. But even if they are, it is unclear how TE would escape the apparent constraint on simultaneous realization that we see in (64).

However, the crucial argument against an account of this sort is that it does not reflect the interpretation of TE: the subject denotatum is not required to be either the cause or the object of emotion. Starting with the former role, TE subjects do not always cause an emotional state. Consider again example (65). It is obvious that Marilyn Monroe herself is not the cause of the speaker's emotional state, but rather the state of affairs expressed by the extraposed complement. On the flip side, subjects need not serve as the object of the speaker's emotion. In (66), the speaker expresses annoyance at Mary's knuckle cracking but

explicitly denies that Mary herself was the target of annoyance. The object of emotion is clearly expressed by the extraposed NP.

- (65) Marilyn Monroe is amazing the way people always read books about her.
- (66) Mary annoyed me the way she kept cracking her knuckles, but I wasn't annoyed at HER.

Therefore, it appears as if the extraposed complement is assigned both roles, if in fact they are distinct. The subject, in contrast, bears neither role consistently. This suggests that we must look elsewhere for plausible semantic constraints on the licensing of the TE subject.

5.2 Property Factoring, Qualification, and Facets

We now turn to three closely related alternative accounts. The first capitalizes on the observation that TE bears a close resemblance to the class of property factoring alternations, in which an individual and one of his/her properties may receive joint (67a) or disjoint (67b) realization (Levin 1993: 72-77). On this analysis, TE (67c) would qualify as a case of disjoint realization. This could be implemented through a construction that selects for an *individual* and a *property*, as Knud Lambrecht (p.c.) suggests, or through comparable lexical semantic representations.³³

- (67) a. Mark's single-mindedness terrifies me.
- b. Mark terrifies me with his single-mindedness.
- c. Mark terrifies me the way he's so single-minded.

Alternatively, one could propose that TE expresses qualification, in which a property is predicated of an individual "under a description" (Anscombe 1957: 11, cf. Szabó 2003).

This means that the property is taken to hold of the individual in certain states (Szabó

³³ Although evaluative adjectives like *great* and *terrible* do not undergo the alternation in (67), in which the property is realized as a *with*-adjunct, they do allow similar alternations:

- (i) Mark's single-mindedness is terrible.
- (ii) Mark's terrible in that he's so single-minded.

2003) or in certain roles (Asher 2006). Example (68) illustrates four types of sentences that presumably involve qualification: *as NP* phrases (68a) and *qua NP* phrases (68b) (Landman 1989, Szabó 2003, Asher 2006), *at V-ing* phrases (68c), and *in that S* phrases (68d), where the main predicate is a subjective adjective (cf. Kamp 1975, Partee 1995).

- (68) a. Lucy is great as a typist.
b. Lucy is great qua typist.
c. Lucy is great at typing.
d. Lucy is great in that she types well.
e. Lucy is great the way she types.

In (68a-d), the individual under discussion is Lucy and the relevant description is Lucy in the role of typist. The predicate *great* holds of Lucy under this description although it does not necessarily hold of her generally (under *all* or *most* or even *many* descriptions). In a qualification account of TE (68e), the main predicate would hold of the subject under the description provided by the extraposed complement.

In the third approach, the denotation of the extraposed complement would be conceptualized as a *facet* of the subject denotatum. Croft and Cruse (2004: 116) state that facets are different ways of construing the same concept, “units that have a significant degree of autonomy, but can be unified to form a global GESTALT.” They argue that the concept *book* has two facets: [TEXT], i.e. informational manifestation, and [TOME], i.e. physical manifestation.³⁴ That the two facets have a certain degree of autonomy can be seen through the fact that it is possible for adjectives to target only one facet: in (69a), *interesting* targets the book’s TEXT facet while in (69b) *red* modifies its TOME facet (Croft and Cruse

³⁴ This proposal has a close parallel within Generative Lexicon Theory (Pustejovsky 1995, 2006, Asher and Pustejovsky 2005), in which it has been claimed that *book* has a complex type structure reflecting both its physical and informational aspects. In fact, Asher (2006) proposes to extend this analysis to qualification phenomena, which he argues involve complex type structures composed of an individual and its description. This underscores the similarities between the qualification and facet-based approaches.

2004: 116). However, the two facets can be unified, as in (69c), where both facets are targeted simultaneously.

- (69) a. an interesting book
b. a red book
c. You'll find that red book on the top shelf very funny.

Croft and Cruse suggest that conceptual representations of human beings arguably have facets as well: for example, *woman* appears to have two facets, [MIND] and [BODY]. However, the former is more essential, as shown by the contrast between (70a) and (70b) below (Croft and Cruse 2004: 126).

- (70) a. I'm not interested in the woman's body, I'm interested in the woman herself.
b. ? I'm not interested in the woman's mind or personality or feelings, I'm interested in the woman herself.

Because the extraposed complement in TE denotes a state of affairs, a facet-based account of the construction would require that individuals be construed as eventualities in which they participate or to which they are otherwise saliently related. I will set aside the question of whether or not this is plausible.

These accounts can be seen as distinct implementations of a single claim: that the main predicate in TE holds of the subject denotatum with respect to one of its components – a property, description, or facet. Most TE sentences are indeed interpreted in this way. For example, (71) is typically interpreted as a statement that the speaker finds Mary annoying with respect to her knuckle-cracking.

- (71) Mary's annoying the way she always cracks her knuckles.

However, a closer look at the interpretation of TE sentences reveals that there is no necessary semantic connection between the subject and main predicate, or between the

subject and the extraposed complement. Although it typically follows from (71) that the speaker finds Mary annoying, this interpretive preference can be overridden (72a). This indicates that the main predicate does not apply to the subject in TE. Compare this to the infelicitous (72b), which shows that the main predicate does hold of the extraposed complement.

- (72) a. Mary's annoying the way she always cracks her knuckles, but she herself isn't annoying.
b. # Mary's annoying the way she always cracks her knuckles, but the knuckle-cracking itself isn't annoying.
c. ? Mary's annoying the way she has no concern for others, but she herself isn't annoying.

As Laura Michaelis-Cummings points out (p.c.), there are cases in which it is difficult, if not impossible, to override the “impression” that the predicate holds of the subject. This is illustrated by (72c). Intuitively, the contrast between (72a) and (72c) is due to the fact that Mary's level of concern for others is likely to be an essential component of the speaker's impression of her, while her knuckle-cracking may or may not be. This suggests that the impression that the main predicate holds is pragmatic rather than semantic in nature. If the subject were an argument of the main predicate, then it would be impossible to deny that a semantic connection holds.

Furthermore, there are TE sentences in which the extraposed complement cannot plausibly be seen as a component of the subject denotatum – be it a property, a description, or a facet. Consider again example (73).

- (73) Marilyn Monroe's amazing the way people still read books about her.

The proposition that people still read books about Marilyn Monroe can hardly be considered an aspect of Monroe herself. It is neither a component of her physical

manifestation nor of her personality; it is not even an event in which she was directly involved. Rather, it is an observation that influences the speaker's impression of her. Crucially, this constitutes a pragmatic link, not a semantic one.

These observations raise the following question. If the main predicate does not apply to the subject, and if the extraposed complement is not necessarily a "description" of the subject denotatum, then why is an interpretation along these lines preferred? I believe that this is a consequence of the fact that TE conventionally expresses a topic-comment information structure. The extraposed complement typically denotes a state of affairs in which the subject referent participates because this establishes a semantic foundation on which to build a topic-comment link. The "impression" that the main predicate holds of the subject emerges from a requirement that the comment influence the speaker's perception of the topic referent. The following section explains these claims further.

6. TE and the Topic-Comment Relation

In Section 3.7, I argued that the semantic basis for the topic-comment link in TE is relatively weak, with the result that it requires an "informativity boost" that comes in the form of exclamation. This section develops this idea in more depth and applies it to further pragmatic constraints on TE. It starts with a discussion of the semantic and pragmatic foundations that underlie the topic-comment link across constructions and across languages. I then argue that TE instantiates a particular type of topic-comment link which I call *subjective topicality*, in which the comment influences the speaker's impression of the topic referent.

6.1 Foundations of the Topic-Comment Link

In some constructions, speakers are in principle free to link topics to comments in idiosyncratic or even downright strange ways. Suppose I utter the sentence in (74). By use of the *As for X* construction, I signal to the hearer that the proposition that Russian names are fabulous is intended to be a comment about guinea pigs. In the absence of supporting context, the unfortunate hearer is unlikely to be able to determine how the comment is relevant to the topic.³⁵

(74) *As for guinea pigs, Russian names are fabulous.*

In actuality, however, topicality is typically grounded through a semantic connection that helps to establish relevance between the topic and the comment. The inventory of semantic connections that serve this function appears to be rather small (cf. Shibatani 1994, Chen 1996, Jacobs 2001). Individual topic-marking construction may require that topicality be grounded in a particular way. It is only because the *As for X* construction is permissive in this respect – as we will see, it appears to place few, if any, constraints on the foundation of the topic-comment link – that (74) is simply difficult to contextualize, rather than ungrammatical.

In the most restrictive topic-marking constructions, the topic referent is an argument of the comment. In Section 3.7 I demonstrated that this is the case for left dislocation and topicalization in English. Example (75) below illustrates Right Dislocation (RD) in German, another topic-marking construction of this sort (Lambrecht 1994: 204). In Right Dislocation, the topic expression, which is the detached NP at the sentence's right boundary, has a co-referring pronoun within the comment.

³⁵ One possibility is that the speaker finds Russian names particularly becoming on guinea pigs.

- (75) Sie wohnen im dritten Stock, die Müllers.
 They live on third floor the Müllers
 ‘They live on the third floor, the Müllers.’

Of course, in these constructions the pragmatic constraints on the topic-comment link coincide with the formal mechanism through which topic-marking occurs: through displacement in topicalization, and through detachment in dislocation.

In other topic-marking constructions, the topic referent may stand in a particular semantic relationship to an argument within the comment. This is evident in the several distinct ways in which the topic-comment link can be grounded across languages in Multiple Subject Constructions (MSCs), in which the first subject is topical and the second is an argument within the comment (cf. Shibatani 1994, Chen 1996). In Lahu MSCs, the topical subject may be the inalienable possessor of the second subject ((76); Li and Thompson 1976: 468). In the Broad Subject Construction in Hebrew, which Heycock and Doron (2003) claim is a MSC, the topical subject may be the inalienable, or social, possessor of the second subject ((77); p. 5). Finally, the Japanese example in (78) illustrates that the topical subject may be a superset of the second subject (Kuno 1973: 64).

- (76) hɔ̄ ̄ na-qhɔ̄ yì ve yò.
 elephant TOP nose long PRT DECL
 ‘Elephants, noses are long.’
- (77) im be’emet dani ha-xavera ʃelo aba ʃela mi-tsarfət,
 if really Dani the-girlfriend his father her from-France
 ex ze ʃe hu af pa’am lo haya ʃam?
 how it that he never not was there
 ‘If indeed Dani’s girlfriend’s father is from France, how come he was never there?’
 (Lit ‘If indeed Dani his girlfriend her father is from France’)
- (78) kono class-wa dansei-ga yoku dekuru.
 this class-TOP male-NOM well are-able
 ‘This class, the boys do well.’

TE exhibits a distinct type of foundation for the topic-comment link. As we have seen,

in most TE sentences, such as those in (79) below, the topic referent is a participant in the state of affairs denoted by the extraposed complement, which is an argument of the comment.

- (79) a. He's terrible the way he goes after business.³⁶
b. She's amazing what she's put up with so far.³⁷
c. He's amazing that he was as effective as he was without playing.³⁸

The topic-comment link can also be grounded through the effect that the state of affairs expressed by the comment has on the topic referent. The effect may either be adversative or benefactive. Shibatani (1994) and Oshima (2006) illustrate that the topical subject of the indirect passive in Japanese is interpreted as being adversely affected by the event that the sentence expresses. This is illustrated by (80), which on one reading implies that Taro was adversely affected by Hanako's praising of the child (Oshima 2006: 148).³⁹

- (80) Taro-ga Hanako-ni kodomo-o home-rare-ta.
Taro-NOM Hanako-DAT child-ACC praise-passive-past
'The child was praised by Hanako, adversely affecting Taro.'
(Lit. 'Taro, the child was praised by Hanako.')

Finally, I propose that in some topic-marking constructions, the comment is taken to affect the speaker's impression of the topic referent, rather than the referent itself. The rather vague term *impression* is intended to include both the speaker's beliefs about the topic referent and her attitude or stance toward it. I call topics of this sort *subjective topics*, indicating that they reflect the speaker's subjective interpretation of the significance of the

³⁶ <http://www.abovetopsecret.com/forum/thread232895/pg3>, accessed 11/9/2009

³⁷ <http://www.dogster.com/dogs/80763>, accessed 11/9/2009

³⁸ <http://www.nytimes.com/2005/05/24/sports/24iht-NBA.html>, accessed 11/9/2009

³⁹ The adversative interpretation disappears if the child praised is taken to be Taro's; see Oshima (2006: 148), as well as the discussion below, for details.

comment. TE sentences such as the much-discussed (81) have subjective topics, because the comment is interpreted as influencing the speaker’s impression of Marilyn Monroe. In Section 6.2 I will argue that in fact all TE subjects are subjective topics.

(81) Marilyn Monroe is amazing the way people still read books about her.

Table 2 summarizes the types of foundations of the topic-comment link that we have seen so far, along with the constructions that have been used to illustrate them. Note that SOA stands for *state of affairs*.

Table 2: Semantic/Pragmatic Foundations of the Topic-Comment Link

Type	Topic	Comment	Representative Topic-Marking Constructions
1	X	PRED(X)	Left Dislocation, Right Dislocation, Topicalization
2	Inalienable Possessor of X	PRED(X)	Lahu MSC
3	Alienable Possessor of X	PRED(X)	Hebrew MSC
4	Superset of X	PRED(X)	Japanese MSC
5	Participant in X	PRED (X: SOA)	TE
6	Affectee of X	X: SOA	Japanese indirect passive
7	Subjective topic of X	X: SOA	TE

Individual topic-marking constructions may specify the grounds on which the topic-comment link must be established. Others are virtually unconstrained, allowing the

speaker to establish relevance in any way that she thinks the hearer will be able to reconstruct. Two constructions of the latter sort are *As for X* and prolepsis in English. The following examples show that both constructions can instantiate all of the foundation types that appear in Table 2. The interpretation of the first five types, illustrated by the sentences in (a-e), is straightforward. In the (f) sentences, which illustrate Type 6, the fact that the deadline was pushed up is taken to have an effect on Sue. This would most likely be an adverse effect of making the deadline more difficult to meet, but it could potentially be beneficial if Sue is ahead on her work and thus one of the few entrants that is likely to finish on time. One possible interpretation of the (g) sentences involves a link of Type 7. If Sue is the instructor of the class, then the speaker may take the proposition expressed by the comment as evidence of Sue's ineffectiveness as a teacher.

- | | | |
|------|---|---------------|
| (82) | a. As for Sue, she left early. | <i>Type 1</i> |
| | b. As for Sue, her feet hurt. | <i>Type 2</i> |
| | c. As for Sue, her sister lives in Toronto. | <i>Type 3</i> |
| | d. As for rodents, guinea pigs make great pets. | <i>Type 4</i> |
| | e. As for Sue, the way she sings is amazing. | <i>Type 5</i> |
| | f. As for Sue, the contest deadline was pushed up. | <i>Type 6</i> |
| | g. As for Sue, everyone in the class failed the exam. | <i>Type 7</i> |
| (83) | a. He said of Sue that she left early. | <i>Type 1</i> |
| | b. He said of Sue that her feet hurt. | <i>Type 2</i> |
| | c. He said of Sue that her sister lives in Toronto. | <i>Type 3</i> |
| | d. He said of rodents that guinea pigs make great pets. | <i>Type 4</i> |
| | e. He said of Sue that the way she sings is amazing. | <i>Type 5</i> |
| | f. He said of Sue that the contest deadline was pushed up. | <i>Type 6</i> |
| | g. He said of Sue that everyone in the class failed the exam. | <i>Type 7</i> |

In contrast, some topic-marking constructions allow only a proper subset of the link types illustrated in Table 2. One example is the Japanese indirect passive, which Shibatani (1994) and Oshima (2006) demonstrate can instantiate links of Types 2, 3, and 6. The latter type was illustrated by (80) above. In fact, the same sentence can have a Type 3

interpretation, in which Taro is taken to be the father of the child that Hanako praised. The following example illustrates a Type 2 link, in which the topical subject is the inalienable possessor of an event participant (Shibatani 1994: 463).

- (84) Taroo-wa Hanako-ni atama-o nagur-are-ta.
Taro-TOP Hanako-DAT head-ACC hit-PASS-PAST
'Taro was hit on the head by Hanako.'
(Lit: 'Taro, head was hit by Hanako.')

Shibatani (1994) advocates a unified account of the multiple interpretations of the indirect passive.⁴⁰ He argues that subjects must be construed as relevant to the scene depicted. For Shibatani, relevance has two dimensions: *physical proximity*, whether the individual in question is physically present at the scene, and *affectedness*, whether the individual is affected by the scene. In cases such as (84), in which a body part of the subject referent appears within the comment, the subject is both physically present and affected. According to Shibatani, the adversative reading emerges when other grounds for establishing relevance are weak, for example when the subject is not physically present. For example, the interpretation of (85) does not require that Taro be present when Hanako eats the meal; however, her doing so must affect him adversely (Shibatani 1994: 469).

- (85) Taroo-wa Hanako-ni gohan-o zenbu tabe-rare-ta.
Taro-TOP Hanako-DAT meal-ACC all eat-PASS-PAST
'Taro was adversely affected by Hanako's eating all the meal.'
(Lit: 'Taro, all the meal was eaten by Hanako.')

Shibatani's analysis raises the possibility that at least some of the links in Table 2 can ultimately be reduced to something more basic. In particular, affectedness seems to be at the core of several link types – most obviously Type 6, but also Types 2 and 3, as Shibatani

⁴⁰ In fact, his analysis extends to a wide range of constructions across languages, including possessor raising constructions and "free" or "ethical" datives (see discussion in Chapter 2).

proposes. Further support for this idea comes from the fact that a Type 3 link, in the absence of affectedness, may not be sufficient grounds for establishing relevance. This can be illustrated through constraints on the Japanese MSC (Doron and Heycock 1999).

- (86) a. John-ga zibun-zini-no hisyo-ga kubi-ni natta (koto)
 John-NOM self-GEN secretary-NOM was-fired (fact)
 ‘(The fact that) John, his secretary was fired’
 b. # John-ga musuko-ga waratta (koto)
 John-NOM son-NOM laughed (fact)
 ‘(The fact that) John, his son laughed’

In both sentences, the topical subject is the social possessor of the second subject, a Type 3 link. However, the topic relation is only licensed in (86a), presumably because it is easy to see how the firing of John’s secretary might affect him; for example, he might face controversy or a particularly high workload. In (86b), it is difficult to envision how John would be affected by the event of his son laughing.⁴¹

However, it is clear that a unified account of the links in Table 2 would require a broader conception of relevance than what is assumed by Shibatani. In particular, it is essential to capture cases in which what is “affected” is not the topic referent itself, but the speaker’s impression of it. As the following examples illustrate, this too can involve a range of links, including Types 4, 5, and 7. (87a) is interpreted as a report on the speaker’s impression of rodents; while (87b-c) reflect the speaker’s impression of Sue.

- (87) a. He said of rodents that guinea pigs make great pets. *Type 4*
 b. He said of Sue that the way she sings is amazing. *Type 5*
 c. He said of Sue that everyone in the class failed the exam. *Type 7*

⁴¹ As Larry Horn points out, it is possible to establish this connection with elaborate contextualization; for example, if John and his son are contestants in a game show in which the team that avoids laughing the longest wins.

In the following subsection, I will make the claim that topicality in TE must be established on subjective grounds; it may involve a link of Type 5 or Type 7.

6.2 TE Subjects as Subjective Topics

It is possible to account for a range of pragmatic constraints on TE if we adopt the proposal that subjects serve as subjective topics. In Section 3.7, I suggested that this is at the root of the requirement that TE sentences be exclamative. In an exclamative utterance, the speaker's judgment is taken to be extreme, contravening his or her expectations. When the speaker takes an extreme stance on a state of affairs that is relevant to a particular individual, as in (88a), it is likely to affect her impression of the individual. When the speaker's evaluation of a state of affairs is more moderate (88b), it is less likely to bear on a her impression of the individual. Hence the exclamativity requirement.

- (88) a. He's amazing the way he writes.
b. # He's acceptable the way he writes.

The subjective topic analysis also provides a means of accounting for the observation that the main predicate is usually taken to hold of the subject, even though this is not required. Given the felicity of examples like (89a), why do we generally interpret (89b) as a statement that Mary is annoying (at least in a particular respect)?

- (89) a. Mary's annoying the way she always cracks her knuckles, but she herself isn't annoying.
b. Mary's annoying the way she always cracks her knuckles.

A plausible answer is that applying the main predicate to the subject referent is an easy way to satisfy the requirement that the comment bear on the speaker's impression of the topic. If the irritation caused by Mary's knuckle-cracking results in annoyance at Mary herself,

then clearly the speaker's impression of Mary has been affected. Because this is the simplest means of satisfying the subjective topic requirement, it will generally hold except if explicitly overridden.

Finally, the proposal accounts for the infelicity of a class of sentences that might otherwise be taken as evidence against the topic-licensing analysis. Consider the sentences below. (90a) is a naturally occurring *it*-extraposition sentence, while (90b) and (90c) are constructed TE counterparts. The actual discourse leading up to (90a) centered on the life of JonBenet Ramsey, establishing both JonBenet and her mother as discourse topics.

- (90) a. [I just got finished looking at some of her other photos and] it was reprehensible the outfits her mother dressed her in!⁴²
b. ... # and she (=JonBenet) was reprehensible the outfits her mother dressed her in!
c. ... and she (=JonBenet's mother) was reprehensible the outfits she dressed her in!

What is striking here is the infelicity of (90b), in which JonBenet is the subject referent. On an account with no restrictions on the topic-comment link, this would be unexplained, given that JonBenet is an established discourse topic and the reprehensible nature of the outfits she wore certainly seems to be relevant to her. That JonBenet is a possible topic of the proposition expressed in (90) is illustrated by the *As for X* constructions below.

- (91) a. As for JonBenet, it was reprehensible the outfits her mother dressed her in.
b. As for JonBenet's mother, it was reprehensible the outfits she dressed JonBenet in.

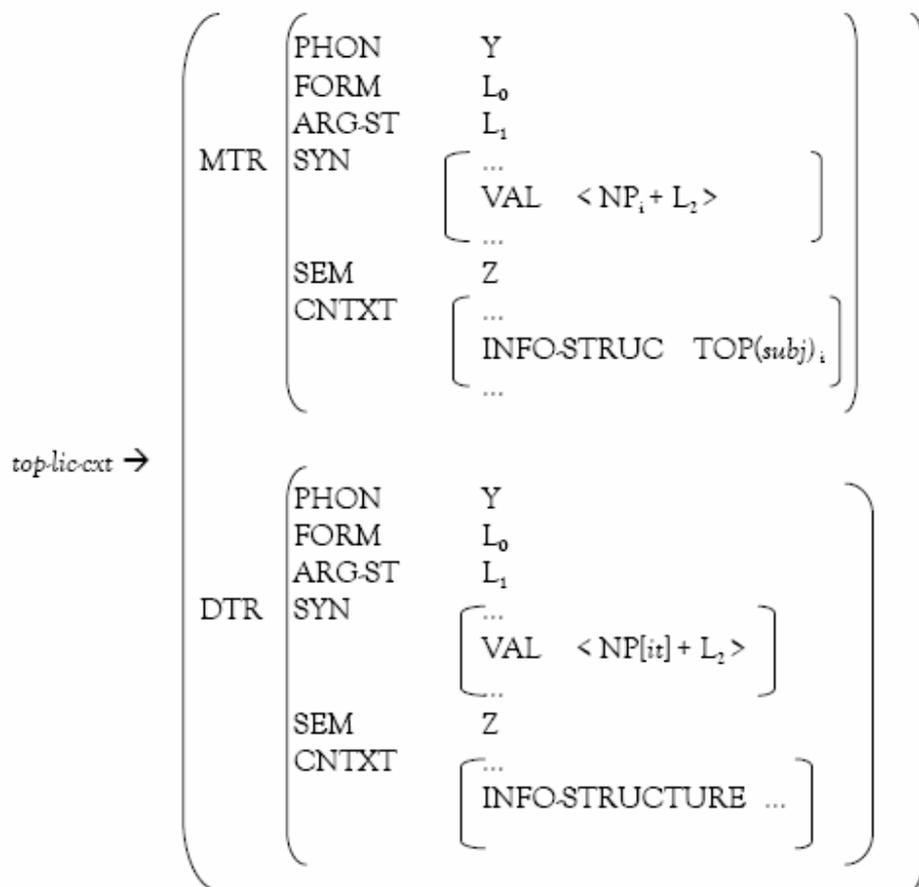
The fairly obvious problem with (90b) is that there is an "impression" that the speaker attributes reprehensibility to JonBenet, an intention we assume the speaker did not have in mind. In contrast, example (90c), in which reprehensibility is taken to hold of JonBenet's mother, is fine, in accordance with our intuitions that the speaker of (90a) likely believed

⁴² <http://www.crimeshots.com/forums/showthread.php?t=1269>, accessed 11/30/2009

just that. This contrast is a natural consequence of the claim that TE subjects are required to be subjective topics. Because JonBenet's mother is the one responsible for the outfits, the speaker's extreme evaluation of them is more likely to affect her impression of the mother than her stance towards JonBenet. Thus, while JonBenet is a suitable topic for the proposition in (90), she is not a suitable subjective topic, which I have argued is a requirement of TE.

Figure 3 presents a revised version of the Topic Licensing construction that underlies TE. In this version, the informational component of the mother sign indicates that the topic must be grounded in the speaker's subjective impressions (*subj*).

Figure 3. The Topic Licensing Construction: Final Version



7. NP predicates and TE

Shifting gears somewhat, this section addresses the nature of the relationship between TE and a superficially similar construction that is headed by a predicative NP. Examples of the latter sentence type appear in (92a) and (93a), which are headed by the predicate nominals *a genius* and *a bit of a nerd*, respectively. I will refer to members of this construction, rather uncreatively, as *NP sentences*.

- (92) a. He's a genius the way he uses the ice.⁴³
b. He's brilliant the way he uses the ice.
c. It's brilliant/*a genius the way he uses the ice.
- (93) a. [Mara was beginning to think that] Niki was a bit of a nerd the way she was trying so hard to take Mara under her wing.⁴⁴
b. Niki was a bit nerdy the way she was trying so hard to take Mara under her wing.
c. It was a bit nerdy/*a bit of a nerd the way she was trying so hard to take Mara under her wing.

Although the interpretation of the NP sentences is intuitively quite similar to that of their TE counterparts (92-93b) sentences, the two constructions have distinct syntactic behavior. Crucially, NP sentences do not permit the alternation with an expletive subject that is characteristic of TE (92-93c), which suggests that the former construction does not involve extraposition. Because extraposition is central to the analysis of TE presented here, the present account cannot be extended to NP sentences. This raises the question of whether to amend the informational account of TE in order to accommodate NP sentences or to maintain the present analysis and treat the NP sentences as an independent phenomenon. This section makes a case for the latter approach. It demonstrates that NP sentences differ from TE with respect to syntax, semantics, and pragmatics, and thus demand a distinct

⁴³ blog.mlive.com/snapshots/.../oilers_coach_mactavish_praises.html, accessed 11/18/2009

⁴⁴ www.writing.com/main/view_item/item_id/1388871/printit/1, 11/19/2009

analysis. The following discussion presents three further ways in which the two constructions diverge: the relationship between the subject and the main predicate (Section 7.1), the syntax and semantics of the post-verbal constituent (Section 7.2), and the relationship with exclamation (Section 7.3).

7.1 “Impression” of Predication vs. Actual Predication

As we have seen, in TE there is generally an “impression” that the main predicate holds of the subject. I have argued that this impression has a pragmatic source, namely the requirement that subjects function as subjective topics. Recall that the impression can be overridden in certain contexts, as in (94a). In contrast, it is impossible to deny that the nominal predicate holds of the subject in NP sentences, reflecting the fact that the two constituents stand in a genuine predicative relationship. This is illustrated by (94b).

- (94) a. He’s nerdy the way he asks questions right at the end of class, but he himself isn’t nerdy.
b. # He’s a nerd the way he always asks questions right at the end of class, but himself isn’t a nerd

This constitutes a second piece of evidence that NP sentences do not involve extraposition, in contrast with TE.

7.2 Syntax and Semantics of the Post-Verbal Constituent

A third piece of evidence is that the post-verbal constituent – in the examples discussed here, the NP headed by *the way* – is not an argument of the main predicate. The predicates that appear in NP sentences (e.g. *genius*, *nerd*) select for individuals, whereas *way* NPs, as we have seen, represent states of affairs. For this reason, it is impossible to realize the post-verbal constituent of an NP sentence (95a) as a subject (95b). This differs from TE (96a),

which as an extraposition construction is associated with a “canonical” alternant in which the extraposed complement appears as subject (96b).

- (95) a. He’s a nerd the way he asks questions right at the end of class.
b. * The way he asks questions right at the end of class is a nerd.
(96) a. He’s nerdy the way he asks questions right at the end of class.
b. The way he asks questions right at the end of class is nerdy.

For these reasons, I assume that NP sentences do not involve extraposition. Instead, they are simple subject-predicate sentences with a post-verbal modifier, the *way* NP. Their superficial resemblance to TE emerges from the fact that *way* NPs can either function as arguments or as adjuncts. Adjunct *way* NPs can be adjoined either sentence-finally (97a) or sentence-initially (97b) with roughly equivalent meanings, while *way* NPs that are extraposed arguments (98a) cannot be moved to a sentence-initial position (98b).

- (97) a. I don’t know if we can accept him the way he behaves.
b. The way he behaves I don’t know if we can accept him.
(98) a. She’s annoying the way she cracks her knuckles.
b. * The way she cracks her knuckles she’s annoying.

Adjunct *way* NPs are also semantically distinct from their argument counterparts, reflecting their function as modifiers. They are quite free with respect to how they are integrated into the propositional content of the sentence. Example (99) illustrates two of the semantic functions associated with adjunct *way* NPs.

- (99) a. I don’t know if we can accept him the way he behaves.
b. She always knows when I’m angry the way I walk in the door.

In (99a) the *way* NP expresses a justification for the speaker’s evaluation. In (99b) it names the epistemic source for the evaluation. Neither of these interpretations is possible within TE, because the *way* NP is uniformly interpreted as an argument of the main predicate.

7.3 Exclamativity

Finally, NP sentences are free from the pragmatic constraints that emerge from the fact that TE subjects function as subjective topics, such as the requirement of exclamative force. In contrast with TE, NP sentences do not always express exclamative judgments. This is illustrated by the felicity of sentences like (100a) in the face of the unacceptability of its TE counterpart (100b).

- (100) a. [Weis would be an idiot to leave and] I don't think he's an idiot the way he floats NFL rumors to get himself a raise.
b. # ... I don't think he's idiotic the way he floats NFL rumors to get himself a raise.

The optionality of exclamative semantics in NP sentences suggests that they do not conventionally express the complex topic-comment structure that is associated with TE.

To wrap up this section, we have seen that several aspects of the syntax, semantics and pragmatics of NP sentences differ considerably from TE. Because the differences are so pervasive, they do not challenge the present account of TE. Rather, they illustrate the need for an independent account of NP sentences.

8. Conclusion

This chapter has demonstrated that evidence for direct licensing can be found in English Topical Exclamatives, in which the matrix subject is licensed to serve as a topic. It has also shown that several of the construction's pragmatic quirks can be captured through the proposal that TE subjects function as *subjective topics*, which are grounded in the speaker's impression of the topic referent. This is formalized via a construction that takes *it*-extraposition as input and returns an otherwise equivalent structure with a subjective topic.

In Chapter 4, I argue that the same construction underlies the licensing of the matrix subject in Copy Raising.

Chapter 4

Direct Licensing II: Copy Raising

1. Introduction

This chapter extends the topic-licensing analysis to the construction that Andy Rogers originally christened as “Richard” but that now goes by the more theoretically-motivated name of “Copy Raising” (Postal 1974: 268, Rogers 1971, 1972, 1973, 1974, Horn 1981a, Lappin 1983, 1984, Heycock 1994, Potsdam and Runner 2001, Asudeh 2004, Asudeh and Toivonen 2005, 2006, 2009). On the present definition of Copy Raising (CR), the main predicate is an epistemic or perceptual verb that takes an obligatorily extraposed prepositional phrase headed by *like* or *as*, that in turn selects for a finite clause. Expletive *it*, the default subject in extraposition structures, is typically replaced by a referential subject.¹ The (a) sentences below illustrate CR; the (b) sentences its expletive-subject alternants.

- (1) a. He looks as though he’s been weaned on a pickle.²
b. It looks as though he’s been weaned on a pickle.
- (2) a. [His style is wooden, old-fashioned and artificial] ... he feels to me like he belongs in another era.³

¹ In a restricted set of environments, CR allows the main-clause subject to be non-referential, as in (i). I will argue that this reflects the fact that there are two routes to subject licensing in CR: the topic licensing construction and a purely syntactic co-instantiation or “raising” construction. See further discussion in Sections 4 and 5.

(i) There seems like there’s a problem.

² www.quotationspage.com/quote/11559.html, accessed 11/22/2009

³ www.amazon.co.uk/review/R3RD409NYAYBYY, accessed 11/22/2009

- b. It feels to me like he belongs in another era.
- (3) a. She sounds as though she's a happy and content cat.⁴
b. It sounds as though she's a happy and content cat.

These examples provide an illustration of the intuition behind the term “Copy Raising”; in the (a) sentences, it is as if the matrix subject has been raised from within the embedded clause, leaving a resumptive pronominal “copy” behind. Much of the cross-linguistic work on CR has focused on its syntactic properties, defining the construction in procedural terms as movement to the matrix subject position from within a finite embedded clause (see e.g. Joseph (1976) for Greek, Chung (1978) for Samoan, McCloskey and Sells (1988) for Irish, Moore (1998) for Turkish, Ura (1998) for Igbo, and Davies (2005) for a related construction in Madurese). In this chapter, I will propose a constructional analysis of English CR, adopting the abbreviation with no commitment to the idea that “copying” or “raising” is involved.

The central aim of this chapter is to account for the licensing of the matrix subject in English CR. Following several previous researchers, I assume that there are two distinct licensing mechanisms, one purely syntactic and one that comes with specific interpretive constraints (Rogers 1973, 1974, Horn 1981a, Potsdam and Runner 2001, Sag 2010, cf. Asudeh and Toivonen 2005, 2006, 2009).⁵ I focus here on the latter. Rogers (1973: 77) observes that CR sentences are often not quite synonymous with their expletive-*it* counterparts. For example, he claims that (4a), but not (4b), presupposes (4c).

⁴ www.dailymail.co.uk/coffeebreak/chat/r/t9901466/index.html, accessed 11/22/2009

⁵ Asudeh and Toivonen (2005, 2006, 2009) concur that there are two routes to licensing in CR (as I have defined it), but argue that both are associated with semantic constraints. See Section 2 for details.

- (4) a. Charley looked to me like he kissed Francine.
b. It looked to me like Charley kissed Francine.
c. I saw Charley.

Subsequent research has attempted to refine this observation. Asudeh and Toivonen (2005, 2006, 2009) argue that it is a consequence of the fact that the CR subject is assigned the semantic role of *perceptual source*.

Lappin (1984) and Heycock (1994: 293-294) discuss interpretive constraints of a different sort. They note that CR subjects require the specific reading of bare NPs, which as we have seen are ambiguous between an existential interpretation and an interpretation which makes specific reference to a kind (cf. Carlson 1977, Laca 1990, Diesing 1992, Kratzer 1995, Chierchia 1998, McNally 1998, Cohen and Erteschik-Shir 2002, Heycock and Doron 2003). Lappin (1984: 241, 245) offers the following contrast: (5a) is acceptable because the embedded clause expresses a property of cows as a kind, while (5b) is not because the embedded clause forces an existential reading of *cows*, which is unavailable in CR.

- (5) a. Cows seem as if they are lethargic to the casual observer.
b. * Cows seem as if they are grazing in Fred's field.

Both Lappin (1984) and Heycock (1994: 288-298) argue that the embedded clause functions as the syntactic predicate of the main-clause subject. Heycock (1994: 265) suggests that this syntactic operation comes with the pragmatics of topicality.

Thus, there are two general approaches to the interpretive constraints on CR: one (arguably) semantic, dealing with the judge's source of evidence for a particular proposition, and one informational, hinging on the notion of topicality. In this chapter, I argue that subjects in CR function as topic expressions. I provide new evidence for this

view, as well as evidence that the subject denotatum does *not* always denote the source of perception. Furthermore, I argue that topicality is at the core of the formal mechanism that licenses CR subjects. The Topic Licensing construction, which I introduced in Chapter 3 in order to account for the licensing of the main-clause subject in Topical Exclamatives (TE), is also responsible for licensing the subject in CR.

The remainder of this chapter proceeds as follows. Section 2 provides an overview of previous research on CR, starting with the traditional raising analysis (Postal 1974, Rogers 1971, 1972, 1973, 1974) and turning to a range of more recent approaches. Section 3 investigates the semantics of CR and its *it*-extraposition counterparts. Section 4 provides evidence that CR subjects are licensed to serve as topics. Section 5 returns to the observation that there are two routes to licensing in CR, and discusses the interesting question of how the two constructions might interact. Section 6 concludes.

2. Previous Accounts of CR

This section reviews previous research on CR, focusing on accounts that have attempted to capture the construction's interpretive constraints. Section 2.1 starts by surveying the evidence that Rogers (1973) presents both for and against a raising analysis of CR. This will also serve to illustrate several of the properties that have made CR particularly challenging to analyze. We then turn to two alternative approaches. Section 2.2 elaborates upon Lappin's (1984) and Heycock's (1994: 288-298) proposal that CR subjects are arguments of a sentential predicate: the embedded clause. In Section 2.3, I introduce Asudeh and Toivonen's analysis, in which the subject of CR is licensed via a syntactic dependency with its pronominal "copy," and is interpreted as a perceptual source (Asudeh 2004, Asudeh

and Toivonen 2005, 2006, 2009). Finally, Section 2.4 discusses the various incarnations of the idea that there are two ways in which CR subjects are licensed.

2.1 CR as Raising

One of the central questions that has guided research on CR is whether the matrix subject is licensed by a syntactic raising rule. This hypothesis was first explored in depth by Rogers (1971, 1972, 1973, 1974; cf. Postal 1974: 268), who points out parallels between CR and other constructions that have been argued to involve raising, most prominently subject-to-subject raising (SSR). CR and its expletive-subject counterpart are illustrated in (6) below; SSR and a parallel “unraised” sentence appear in (7).

- (6) a. He seems like he’s a good candidate for the job.
b. It seems like he’s a good candidate for the job.
- (7) a. He seems to be a good candidate for the job.
b. It seems that he is a good candidate for the job.

In the standard transformational analysis of SSR, the embedded-clause subject raises to its surface position as subject of the matrix clause. Rogers’s raising analysis of CR is the same, except that the raising operation leaves behind a pronominal copy. The two constructions also differ syntactically in one additional respect: the referential-expletive subject alternation associated with SSR co-occurs with a finiteness alternation in the embedded clause, while the CR alternation does not. There is a partial overlap in the sets of predicates that can serve as the main predicate in each construction. The epistemic verbs *seem* and *appear* are found in both CR and SSR, while other predicates are found in only one construction or the other, e.g. raising adjectives such as *likely* (SSR only) and perceptual resemblance verbs such as *feel*, *taste*, and *smell* (CR only).

In this section I review some of the traditional arguments for a raising account of SSR and discuss its extension to CR by Rogers. During the discussion it will become clear that a distinction must be made between arguments that point to a syntactic dependency between the subject and its pronominal copy and arguments that specifically support a raising analysis. Some of Rogers's original arguments have been interpreted by later researchers as evidence that the matrix subject is licensed via some sort of syntactic dependency, but not necessarily movement (Potsdam and Runner 2001, Asudeh 2004, Asudeh and Toivonen 2005, 2006, 2009, Sag 2010).

2.1.1 Evidence for a Raising Analysis

One essential component of the raising analysis is that the matrix subject position is non-thematic: for example, in (6-7a) *seem* does not assign a thematic role to *he*. One key argument in favor of this proposal is that the class of elements which can be “raised” contains non-referential constituents which cannot be assigned thematic roles, such as expletives and idiom chunks. This argument has been made in order to support raising accounts of SSR ((8); Rosenbaum 1967) and CR ((9); Rogers 1973: 81-83), as well as other constructions, including Subject-to-Object Raising ((10); Postal 1974), in which the embedded clause subject is argued to have raised to the matrix object position. The data below demonstrate that expletives (exemplified by dummy *there*) and non-referential idiom chunks (exemplified by the idiom *the shit hit the fan*) can appear in the “raised” position for all three constructions.

- (8) a. It seems that there is a problem.
- b. There seems to be a problem.
- c. It appears that the shit just hit the fan.

- d. The shit appears to have just hit the fan.
- (9) a. It seems like there is a problem.
- b. % There seems like there is a problem.
- c. It seems like the shit just hit the fan.
- d. % The shit seems like it just hit the fan.
- (10) a. She expects that there is a problem.
- b. She expects there to be a problem.
- c. She expected that the shit would hit the fan.
- d. She expected the shit to hit the fan.

The symbol % represents dialectal variation; as Horn (1981a) and Potsdam and Runner (2001) point out, some speakers are reluctant to accept non-referential subjects in CR. Crucially, though, the raised sentences in (9) are much better than superficially similar structures in which the subject is uncontroversially assigned a semantic role.

- (11) a. * There acted like there was a problem.
- b. * The shit acted like it hit the fan.

2.1.2 Problems for a Raising Analysis

On the classic diagnostics for raising that we have seen so far, CR patterns with SSR. However, Rogers (1973: 94-98, cf. Potsdam and Runner 2001) points out that the raising analysis is only well-supported for structures in which the raised constituent originates in the embedded subject position. Although most CR sentences have subject “copies,” structures with a co-referring pronoun in another syntactic position are in fact quite common (cf. Asudeh and Toivonen 2009).

- (12) a. She_i sounded like I offended her_i.⁶
- b. [Murat’s nose was bleeding and his face had reddened even more by this time, and] he_i looked like someone had painted his_i face with a bucket of red paint.⁷

⁶ launch.groups.yahoo.com/group/jameshallandwap/message/221, accessed 11/23/2009

⁷ <http://www.boxingnews24.com/2008/09/murat-defeats-campillo/>, accessed 11/23/2009

- c. [The sprouts overpower the entire sandwich and] they taste like someone just walked out back and picked them up and didn't clean them.⁸

The tests for raising that we have seen so far fail for structures of this sort. Rogers (1973: 96-97) shows that expletives and idiom chunks can only raise from the subordinate subject position. Potsdam and Runner (2001: 3-4) illustrate this with the following contrasts.

- (13) a. % There looks like there's gonna be a riot.
b. # There looks like John expects there to be an election.
(14) a. % The shoe looks like it's on the other foot.
b. # The other foot looks like the shoe is on it.

An even greater challenge to the raising account comes from sentences in which the matrix subject lacks a co-referring pronoun within the subordinate clause. Rogers (1973: 99-100) first made note of structures like those below, which are also discussed by Heycock (1994: 292), Potsdam and Runner (2001: 5), and Asudeh and Toivonen (2009: 4). The sentences in (15) were constructed by Rogers; those in (16) are naturally-occurring.

- (15) a. The soup tastes to me like Maude has been at the cooking sherry again.
b. The orchestra sounds to me like Mehta has been having a good night.
c. The peanut butter feels to me like David forgot to put the lid on the jar.
(16) a. [Andes Mocha Mint Indulgence]: They taste like someone in the flavoring department took a shortcut.⁹
b. She looks like there's some evil in there.¹⁰
c. [And they're not even good accents], they sound like someone in the control room said "make him Scottish."¹¹

⁸ <http://www.yelp.com/biz/eriks-delicafe-pleasanton>, accessed 11/23/2009

⁹ candyrecapper.com/archives.html, accessed 11/23/2009

¹⁰ www.dlisted.com/node/34319, accessed 11/23/2009

¹¹ forums.comicbookresources.com/archive/index.php/t-63488.htm, accessed fall 2008

The obvious problem for a raising analysis is that the matrix subject cannot have been licensed via raising if there is no position within the embedded clause in which it may have originated.¹² Thus it appears that a raising account is plausible when there is a pronominal copy in the embedded subject position, implausible when there is a pronominal copy elsewhere, and impossible when there is no copy at all. Rogers (1973: 94-103) discusses this problem at length, ultimately settling on a transformational account of the subject copy cases and an alternative approach to the licensing of the matrix subject in the other cases. Other researchers deal with the mixed properties of CR by dealing with only part of the construction as I have defined it. Potsdam and Runner (2001, cf. Moore 1998, Ura 1998, Davies 2005, Fujii 2005) argue that “true” CR exists only when there is a pronominal copy in subject position and thus restrict the investigation to those cases. Asudeh (2004) and Asudeh and Toivonen (2005, 2006, 2009) restrict the definition of CR to cases in which there is a pronominal copy in any position, thus eliminating structures headed by perceptual verbs, which commonly lack a copy pronoun. All of these researchers come down in favor of a syntactic analysis of CR on their respective definitions of the construction, but recognize that alternative modes of licensing are required outside these boundaries.

¹² As Heycock (1994: 292) points out, for the purposes of this argument it is irrelevant whether the subject referent is an implicit argument of the embedded clause. Crucially, it is not associated with a covert syntactic position, a necessary component of a raising analysis. This is shown by the fact that the subject NP cannot head a relativization of the subordinate clause, e.g. for (15a) **the soup that Maude has been at the cooking sherry again*.

2.1.3 The Significance of (Non)-Synonymy

The claim that raising structures have non-thematic subject positions suggests that they should be truth-conditionally equivalent to their expletive-subject counterparts. However, it is well-known that “raised” and “unraised” structures are often associated with interpretive differences – some subtle, some less so. This is illustrated by a famous contrast that Postal (1974: 357) attributes to Bill Cantrall.

- (17) a. It just now struck me that my wife has been dead two years tomorrow.
b. * My wife just struck me as having been dead two years tomorrow.

The crucial question is whether meaning distinctions of this sort are problematic for a raising account of CR. The answer depends on the assumption one makes about the relationship between movement rules and meaning. In the early 1970’s, when Postal and Rogers first proposed the raising analysis of CR, it was widely assumed that transformations are meaning-preserving. However, Partee (1971) demonstrated that several alternations that were hypothesized to emerge via movement involve semantic differences ranging from the subtle to the significant. One alternation that she discusses is Object-to-Subject Raising (OSR; also known as *tough*-movement, e.g. Bolinger 1961, Postal and Ross 1971, Akmajian 1972), in which the subordinate clause object is hypothesized to raise to the matrix subject position. The rough synonymy of sentence pairs like (18) was initially taken to be evidence in favor of a raising account. However, Partee points out that in some cases, such as (19), OSR seems to make a significant semantic contribution (p. 17, cf. Ross 1967).

- (18) a. It is impossible to outdo Lauren.
b. Lauren is impossible to outdo.

- (19) a. It is particularly easy to get this baby into these overalls.
b. This baby is particularly easy to get into these overalls.
c. These overalls are particularly easy to get this baby into.

A raising analysis of OSR, in tandem with the assumption that transformations are meaning-preserving, would make the prediction that all three alternants should be synonymous. However, there is a relatively strong contrast between (19b) and (19c): the former seems to say something about the baby while the latter seems to comment on the overalls. To maintain that the two structures are synonymous would at best require further justification, specifically an account of which aspects of meaning are relevant to the notion of synonymy and which aspects are not. Partee suggests that it may be possible to preserve the traditional transformational account (along with its assumptions) if it can be shown that alternants derived through movement differ only in non-truth-conditional aspects of meaning.

The debate over (non)-synonymy in putative cases of raising has framed much of the literature on CR, starting with Rogers (1971, 1972, 1973, 1974). Although Rogers advocates a raising account, he nevertheless argues that CR sentences and their unraised counterparts are non-synonymous, differing in two distinct elements of meaning. As discussed in Section 1, he shows that CR sentences imply that the speaker has had firsthand perceptual experience of the matrix subject referent. As Postal (1974) observes, SSR tends to generate a similar implication. However, it is more easily overridden in SSR than in CR, as the following examples demonstrate.¹³

- (20) Iris: Two of our committee members seem to be sick today.
Ezra: Actually, they're just late.

¹³ These sentences are adapted from an example of Charles Fillmore's that is discussed in Ruppenhofer (2004: 222-223).

- (21) Iris: Two of our committee members seem like they're sick today.
Ezra: # Actually, they're just late.

The dialogue in (20) is acceptable because the form of Iris's utterance does not imply that she has seen the two committee members in question. It could be uttered on the basis of her having noticed two empty chairs in the meeting room. For this reason, Ezra's response, which implies a lack of firsthand perceptual experience, is felicitous. This indicates that SSR does not require direct perceptual experience of the subject referent. In contrast, the perceptual implication is quite strong in CR, rendering the dialogue in (21) infelicitous. As I will argue in this chapter, it is not an absolute requirement, but it is strongly preferred in certain contexts.

In addition, Rogers (1973: 78-79) claims that raised sentences do not necessarily entail their unraised counterparts. He argues that (22a) does not entail (22b) on the following grounds: (22a) may be uttered felicitously if Charley's appearance or behavior gives off the impression that he kissed Francine, even if the speaker does not believe that any kissing has taken place. (22b), in contrast, requires the speaker to believe that an event of kissing occurred, and is therefore not entailed by (22a).

- (22) a. Charley seems like he kissed Francine.
b. It seems like Charley has kissed Francine.
(23) a. Charley seems to have kissed Francine.
b. It seems that Charley has kissed Francine.

This contrast could be seen as a more serious challenge to the classic raising account (coupled with the assumption that transformations preserve meaning), as it clearly intrudes into truth conditional semantics. Note that in SSR (23) this contrast is also present,

although perhaps less strong: (23b) suggests more forcefully that Charley has kissed Francine.

In Sections 3 and 4, I argue that Rogers's account of the contrast in (22) is not correct. I will show that neither CR nor *it*-extraposition requires that the speaker believe the proposition expressed by the embedded clause (in (22), that Charley kissed Francine). Instead, this is an inference that is sometimes drawn by the hearer: on the basis of the speaker's assertion that there is perceptual evidence consistent with a particular proposition, the hearer may infer that the speaker believes that proposition. The question is why this inference is more likely to be drawn in CR than in expletive-*it* sentences. I will argue that this is a consequence of the fact that CR subjects function as topics.

2.1.4 SSR and Information Structure

Others have suggested that raising is correlated with informational constraints. Two researchers who take this approach to SSR are Langacker (1995) and Achard (2000), both working within the framework of Cognitive Grammar (Langacker 1987, 1991). Langacker (1995) argues that raised constituents are pragmatic "reference points" that serve as a point of entry into the propositional content, a notion he suggests is roughly equivalent to that of topicality.¹⁴ Achard (2000) provides support for this analysis with corpus data from French, which reveal that raised constituents are more frequently discourse-old than unraised constituents. However, Ruppenhofer's (2004) larger corpus study of SSR in English shows no significant correlation between information status and the distribution of raised and

¹⁴ Langacker does not consider the phenomenon of raised expletives to be problematic for his account because he treats expletives as meaningful expressions of event settings that are capable of serving as "reference points."

unraised structures. In my view, it is an open question whether this discrepancy is due to methodological differences or genuine distinctions between the pragmatics of SSR in French and English.

On the topic of pursuing an informational account in light of Ruppenhofer's (2004) null findings for English SSR, the following comments are in order. First, Ruppenhofer's definition of topicality, which hinges on information status, is ultimately independent of the aboutness definition that I am adopting (see Chapter 2 for discussion). We will see in Section 4 that CR subjects are aboutness topics; I leave the investigation of their information status for future work. Second, we have seen that SSR and CR differ in several significant ways. On the syntactic level, SSR subjects are associated with an obligatory "gap" in the embedded subject position. CR subjects are never associated with a gap, but are sometimes associated with a co-referring pronoun, which can appear in any position within the embedded clause. At the semantic level, CR carries a stronger implication of direct perception than SSR. For this reason, the generalizations that have been made about the pragmatics of SSR do not necessarily carry over to CR. Moreover, given the existence of CR cases in which there is *no* pronominal copy (and hence no possible syntactic dependency between the subject and an element of the embedded clause) it is debatable whether SSR even constitutes a particularly appropriate comparison construction.

In fact, CR bears a stronger resemblance to TE. In both constructions, the main-clause subject may have a pronominal "copy" in the embedded subject position, in a lower syntactic position, or none at all. The syntactic parallels between the two constructions

reflect the fact that they are the output of the same licensing mechanism, the Topic Licensing construction.

2.2 CR Subjects as Arguments of Sentential Predicates

Both Lappin (1984) and Heycock (1994: 288-298) propose that the subject in CR is an argument of the embedded clause: in (24a), that *cows* is an argument of the sentential predicate (*as if*) *they are lethargic to the casual observer*, and in (24b), that *the soup* is an argument of (*as if*) *Maude has been at the cooking sherry again*.¹⁵

- (24) a. Cows seem as if they are lethargic to the casual observer.
b. The soup tastes to me as if Maude has been at the cooking sherry again.

The syntactic details of the proposals, which are shaped by the assumptions of Government and Binding Theory (Chomsky 1981), are not crucial for the present purposes. I focus here on their semantic/pragmatic content. Lappin (1984: 247-248) suggests that *seem* can mark its embedded-clause complements headed by *as if*, *as though*, or *like* either as full propositions or propositional functions, the latter being unsaturated and thus capable of combining with an argument. In CR sentences, the embedded clause is a propositional function that takes the main-clause subject as an argument. In expletive-subject sentences, the embedded clause is a full proposition. Heycock takes a different approach, assimilating the pragmatics of CR to that of Multiple Subject Constructions (MSCs) in Japanese, in which the first subject is integrated into the clause via topicality (p. 265).

¹⁵ Lappin includes the particles *as if/as though/like* within the sentential complement, while Heycock excludes them. I set this distinction aside here.

Heycock's analysis lends itself to a natural account of the pragmatic type restrictions on subjects in CR, which Lappin originally observed but had no explanation for (p. 244). Lappin (1984: 241), citing Lappin (1983), makes the following generalization.

“ ... the NP subject of a ‘seems as if’ construction can bind a pronoun in the complement clause only if (i) the complement can be understood as making an assertion about a determinate entity (or entities) that serves as the value of the bound pronoun or (ii) the VP complement can be understood as a generic predicate of the matrix subject NP.”

This mirrors the pragmatic type restrictions associated with topicality: clause (i) states that the CR subject must make reference to a specific individual or set, while the “generic” reading referenced in clause (ii) corresponds to the specific kind reading of bare NPs. Heycock (1994: 293) makes this connection, noting the parallel interpretive restrictions in Japanese MSCs. In the present analysis, I build upon Heycock's observation that CR subjects function as topics by proposing a licensing mechanism in which the role of topicality is made explicit.

2.3 CR Subjects as Perceptual Sources

Asudeh and Toivonen (2005, 2006, 2009, cf. Asudeh 2004) propose an alternative analysis of CR couched within Lexical-Functional Grammar (Bresnan 2001). They argue that the licensing of the main-clause subject depends on the presence of a co-referring pronoun, which can appear anywhere within the embedded clause. The main verb selects for a “manager resource” that removes the pronominal copy from composition, enabling the licensing of the matrix subject (see Asudeh 2004: 361-392 and Asudeh and Toivonen 2009 for details). In contrast with the present analysis, Asudeh and Toivonen assume that

sentences that lack pronominal copies, such as (15)-(16) above, are not “true” CR and require an independent analysis.

Furthermore, Asudeh and Toivonen claim that the matrix subject in CR is assigned the semantic role of *perceptual source*. They support this claim with two “puzzles” concerning the interpretation of CR in English and Swedish. The “puzzle of the absent chef” (Asudeh and Toivonen 2005: 7-8) is illustrated by (25): in the context described, why is it felicitous to use SSR but not CR?

- (25) [Context: *A and B walk into Tom’s kitchen. There’s no sign of Tom, but there are various things bubbling away on the stove and there are several ingredients on the counter, apparently waiting to be used.*]
a. Tom seems to be cooking.
b. # Tom seems like he’s cooking.

The Swedish “*på* puzzle” (Asudeh and Toivonen 2005: 10-11) is shown in (26): why are adjuncts headed by *på* incompatible with CR?

- (26) a. Det verkar på Tom som om han har vunnit.
It seems on Tom as if he has won
'Tom gives the impression that he has won.'
b. * Tom verkar på Lisa som om han har vunnit.
Tom seems on Lisa as if he has won
(intended: 'Lisa gives the impression that Tom seems like he won.')

Asudeh and Toivonen argue that both puzzles can be solved if we adopt an analysis in which CR subjects are interpreted as perceptual sources. (25b) is infelicitous because interpretive constraints on CR require that Tom be the perceptual source, which in the given context he clearly is not. CR is incompatible with Swedish *på*-phrases because they too denote perceptual sources, and there are presumably constraints disallowing two expressions of the same event role in a single sentence.

In the present chapter, I claim that CR subjects are not assigned the role of perceptual source, in spite of the evidence that Asudeh and Toivonen present. Instead, I will argue that they are often interpreted as perceptual sources because this is one means of satisfying the relevance requirement that comes with topicality.

2.4 Two Routes to Licensing

Asudeh and Toivonen's is the most recent of several analyses that have claimed that there are two routes to licensing in CR, as I have defined it.¹⁶ Rogers (1973: 94-103) was the first to propose a dual-route analysis of CR, a move that has since been supported by Horn (1981a), Potsdam and Runner (2001), and Asudeh (2004), Asudeh and Toivonen (2005, 2006, 2009), and Fujii (2005). All accounts except for that of Asudeh and Toivonen propose to treat structures with embedded-subject copies differently from those that have copies elsewhere (or not at all). We saw a compelling piece of evidence for this in Section 2.1: for many speakers, sentences with embedded-subject copies allow idiom chunks and expletive *there* to appear as the matrix subject (27), while sentences with copies elsewhere (28) do not.

- (27) a. % The shit seems like it just hit the fan.
 b. % There seems like there is a problem.
(28) a. * The fan seems like the shit just hit it.
 b. * There seems like John expects there to be a problem.

This leads Rogers (1973) and Potsdam and Runner (2001) to propose that only in sentences with embedded-subject copies is the subject syntactically linked to the co-

¹⁶ Several researchers, including Potsdam and Runner (2001), Asudeh and Toivonen (2005, 2006, 2009) and Sag (2010) have dealt with this issue by proposing a narrower definition of "true" CR, then claiming that it has only a single licensing mechanism. Regardless, all of these accounts acknowledge that two licensing mechanisms are required in order to account for the full range of structures that I address here.

referring pronoun. In Rogers's account this is done through raising, while in Potsdam and Runner's analysis both subjects are base-generated but form a syntactic chain. In the present analysis, I follow the same general line of reasoning. I propose that there are two constructions that underlie subject licensing in CR. The first is a counterpart to Rogers's raising rule and Potsdam and Runner's syntactic chain: it requires that the matrix subject be co-indexed with the embedded-clause subject, with no associated interpretive constraints (cf. Sag 2010: 47). The second is the Topic Licensing construction, which licenses a topical main-clause subject without placing any constraints on the composition of the embedded clause. This will be discussed further in Section 5.

Asudeh (2004: 366) and Asudeh and Toivonen (2005, 2006, 2009) propose a distinction along different lines. Recall that in their analysis, "true" CR requires the presence of a co-referring pronoun. They claim that epistemic verbs (*seem*, *appear*) require a co-referring pronoun within the embedded clause, while perceptual resemblance verbs (*look*, *sound*, *feel*, *smell*, *taste*) do not, and thus that only epistemic verbs permit CR (on their respective definition). Asudeh and Toivonen (2009: 4) provide the following judgments.

- (29) a. * Tina seems/appears like/as if/as though Chris has been baking sticky buns.
b. Tina smells/looks/sounds/feels/tastes like/as if/as though Chris has been baking sticky buns.

While for many speakers there is indeed a distinction between (29a) and (29b), it is clearly not categorical. The results of a grammatical judgment survey that Asudeh and Toivonen performed demonstrates that the distinction is more subtle than their judgments on (29) would suggest. They report (2009: 8) that 6.35% of English speakers accept copy-less sentences headed by epistemic verbs (29a), while 30% accept copy-less sentences with

perceptual resemblance verbs (29b). While the difference between the two verb types certainly requires an explanation, it is crucial to note that some speakers *do* accept the epistemic verb sentences. I have also found naturally-occurring examples of this sort, as illustrated below.

- (30) a. [I'm having those pains. They're terrible and] they seem as if someone's stabbing me.¹⁷
b. [The sentences lack eloquence and flow.] They seem as though someone was trying to write the story while bull riding.¹⁸
c. Some people seem like it's the end of the world.¹⁹

Because there is not a categorical distinction between the two verb types, I will not treat them differently in this analysis.

It is also significant that most speakers rejected all of the copy-less sentences, regardless of verb type. My speculation is that this is (at least in part) because the sentences were presented in isolation. If CR subjects function as topics, then it is expected that they will be less acceptable if there is no context that establishes the grounds for topicality. This, I will argue, is particularly the case in copy-less sentences, in which the connection between the topic and the comment must be inferred.

3. Semantics and Argument Realization in CR

This section provides a range of observations regarding the relationship between semantics and argument realization in CR and its expletive-subject counterparts, which I hereby refer to as *looks like* structures. The main predicates of *looks like* structures are lexical evidentials,

¹⁷ http://www.steadyhealth.com/severe_pains_what_s_wrong_with_me_t104115.html, accessed 11/29/2009

¹⁸ http://www.meghan-mccarthy.com/articles_marthastuartbook.html, accessed 11/29/2009

¹⁹ www.gamespot.com/pages/forums/show_msgs.php?topic_id=26400910, accessed 11/29/2009

as suggested by Anderson (1986), Mithun (1986), Gisborne (1998), Ifantidou (2001) and Gisborne and Holmes (2007). This means that their basic function is to indicate that there is evidence in support of a particular proposition. Faller (2006) argues that there are two types of evidential markers: those that contribute to the propositional content of the sentence and those that do not, instead adding a “comment” on the speaker’s source of evidence. Faller claims that the former class can be targeted by disagreement, while the latter class cannot. Apparently, the main predicates of *looks like* structures contribute to propositional content, as they can serve as the locus of disagreement (31B1). So can the proposition expressed by the embedded clause (31B2).

- (31) A: It looks like he’s a good candidate for the job.
B1: No, it doesn’t!
B2: No, he isn’t!

This indicates that the core meaning of a *looks like* structure is a proposition (ϕ_1) that there exists evidence compatible with a second proposition (ϕ_2). In (31A), ϕ_1 corresponds to the whole sentence (minus the expletive subject), while ϕ_2 corresponds to the embedded clause. The basic meaning of a *looks like* structure can be augmented in several ways, two of which are correlated with subject choice. First, the sentence can specify the *evidential source* for ϕ_2 ; as we will see, this is sometimes, but not always, expressed through the main-clause subject. Second, it can undergo *epistemic strengthening*, in which the hearer infers that the speaker believes ϕ_2 . This too is correlated with subject selection, in that there is a stronger preference for epistemic strengthening in sentences with expletive subjects than in CR. I will argue that both effects are consequences of the fact that CR subjects function as topics, while the parallel expletive-subject construction is not topic-marking.

This section begins with a discussion of the basic grammatical properties of the two verb classes that appear in *looks like* structures: epistemic and perceptual verbs (Section 3.1). We then turn to an in-depth discussion of the relationship between subject selection and the specification of the evidential source (Section 3.2). Finally, we explore the connection between epistemic strengthening and the expletive-referential subject alternation in *looks like* structures (Section 3.3).

3.1 Epistemic and Perceptual Resemblance Verbs

This section takes a closer look at the argument realization patterns of the main predicates in CR: the epistemic verbs *seem* and *appear* and the perceptual resemblance verbs *look*, *sound*, *feel*, *taste*, and *smell*. The following examples demonstrate that all members of both classes participate in the *looks like* alternation; the CR alternants of (32b-g) are naturally-occurring.

- (32) a. He/it seems like he'd be a good candidate for the job.
b. Sadly, she/it appears like she thinks she is hot stuff.²⁰
c. They/it sound(s) like they have a third subwoofer stuffed in there somewhere.²¹
d. She/it looks to us like she's part Greyhound or Whippet or something.²²
e. He/it feels to me like he belongs in another era.²³
f. The cookies/it tasted like they were made yesterday.²⁴
g. We/it smelled like we were deep-fried by the time we left.²⁵

²⁰ justjared.buzznet.com/2008/01/30/ruiner-willis-verizon/comment-page-2/, accessed 3/2/2010

²¹ www.amazon.com/review/R1F1RKPF64UNCM, accessed 11/29/2009

²² jellyjules.com/?m=200604, accessed 11/29/2009

²³ amazon.co.uk/review/R3RD409NYAYBYY, accessed 11/29/09

²⁴ tpinotandprose.blogspot.com/2008/03/one-where-fortune-cookies-amaze-me, accessed 11/29/2009

²⁵ www.yelp.com/biz/kellys-roast-beef-natick, accessed 11/29/2009

Looking beyond *looks like* structures, the two verb classes have partially overlapping argument realization patterns. Both permit predicative raising, in which the verb is followed by a predicative AP (33a). They differ with respect to the availability of various types of clausal complements. Epistemic verbs permit extraposed *that*-complements and infinitival clauses with “raised” subjects, while for perceptual verbs *that*-clauses are out and infinitival clauses are lexically and dialectally conditioned (cf. Matushansky 2002). This is illustrated in (33b-c).

- (33) a. The cake seems / appears / looks / sounds / smells / feels / tastes incredible.
 b. It seems / appears / *looks / *sounds / *smells / *feels / *tastes that the cake is incredible.
 c. The cake seems / appears / %looks / %sounds / *smells / *feels / *tastes to be incredible.

These distinctions presumably emerge from the subcategorization constraints associated with each verb.

3.2 Specifying the Evidential Source

This section turns to the ways in which evidential sources are specified in *looks like* structures. We start with the observation that verbs differ with respect to the constraints they place on the modality through which the evidence must be obtained, with consequences for subject selection. All five perceptual resemblance verbs are associated with a particular sensory modality, whereas *seem* and *appear* are not (though they often carry at least a slight preference for visual evidence). In the typical case, sentences headed by perceptual resemblance verbs incorporate evidence that is gathered via the lexically-specified modality. For example, *sounds like* sentences typically often involve auditory evidence. However, the verbs *look*, *sound*, and *feel* also permit *evidential bleaching*, in which

evidence comes from outside the lexically-specified modality (cf. Heycock 1994: 289-290). In fact, the evidential source for these verbs need not be perceptual at all; evidence may also come from hearsay or inference. In contrast, *taste* and *smell* usually require modality-specific perceptual evidence.²⁶ The sentences in (34a) do not entail that the speaker has visual, aural, or tactile evidence in favor of the embedded proposition; she may have reached her conclusion after having overheard a conversation between the judges of the competition. In contrast, the sentences in (34b) require that the speaker possess olfactory or gustatory evidence.

- (34) a. Ezra's cake looks/sounds/feels/seems/appears like it is going to win the prize.
b. Ezra's cake smells/tastes like it is going to win the prize.

Furthermore, as the naturally-occurring sentence in (35a) and its constructed counterpart in (35b) indicate, it is possible to target both the modality-specific and the evidentially-bleached readings of *look* and *sound*, indicating that the two readings are distinct. Because both verbs undergo bleaching, they permit an interpretation in which the matrix verb is bleached but the embedded verb is not; (35a) means something along the lines of "evidence from inference indicates that making a bathing suit is harder than is suggested by visual evidence."

- (35) a. Judging from the various ass and cooch-fitting issues, it looks like making a bathing suit is probably harder than it looks.²⁷

²⁶ As Larry Horn points out, there are fixed expressions in which *smell* is evidentially bleached, e.g. *that smells fishy*. In addition, I found one CR structure with evidentially bleached *smell*, which has an innovative flavor.

(i) With Miramax backing it, stars like Meryl Streep and Philip Seymour Hoffman cast in the leads and an opening date set in December, the project smells like someone's seriously looking to score a few awards.
<http://www.moviehamlet.com/Blog/1760/someones-looking-for-more-oscars>, accessed 11/29/2009

²⁷ <http://projectrunday.blogspot.com/search/label/Season%201%20-%20Episode%206>, accessed 11/29/2009

- b. Judging from the disastrous results of the music composition seminar, it sounds like writing 12-tone music is probably harder than it sounds.

The (non)-availability of evidential bleaching is associated with different constraints on the interpretation of main-clause subjects. Verbs that forbid evidential bleaching require that the subject serve as the perceptual source. In order to felicitously utter (34b), the speaker must have had firsthand perceptual experience of Ezra's cake. In contrast, the verbs that allow evidential bleaching do not require this interpretation. Note that in (34a), *Ezra's cake* makes an acceptable subject for *sound* despite the fact that it is impossible for a cake to provide auditory evidence. From this point out, I will focus on the verbs that permit evidential bleaching.

There is considerable additional evidence that the subjects of *look*, *sound*, and *feel*, in addition to the epistemic verbs *seem* and *appear*, need not be evidential or perceptual sources (contra Asudeh and Toivonen 2005, 2006, 2009). The sentences in (35) above demonstrate that the evidential source can be expressed through an adjunct phrase: in (35b), the evidential source is the disastrous results of the music seminar. It is possible for CR to co-occur with an adjunct denoting the source of evidence, indicating that the subject is not assigned that role. The naturally-occurring sentences in (36) illustrate this.

- (36) a. Judging from the short he looks like he's going to be fantastic.²⁸
b. Judging from the press conference she seems like she needs/wants the attention.²⁹

This stands in contrast with Asudeh and Toivonen's (2005: 10-11) observation that in Swedish, CR subjects cannot co-occur with *på*-adjuncts, which denote the perceptual

²⁸ www.metafilter.com/46809/Newest-Doctor-Who, accessed 11/29/2009

²⁹ www.topix.com/forum/family/T94T7O9LQ4888P607, accessed 11/29/2009

source. They argue that the constraint emerges from the fact that CR subjects express perceptual sources, and that multiple instances of the same event role within a single sentence are forbidden. The same argumentation leads us to a different conclusion regarding CR in English: because CR can co-exist with adjuncts that denote the source of evidence, they do not conventionally function as evidential sources.

There also exist CR sentences in which the subject referent is not a plausible evidential source for the proposition in question. This is exemplified by (37).

- (37) For instance, this year Barack Obama looks like he's ahead, on average, by two or three percentage points.³⁰

Barack Obama himself could hardly serve as evidence that he is up two to three percentage points in the polls. It is certainly possible that his appearance or behavior might give the impression that he is ahead, but it is unlikely to be the basis for a precise statistical claim.

Finally, there are cases in which the reverse holds: in which the evidential source, realized as an adjunct, would not be an acceptable subject. This is illustrated by the contrasts between the (a) and (b) sentences below.

- (38) a. [Hopefully Tom brings a bit more charisma to the table than the old Gethbot did] – judging from the Internet videos already provoking ringer outrage, he looks like he'll probably be an even better dancer as well.³¹
b. # The Internet videos already provoking ringer outrage look like he'll probably be an even better dancer as well.
- (39) a. [I doubt her motives were pure.] Judging from the press conference she seems like she needs/wants the attention.³²
b. # The press conference seems like she needs/wants the attention.

³⁰ blog.newsweek.com/blogs/stumper/archive/2008/08/20/expertinent-why-obama-needs-a-big-convention-bump.aspx, accessed 11/29/2009

³¹ <http://monkseal.wordpress.com/2008/08/28/strictly-come-dancing-preview-2/>, accessed 11/29/2009

³² www.topix.com/forum/family/T94T7O9LQ4888P607, accessed 11/29/2009

Despite the ample evidence that CR subjects need not be evidential sources, it is undeniable that they are interpreted this way most of the time. In Section 4, I argue that this emerges naturally from the topicality analysis presented here.

3.3 Epistemic Strengthening

Epistemic strengthening, in which the hearer infers that the speaker believes the proposition expressed by the embedded clause, also has consequences for the interpretation of subjects in *looks like* structures. Generally, epistemic strengthening takes place unless the speaker provides some indication that it should not. Thus, upon hearing (40a) the hearer is likely to infer that the speaker believes that Charley kissed Francine. The inference can be blocked in at least three ways. The first is explicit disavowal of the embedded proposition, illustrated by (40b). The second is contrastive stress on the main verb, which highlights a discrepancy between what the evidence suggests and what the speaker actually believes (40c). The inference can also be blocked by an embedded proposition that is clearly implausible (40d). In most cases, the hearer of (40d) will infer that the speaker intends to draw a comparison between Charley's appearance and the proposition that he spent time in a black hole, without commitment to that proposition.³³

- (40) a. Charley looks like he kissed Francine.
b. Charley looks like he kissed Francine, but he didn't.
c. Charley LOOKS like he kissed Francine.
d. Charley looks like he spent a few weeks in a black hole.

³³ Only a particularly clueless hearer would target this proposition for disagreement, as illustrated by (i) – unless, of course, the speaker is known to have unusual beliefs about space travel.

(i) A: Charley looks like he spent a few weeks in a black hole.
B: (?) But he didn't!

Recall that Rogers (1973: 78-79) argues that CR and its expletive-subject counterparts differ with respect to the epistemic commitment of the speaker. He claims that expletive-subject sentences, unlike CR sentences, require epistemic commitment, resulting in a truth-conditional difference between the two constructions. However, this conclusion is too strong. In this section, I will demonstrate that both constructions allow (but do not require) epistemic strengthening, though it is more strongly preferred in expletive-subject sentences.

I have identified several diagnostics that indicate whether a particular *looks like* structure requires epistemic strengthening. They can be illustrated by a third *looks like* construction in which epistemic strengthening is obligatory. This construction has *no* overt subject; for this reason, I call it *zero*. While zero sentences are prescriptively dispreferred, they are nevertheless quite common in conversation and in informal written genres. The following example illustrates the three-way subject alternation.

- (41) a. He looks/seems like he's from Wisconsin.
b. It looks/seems like he's from Wisconsin.
c. Looks/seems like he's from Wisconsin.

For an in-depth investigation of the semantics and pragmatics of zero sentences, see Mack and Fuerst to appear. For the present purposes, it is sufficient to demonstrate that zero sentences require epistemic strengthening. This is shown through their incompatibility with the following phenomena: overt disavowal, contrastive focus on the main predicate, interpretations in which the judge of evidence is not the speaker, and the felicitous reading of a certain type of ambiguity that emerges in equative and comparative sentences. In

contrast, CR and expletive-subject sentences pass all of these diagnostics, indicating that they do not require epistemic strengthening.

Disavowal

If the speaker can felicitously deny that she believes in the proposition contained within a *looks like* structure, it follows that the construction she chose does not require epistemic strengthening. Both CR (42a) and expletive-subject sentences (42b) allow explicit disavowal, albeit the latter less readily. In contrast, zero sentences (42c) cannot be followed by disavowal, indicating that they conventionally express epistemic commitment.

- (42) a. Charley looks like he kissed Francine, but he didn't.
b. (?) It looks like Charley kissed Francine, but he didn't.
c. # Looks like Charley kissed Francine, but he didn't.

Contrastive Stress

As discussed above, contrastive stress on the main predicate of a *looks like* structure indicates that the speaker intends to highlight a discrepancy between the conclusions that could be drawn from a body of evidence and what she actually believes. In (43a-b), the speaker acknowledges that there is evidence in support of the proposition that Charley kissed Francine, but nevertheless resists an epistemic commitment. Again, we observe a sharp contrast between zero sentences and sentences with overt main-clause subjects; the latter forbid contrastive stress, while the former allow it.

- (43) a. Charley LOOKS like he kissed Francine.
b. It LOOKS like Charley kissed Francine.
c. # LOOKS like Charley kissed Francine.

Speaker-Oriented Meaning

A similar contrast emerges when we examine the distribution of possible judges in *looks like* structures. The *judge* is the individual that evaluates whether a given proposition holds (cf. Lasersohn 2005, Stephenson 2007). In *looks like* structures, the judge is often covert, but it can also be realized as an adjunct prepositional phrase that follows the verb, as in the examples below.

- (44) a. As far as she's concerned, she's in a habitat where it looks to her like there are no males around.³⁴
b. And what exactly IS the gooey center? It looks and tastes to me like they just got all the ingredients and stuck them in a big masher.³⁵

When the judge is not overtly expressed, it is typically interpreted as being the speaker. This corresponds to Lasersohn's (2005) claim that the *autocentric*, i.e. speaker-oriented, interpretation is the default in perspective-dependent sentences. However, when epistemic strengthening occurs, the sentence serves as an expression of the speaker's beliefs. Thus, we would expect that when epistemic strengthening holds, the autocentric interpretation is in turn strengthened, from a preference to a requirement. This is exactly what we observe. CR (45a) and expletive-subject sentences (45b), which do not require epistemic strengthening, permit the realization of a judge other than the speaker. However, some speakers find non-autocentric judges easier to accommodate in CR than in expletive-subject sentences. Zero sentences (45c) do not allow them at all, due to the obligatory status of epistemic strengthening. Crucially, the source of this contrast is not syntactic, as all three constructions allow the speaker to be overtly specified as the judge (46).

³⁴ en.wikinews.org/wiki/Virgin_lizard_reproduces, accessed 11/26/2009

³⁵ www.chocablog.com/reviews/cadbury-starbar/, accessed 11/26/2009

- (45) a. Charley looks to Sam like he kissed Francine.
 b. It looks to Sam like Charley kissed Francine.
 c. # Looks to Sam like Charley kissed Francine.
- (46) a. Charley looks to me like he kissed Francine.
 b. It looks to me like Charley kissed Francine.
 c. Looks to me like Charley kissed Francine.

“Certain Ambiguities”

The final diagnostic that detects constructions that require epistemic strengthening is a certain class of ambiguities that arise in equative and comparative sentences (Lakoff 1970, Horn 1981a). An example appears below (Horn 1981a: 325).

- (47) Jack assumed Mary was older than she was.

There are two readings of (47): a felicitous *de re* reading in which Jack is mistaken about Mary’s age and an infelicitous *de dicto* reading in which Jack makes a contradictory assumption. Horn points out that CR sentences, but not their expletive-subject counterparts, readily permit the felicitous reading, as the examples in (48) demonstrate (Horn 1981a: 356; his judgments). On this basis, he constructs the following argument that at least some CR subjects are licensed through raising. Because CR sentences allow the *de re* reading, example (49), which also allows it, must be an example of CR. This means matrix subject of (49) is *weather-it*, which is selected only by meteorological predicates (*rain*, *snow*), and therefore must have been raised from the embedded-clause subject position.

- (48) a. John looks/sounds like he’s sicker than he (really) is.
 b. ?# It looks/sounds like John is sicker than he (really) is.
- (49) It looks/sounds like it’s raining harder than it is.

I will pursue an alternative account of the distinction between (48a) and (48b) that is based on epistemic strengthening. When epistemic strengthening occurs, only the

infelicitous reading of the ambiguity is possible. This is because the speaker expresses her belief in the proposition expressed by the embedded clause, which in the examples we have seen is contradictory. When epistemic strengthening is absent, the felicitous *de re* interpretation emerges, because the speaker simply asserts that there is evidence compatible with the proposition, but not that she believes it. Therefore, (48a) is better than (48b) because, as we have seen, CR can escape from epistemic strengthening more easily. However, this is also possible for expletive-subject sentences; example (50), which blocks epistemic strengthening via contrastive stress on the main verb, allows the felicitous *de re* interpretation of the ambiguity. As further evidence for this analysis, also note that zero sentences (51) allow only the contradictory *de dicto* interpretation.

- (50) It LOOKS like John is sicker than he actually IS.
(51) # Looks/seems like John is sicker than he is.

To summarize this section, I have presented several strands of evidence in support of the claim that neither CR nor expletive-subject sentences require epistemic strengthening, contra Rogers (1973). Still, we are left with the question of why epistemic strengthening is more strongly preferred in expletive-subject sentences. I will argue that this is the result of the tendency to interpret CR subjects as evidential sources, which in turn is a consequence of the fact that they function as topics. We will revisit this point at the end of the following section.

4. CR Subjects as Topics

This section presents evidence that CR subjects are licensed to serve as topics. Section 4.1 starts with the evidence that their interpretation is shaped by the pragmatic type

restrictions associated with topicality. We also discuss the significance of a small body of data that violate these constraints. Section 4.2 demonstrates that the distribution of CR is affected by a constraint that favors topic continuity within sentences, while Section 4.3 extends the observation to adjacent topic-marking constructions. In Section 4.4, I present evidence that CR subjects, like subjects in TE, function as *subjective topics*, in which relevance is established through the comment's influence on the speaker's impression of the topic referent. In Section 4.5, I argue that topicality is the missing link that is necessary in order to understand the connections between evidential bleaching, epistemic strengthening, and subject selection that were observed in Section 3.

4.1 Pragmatic Type Restrictions

This section demonstrates that in general, CR subjects must be interpreted as making specific reference, a precondition for topicality. It starts with the evidence that Lappin (1984) and Heycock (1994) present concerning the interpretation of bare NPs, then turns to a range of new data, including indefinite NPs, NPs with numeric quantifiers, and copular clauses. At the end of the section, we discuss the challenges posed by CR sentences that violate the pragmatic type restrictions associated with topicality. I will claim that they are evidence for a second, purely syntactic route to licensing in CR (cf. Rogers 1973, 1974, Horn 1981a, Potsdam and Runner 2001).

4.1.1 Bare NPs

As discussed in Chapters 2 and 3, bare plurals and mass nouns are ambiguous between a specific *kind* reading and a non-specific *existential* reading (Carlson 1977, Laca 1990,

Diesing 1992, Kratzer 1995, Chierchia 1998, McNally 1998, Cohen and Erteschik-Shir 2002, Heycock and Doron 2003). We have seen that topic expressions require the former reading, due to the fact that topicality can only be established if the speaker has a specific individual or set in mind. Lappin (1984: 241) demonstrates that when bare NPs are licensed as subjects in CR, they are obligatorily interpreted as specific. The following paradigms are adapted from his examples.

- (52) a. It seems/looks as if cows are extremely intelligent.
b. Cows seem/look as if they are extremely intelligent.
c. It seems/looks as if cows are grazing in Fred's field.
d. # Cows seem/look as if they are grazing in Fred's field.
- (53) a. It seems/looks like sand is composed of tiny particles.
b. Sand seems/looks like it is composed of tiny particles.
c. It seems/looks like sand is blowing all over the backyard.
d. #Sand seems it is blowing all over the backyard.

The subordinate complement clauses in the (a) and (b) sentences of each paradigm favor a kind interpretation of the subject, while the complements of the (c) and (d) sentences favor an existential reading. Because CR subjects must be interpreted specifically, this leads to infelicity in the (d) sentences, in which the specific reading of the subject conflicts with the existential semantics of the complement. This suggests that CR subjects are constrained by the type restrictions that are characteristic of topicality.

Crucially, SSR is not subject to the same constraints. Consider the SSR counterparts of the paradigms above: bare NPs can be interpreted either as kinds or existentially.

- (54) a. It seems that cows are extremely intelligent.
b. Cows seem to be extremely intelligent.
c. It seems that cows are grazing in Fred's field.
d. Cows seem to be grazing in Fred's field.

- (55) a. It seems that sand is composed of tiny particles.
 b. Sand seems to be composed of tiny particles.
 c. It seems that sand is blowing all over the backyard.
 d. Sand seems to be blowing all over the backyard.

4.1.2 Indefinites and NPs with Numeric Quantifiers

Further evidence comes from the interpretation of indefinites and numerically quantified NPs, which may or may not be interpreted specifically, making reference to a particular individual or set. Previous researchers have argued that topical indefinites are obligatorily interpreted as specific (Cresti 1995, Portner and Yabushita 2001, Krifka et al. 2006, Endriss and Hinterwimmer 2007, Endriss 2009). This constraint conditions the interpretation of CR subjects, as shown by the sentences in (56). Expletive subject sentences (56a) permit non-specific readings of indefinite embedded-clause subjects, as shown by the felicity of the continuation sentence. In contrast, CR sentences (56b) require the specific interpretation of the matrix subject. Example (57) extends the observation to subject NPs with numeric quantifiers. Again, SSR subjects are free from the specificity requirement (58-59).

- (56) a. It looks like an organizer is angry. (But I don't know which one.)
 b. An organizer looks like he's angry. (# But I don't know which one.)
 (57) a. It sounds like two dishes contain cilantro. (But I don't know which ones.)
 b. Two dishes sound like they contain cilantro. (# But I don't know which ones.)
 (58) a. It seems that an organizer is angry. (But I don't know which one.)
 b. An organizer seems to be angry. (But I don't know which one.)
 (59) a. It seems that two dishes contain cilantro. (But I don't know which ones.)
 b. Two dishes seem to contain cilantro. (But I don't know which ones.)

A possible objection to my interpretation of these contrasts is that in sentences like (56b), the subject is interpreted as the evidential source and thus specificity comes for free: if the judge is reporting on direct perceptual experience of the subject referent, it follows

that she has that referent in mind. However, the specificity requirement is independent of the preference to interpret subjects as evidential sources. In (57b), the subject is not an evidential source but it is nevertheless interpreted specifically.

A closely related observation is that topical quantified NPs tend to be interpreted with wide scope (Portner and Yabushita 1998, 2001, Krifka et al. 2006, Endriss and Hinterwimmer 2007, Endriss 2009). Krifka et al. (2006) argue that this is a consequence of the fact that topics appear within the restrictive clause of a quantifier, rather than in its nuclear scope (cf. Partee 1992). Previous researchers have observed that CR subjects obligatorily take wide scope with respect to the main predicate (Lappin 1984: 240, Heycock 1994: 294, Potsdam and Runner 2001: 11). Potsdam and Runner illustrate this through the following contrast.

- (60) a. Two people seem to have won the lottery.
 [2>*seem*, *seem*>2]
 b. Two people seem like they won the lottery.
 [2>*seem*, **seem*>2]

They claim that in SSR (60a), the subject quantifier can take either wide or narrow scope with respect to *seem*, while in CR (60b) only wide scope is available. Note that in their analysis, scope is detected through restrictions on the evidential source. In the wide scope reading, the subject is interpreted as the evidential source, resulting in a reading in which two people's appearances give the impression that they won the lottery. In the narrow scope reading, the two people under discussion are not the evidential source.

There are multiple ways to interpret the scope data. For Potsdam and Runner, it supports the view that the matrix subject in CR is base-generated, rather than raised from

the embedded clause as in SSR. On the present analysis, it is a direct consequence of the requirement that subjects function as topics.

4.1.3 Copular Clauses

More support for the topicality analysis comes from constraints on the distribution of copular clauses. Within the class of copular clauses, there is a well-known distinction between *predicational* clauses such as (61a) and *specificational* clauses like (61b).

- (61) a. Obama is the most popular candidate.
b. The most popular candidate is Obama.

Mikkelsen (2004: 1), citing Akmajian (1979), describes the distinction in the following way: “a predicational copular clause tells us something about the individual denoted by the subject NP, whereas a specificational copular clause tells us who (or what) someone (or something) is.” This semantic distinction is accompanied by morphosyntactic differences, for example with respect to pronominalization and distribution (Akmajian 1979, Heycock and Kroch 1999, 2002, Mikkelsen 2004).

Both clause types can appear within the complement clause in *looks like* structures, as illustrated by (62). However, main-clause subjects in CR can only co-refer with the subjects of predicational clauses (63a), not with the subjects of specificational clauses (63b).³⁶

³⁶ There is a comparable restriction in small clauses:

- (i) I consider iced tea the most popular choice.
(ii) (#) I consider the most popular choice iced tea.

This may be related to another parallel between the two constructions. Like CR subjects, subjects of small clauses are preferentially interpreted as objects of direct perception (Borkin 1973, Newman 1982); consider the following contrast of Borkin's, where the implication of direct perception is stronger in the small clause than in the full clause.

- (iii) I find this chair uncomfortable
(iv) I find that this chair is uncomfortable.

- (62) a. It seems like Obama is the most popular candidate.
 b. It seems like the most popular candidate is Obama.
 (63) a. Obama seems like he's the most popular candidate.
 b. # The most popular candidate seems like it's Obama.

This is due to the requirement that topic expressions refer, which is even more basic than the specificity constraint. At least since Donnellan (1966), it has been recognized that definite descriptions have a *referential* reading and an *attributive* reading, in which the NP functions as a predicate (or in Mikkelsen's 2004 terms, a *property*). Definite descriptions receive only the referential reading when they function as topics, as in CR. This accounts for the fact that the subject NP in (64) can only be interpreted as referential.

- (64) The most popular candidate seems like he's from Hawaii.

4.1.4 Violations of Pragmatic Type Restrictions

Despite the compelling evidence for the topicality analysis that we have seen so far, there exists a small body of data which clearly violate the pragmatic type restrictions associated with topicality. First, a non-referential NP can appear as the subject in CR if it has a co-referring pronoun in the subject position of the embedded clause. We have already seen examples involving expletive *there* (65a), idiom chunks (65b), and weather *it* (65c), and observed that licensing of a non-referential subject is impossible when the co-referring pronoun appears elsewhere (66). (Due to the formal identity between weather *it* and the *it* of extraposition, it is impossible to construct an unambiguous counterpart sentence for (65c).)

- (65) a. % There seems like there's a problem.
 b. % The shit seems like it hit the fan.

It may be possible to provide a unified account of these phenomena if we assume, following Basilico (2003), that the subjects of adjectival small-clauses, like CR subjects, are topics.

- c. It looks like it's raining harder than it is.
- (66) a. * There seems like John expects there to be a problem.
- b. * The fan seems like the shit hit it.

For many speakers, quantified NPs that are not possible topics can also appear as subjects when they are associated with an embedded-subject copy. Lappin (1984: 242) notes that the quantifier *no* is possible (67a); I extend his observations to include *every* (67b). Both quantifiers are generally banned in topic expressions, as shown by their ungrammaticality in left dislocation (69).³⁷ Like the non-referential NPs discussed above, NPs quantified by *no* and *every* are impossible when there is a co-referring pronoun outside of the embedded-clause subject position (68).

- (67) a. % No referee seems like he saw the goalie.
- b. % Every referee seems like he saw the goalie.
- (68) a. * No goalie seems like the referee saw him.
- b. * Every goalie seems like the referee saw him.
- (69) a. * No referee, he saw the goalie.
- b. * Every referee, he saw the goalie.

This raises the question of how to reconcile the considerable evidence for the topicality analysis with the observation that some CR subjects are clearly not topical. In fact, the situation is even more complicated than this. Consider that all of the examples discussed in Sections 4.1.1-4.1.3, in which the NP type of the subject is referentially ambiguous, contain a co-referring pronoun in the embedded-clause subject position. If sentences with subject copies were free from the topicality requirement, as the evidence presented in this subsection suggests, it would fail to predict the effects of pragmatic type restrictions that we have observed. The generalization, odd as it is, appears to be as follows. NPs that *can* be

³⁷ See Endriss (2009) for a formal account of why topic expressions can only contain certain quantifiers.

interpreted as topics *are*, regardless of the location of the co-referring pronoun. NPs that are not possible topics are permitted only when they have a subject copy.

I believe that this constitutes strong evidence for dual mechanisms of subject licensing in CR, as suggested by Rogers (1973, 1974), Horn (1981a), Potsdam and Runner (2001), Asudeh and Toivonen (2005, 2006, 2009), and Sag (2010). I will return to the details of my proposal in Section 5.

4.2 Intra-Sentential Topic Continuity

This section presents a second strand of evidence that CR subjects are topic expressions: there is a strong preference for main-clause subjects to co-refer with the topic of the embedded clause, if there is one. This is shown by constraints on the interpretation of matrix subjects when the embedded clause is headed by a symmetric predicate, such as *resemble*, *be similar to*, and *be married to* (70-72). Recall from Chapter 2, Section 3.1.5, that it has previously been proposed that the subject of a symmetric predicate conventionally functions as a topic (Dowty 1991, Gleitman et al. 1996, Bowdle and Gentner 1997).

- (70) a. Iris resembles Ezra.
b. Ezra resembles Iris.
- (71) a. Iris is similar to Ezra.
b. Ezra is similar to Iris.
- (72) a. Iris is married to Ezra.
b. Ezra is married to Iris.

This accounts for the observation that the matrix subject in CR typically co-refers with the embedded subject of a symmetric predicate (73-75a). When there is no co-reference, it results in a certain degree of infelicity (73-75b).

- (73) a. Iris sounds like she resembles Ezra.
b. ? Iris sounds like Ezra resembles her.

- (74) a. Bill seems like he met John.
 b. ? Bill seems like John met him.
- (75) a. Bill feels like he's similar to John.
 b. ? Bill feels like John is similar to him.

This observation extends to predicate pairs like *taller* and *shorter*, which encode identical relations from different perspectives, with the subject presumably serving as a topic. Here too, the matrix subject in CR preferentially co-refers with the embedded subject.

- (76) a. Bill sounds like he's taller than John.
 b. ? Bill sounds like John is shorter than him.

4.3 Topic-Marking Constructions

A related observation is that there is a strong tendency for CR subjects to co-refer with the topic expressions of adjacent topic-marking constructions, such as *What about X?* (77) and prolepsis (78). Because *Iris's cookies* are the topic of (77-78a), it is far more acceptable for the same referent (77-78B1) than for a different referent (77-78B2) to be realized as the subject of CR.

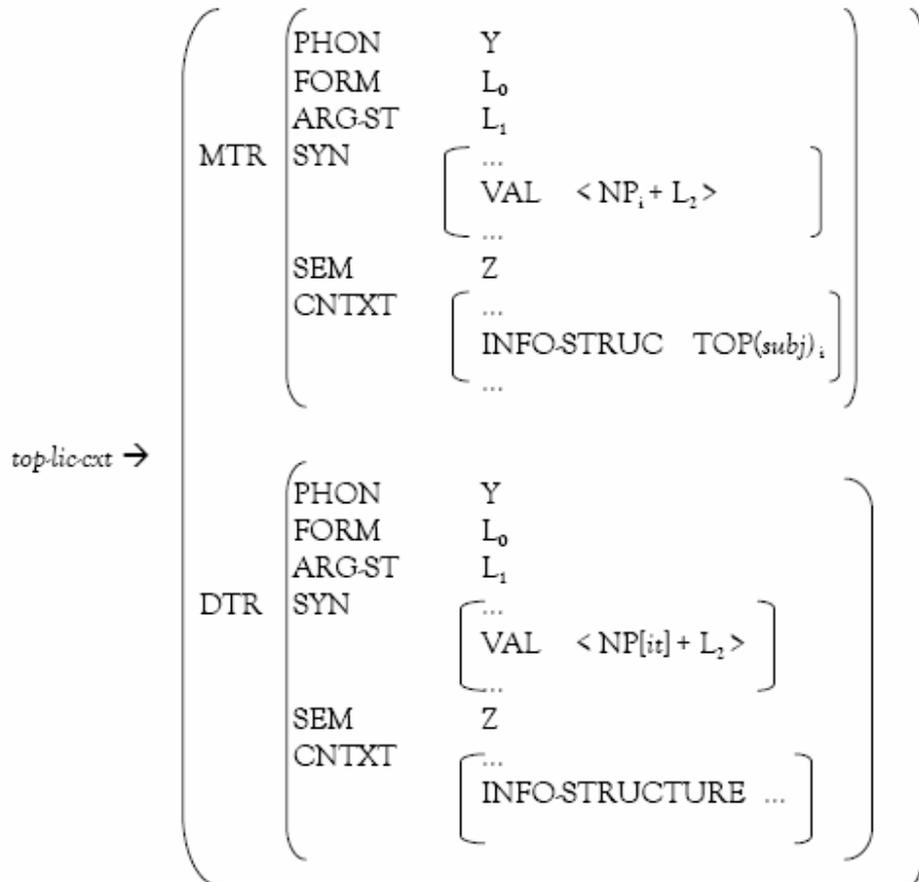
- (77) A: What about Iris's cookies?
 B1: They sound like they were beaten by Ezra's cake.
 B2: (#) Ezra's cake sounds like it beat them.
- (78) a. She said about Iris's cookies that they sound like they were beaten by Ezra's cake.
 b. (#) She said about Iris's cookies that Ezra's cake sounds like it beat them.

4.4 CR Subjects as Subjective Topics

In Chapter 3, I argued that TE instantiates a particular type of link between the topic and the comment: the subject serves as a *subjective topic*, meaning that the speaker's impression of the topic is affected by the comment. I proposed that this constraint is part of the Topic

Licensing construction, which also underlies CR. For reference, the Topic Licensing construction appears in Figure 1 below.

Figure 1. Topic Licensing: Final Version



There is evidence that CR subjects are also subjective topics, as this analysis predicts. This can be seen most clearly in copy-less structures, where there is no direct way in which to establish the topic-comment link. In sentences with pronominal copies, the topic referent is an argument of the embedded proposition, which is itself an argument of the comment, so topicality can be established on this basis.³⁸ In copy-less structures such as (79) below (Rogers 1973: 99-100), the subject is licensed purely on the basis of its function as a

³⁸ In terms of the typology presented in Chapter 3, this is a link of Type 5.

subjective topic. In (79b), for example, the speaker's use of CR is a signal to the hearer that his stance towards the soup has been influenced by the proposition that it tastes like Maude has been at the cooking sherry.

- (79) a. The soup tastes to me like Maude has been at the cooking sherry.
b. The orchestra sounds to me like Mehta has been having a good night.
c. The peanut butter feels to me like David forgot to put the lid on the jar.

This further supported by the fact that some evidential sources, which make perfectly good topics in general, are not plausible subjective topics and thus are odd as subjects in CR. In Section 3.2, we saw that the adjunct evidential sources in (80-81a) are not felicitous as subjects (80-81b).

- (80) a. [Hopefully Tom brings a bit more charisma to the table than the old Gethbot did] – judging from the Internet videos already provoking ringer outrage, he looks like he'll probably be an even better dancer as well.³⁹
b. # The Internet videos already provoking ringer outrage look like he'll probably be an even better dancer as well.
- (81) a. [I doubt her motives were pure.] Judging from the press conference she seems like she needs/wants the attention.⁴⁰
b. # The press conference seems like she needs/wants the attention.

This is because they are not subjective topics. It is intuitively clear that the comments in (80-81a) reflect the speaker's impression of the subject referent, rather than the evidential source. Infelicity in (80-81b) results from the fact that CR subjects are required to be subjective topics. In contrast, as I argued in Chapter 3, the *As for X* construction allows a wide range of topic-comment links. In this pragmatically permissive environment, the evidential sources of (80-81a) are possible topics.

³⁹ <http://monkseal.wordpress.com/2008/08/28/strictly-come-dancing-preview-2/>, accessed 11/29/2009

⁴⁰ www.topix.com/forum/family/T94T7O9LQ4888P607, accessed 11/29/2009

- (82) a. As for the internet videos, it looks like he'll be an even better dancer as well.
b. As for the press conference, it seems like she needs/wants the attention.

4.5 Topicality, Evidential Sources, and Epistemic Strengthening

We are now in a position to return to the connections between subject selection, the expression of the evidential source, and the availability of epistemic strengthening, initially discussed in Section 3. I propose that these connections are mediated by topicality. We start with the connection between topicality and the expression of the evidential source. Because CR subjects are topics, they are typically – but as we have seen, not invariably – interpreted as evidential sources. This is because an evidential assertion that influences the speaker's impression of an individual will often (but not always) come about through direct perceptual experience with that individual.

The expression of the evidential source is in turn related to the possibility of blocking epistemic strengthening. When a judge acquires evidence through direct perception that is compatible with a particular proposition, it is entirely possible for her to withhold belief in that proposition, perceptual evidence notwithstanding. For this reason, when the hearer encounters a *looks like* structure that clearly specifies a perceptual source, she can easily block epistemic strengthening if she feels it conflicts with the speaker's intent. In contrast, when the speaker's evidence for a proposition comes about via inference, it is quite difficult (if not impossible) for her to fail to believe it. Thus, if the hearer suspects that the evidence for a *looks like* claim came about via inference, she is very unlikely to block the inference of epistemic strengthening.

My proposal is that CR blocks epistemic strengthening more readily than *it*-extraposition structures because the subject is typically interpreted as an evidential source (which I have argued is a consequence of topicality). Because the subject of (83a) is preferentially taken to be the source of evidence, it facilitates the blocking of epistemic strengthening. In contrast, (83b) does not specify an evidential source, leaving open the possibility that the supporting evidence is inferential. This makes it relatively difficult, though not impossible, to block epistemic strengthening.

- (83) a. Charley looks like he kissed Francine, but he didn't.
b. (?) It looks like Charley kissed Francine, but he didn't.

This effect is found not only with CR subjects, but also with adjuncts that express the source of evidence. When the evidential source is specified through an adjunct, as in (84), it enables the hearer to block epistemic strengthening, even in expletive-*it* sentences.

- (84) From the lipstick on Charley's collar it looks like he kissed Francine, but he didn't.

Thus, I conclude that the correlations between subject selection and semantics in *looks like* structures discussed in Section 3 can be captured through the topicality analysis. Topics are typically interpreted as evidential sources, which in turn provide a means of blocking epistemic strengthening.

This approach also accounts for the observation that CR subjects are more strongly interpreted as evidential sources than their SSR counterparts. Recall the following contrast from Section 2.

- (85) Iris: Two of our committee members seem to be sick today.
Ezra: Actually, they're just late.
(86) Iris: Two of our committee members seem like they're sick today.
Ezra: # Actually, they're just late.

As I demonstrated in Section 4.1, SSR subjects are not required to function as topics. For example, they are free from the pragmatic type constraints that apply to CR subjects. The close relationship between topicality and the source of evidence provides an explanation for the contrast between (85) and (86): because the CR subject is a topic, there is a stronger inclination to interpret it as an evidential source.

5. The Second Route to Licensing

This section touches briefly on the second way in which CR subjects are licensed: through a purely syntactic construction in which the main-clause subject is associated with a co-referring pronoun in the embedded-clause subject position. Working within SBCG, Sag (2010: 47) proposes that verbs that license CR are associated with the ARG-ST list shown in (87).

(87) [ARG-ST: <NP,_i PRT[*like*], S [XARG <NP_i[*pro*]>]>]

This states that CR verbs select for an NP, a particle (specifically *like*), and a clausal complement whose external argument is a pronoun co-indexed with the first ARG-ST member.

I will suggest two modifications to Sag's analysis. First, I propose that subject-to-subject CR is implemented through a construction that alters the valence, rather than the argument structure, of the main predicate. This is because I assume that the input to subject-to-subject CR, like that of Topic Licensing, is an extraposition structure. Sag himself (2010: 41) proposes that extraposition is a post-inflectional construction that alters valence (see further discussion in Chapter 3, Section 5). To use a procedural metaphor, because both constructions that license CR subjects must "follow" *it*-extraposition, they too must operate

on the valence list. Second, following several researchers, I assume that the extraposed complement in CR is a prepositional phrase that in turn selects for a finite clause (Maling 1983, Heycock 1994: 288-298, Potsdam and Runner 2001, Asudeh 2004: 373-377).

The constructional representation that I propose for Subject-to-Subject Copy-Raising (SSCR) appears in Figure 2.

Figure 2: Subject-to-Subject Copy Raising

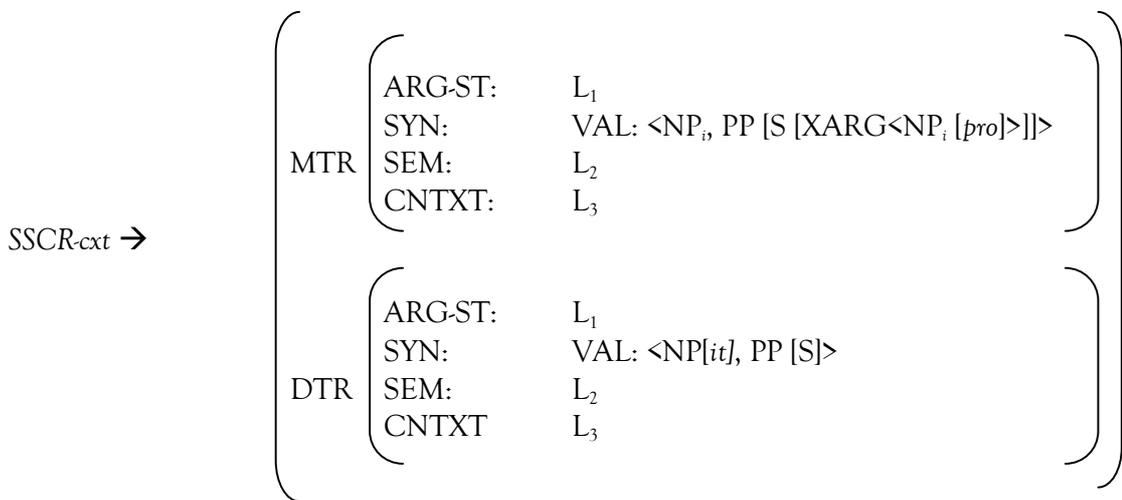


Figure 2 states that an SSCR construct ($\text{SSCR}_{\text{-cxt}}$) is well-formed if the valence list of the input sign contains expletive *it* and a PP that takes a sentential complement, while the valence list of the output sign replaces expletive *it* with a NP that is co-indexed with the external argument of the embedded S. All other levels of representation are unchanged.

My analysis predicts that sentences like (88), which contain a referential matrix subject with an embedded-subject “copy,” should have two possible source constructions: Topic Licensing and SSCR.

- (88) a. John looks like he's sick.
b. Cows seem like they're extremely intelligent.
c. A man from Hawaii sounds like he's the most popular candidate.

However, we have seen that potentially referential subjects such as bare plurals (88b) and indefinites (88c) must obey the type constraints that the Topic Licensing construction imposes. This is unexpected if they could alternatively be licensed via SSCR.⁴¹ This relates to the generalization that was discussed in Section 4: CR subjects with embedded-subject copies are interpreted as topics if they can be, but otherwise non-topic NPs are permitted. I do not presently have an explanation for why this should be the case, but I will present some ideas that previous researchers have offered.

In order to account for similar observations about CR, Rogers (1974) considers proposing a trans-derivational constraint that blocks the purely syntactic route to licensing (for Rogers, a raising operation) when the structure of the sentence is consistent with the route that comes with semantic constraints. He concludes, however, that this approach is more of a description of the problem than a solution (p. 556).

Schmerling (1978) proposes a raising-to-object analysis of the verb *allow* that has interesting parallels with the case of CR. She starts with the observation that pairs such as (89) are usually judged as non-synonymous. This has led previous researchers to claim that *allow* is an object-control verb that assigns a semantic role to the “permissée” (in (89a), *the doctor*) which also serves as the controller of the embedded clause. However, there are also cases in which the object of *allow* is clearly not the permissée, e.g. (90a), in which the object

⁴¹ Similarly, Rogers (1974: 552) notes that the lack of ambiguity of sentences like (88) is “embarrassing” for a dual-route licensing account, which he nevertheless endorses.

is expletive *there*, and (90b), given that it would be odd to grant permission to a newborn baby.

- (89) a. I allowed the doctor to examine John.
b. I allowed John to be examined by the doctor.
- (90) a. I allowed there to be an investigation.
b. I allowed the newborn baby to be examined by the doctor.

This raises the possibility that *allow* is ambiguous between a control and a raising verb, and thus there are two possible routes to licensing the object in sentences such as (89). However, Schmerling rejects an analysis along these lines, arguing instead that *allow* is uniformly a raising verb. She claims that the control-like interpretation of the sentences in (89) emerges from a general preference for maximally complex interpretations.

On this basis, we might initially entertain the possibility that the apparent dual route to licensing in CR is actually an illusion. However, there are compelling reasons to believe that this is not the case. First, recall that SSCR applies only to structures with co-referring subject pronouns, while Topic Licensing places no restrictions on the syntactic position of the copy, if indeed there even is one. We need Topic Licensing in order to account for CR sentences that lack embedded-subject copies, and we need SSCR to account for sentences containing subject copies in which the matrix subject is non-topical. Second, it is not possible to overrule the pragmatic type constraints on CR, in the way that Schmerling shows is possible for *allow* (90b). If pragmatic type constraints were simply the result of a preference for pragmatic complexity, then it would be possible to overrule them in sentences like (91). As we have observed, it is not.

- (91) a. * Cows seem like they're grazing in Fred's field.
b. * Snow seems like it's blowing all over the backyard.

Accordingly, the nature of the relationship between Topic Licensing and SSCR remains mysterious at the present time, and constitutes an interesting question for future research.

6. Conclusion

This chapter has provided evidence that the Topic Licensing construction that underlies subject licensing in TE also extends to CR. Furthermore, I have argued that the observed connections between subject selection, epistemic strengthening, and the specification of the evidential source are consequences of the topicality account. Together with TE, the case of CR illustrates that direct licensing is one way in which information structure influences subject selection in English. In Chapter 5, we turn to the evidence that subject selection is also shaped by resolution.

Chapter 5

Resolution: Instrument Subjects

1. Introduction

Chapters 3 and 4 investigated the phenomenon of direct licensing by information structure, demonstrating that it underlies subject selection in a restricted set of environments in English. This chapter turns to the second component of the interface model: resolution, in which information structure selects a subject from multiple candidates that satisfy the constraints imposed by the lexical semantic linking system. I will focus on the Instrument Subject (IS) alternation, illustrated below.

- (1) a. John broke the vase with the club.
b. The club broke the vase.
- (2) a. John opened the door with the key.
b. The key opened the door.
- (3) a. Melville wrote Moby Dick with this pen.
b. (#) This pen wrote Moby Dick.
- (4) a. John ate the spaghetti with the fork.
b. * The fork ate the spaghetti.

In the IS alternation, an instrument can either be realized as an adjunct, with an agent occupying the subject position (1-4a), or as subject (1-3b). As the examples above show, the distribution of instrument subjects in English is tightly constrained: *this pen* in (3) is acceptable as an adjunct but needs strong supporting context to be linked as subject, while

the fork in (4) is ungrammatical as subject.¹ In this chapter, I provide an account of the alternation in which both lexical semantic and informational constraints underlie the distribution of instrument subjects.

The IS alternation has historically played a central role in theories of argument realization, for a range of reasons. First, instrument subjects serve as a putative counterexample to the generalization that subjects of active sentences tend to be agents (though see e.g. Schlesinger 1989, Jackendoff 1990, DeLancey 1991, and Alexiadou and Schäfer 2006 for arguments that (certain) instrument subjects are conceptualized as agents, a claim that will be revisited in Section 2). For this reason, they pose something of a challenge to attempts to make generalizations about the semantics of subjecthood. Second, since the work of Fillmore (1968) the IS alternation has served as an illustration of the relationships that hold between participants in an event and how they are reflected in grammar. Specifically, Fillmore built upon the observation that instruments are potential subjects only when no agent is realized (as in 1-3b above) to argue for a thematic hierarchy in which agents outrank instruments for argument realization. Theories of linking that do not make use of thematic hierarchies (including the present one) must contend with this and similar observations. Finally, the IS alternation has played a key role in the study of the typology of linking systems. A large body of research has demonstrated that English exhibits more freedom in linking instruments as subjects than many other languages, including Dutch (Van Voorst 1996), French (Van Voorst 1996), German (Hawkins 1985), Irish (Guilfoyle 2000) and Korean (Wolff et al. 2009). Thus theories of linking are faced

¹ *The fork* is ungrammatical as subject when describing the same scene as (4a). It is acceptable only under a “fairy tale” reading in which a fork actually consumes spaghetti – but then, of course, it does not function as an instrument in the event.

not only with the question of why English permits instrument subjects relatively freely, but also why many other languages do not. For these reasons, the IS alternation provides an ideal environment in which to illustrate how the interface model of licensing fares when it is brought into direct contact with the foundational issues of a lexical semantic theory of linking.

The account of the IS alternation that I propose in this chapter hinges on the notion of resolution. For a particular class of event representations containing an agent and an instrument, the lexical semantic linking system does not determine which is to be linked as subject. The indeterminacy is “resolved” via pragmatic constraints on subject selection. Lexical semantic linking constraints are formulated as constraints on lexeme classes, specifically on the relationship between semantics (SEM) and argument structure (ARG-ST). This presupposes a hierarchical model of the lexicon (Davis and Koenig 2000, Davis 2001, Sag 2010) in which the features of more general classes of lexemes (e.g. change-of-state verbs) are inherited by more specific lexeme classes (e.g. causative change-of-state verbs). I propose one very general linking constraint, operating on lexemes whose semantics includes the primitive predicate ACT, denoting the transmission of force. The constraint states that the subject must be an *actor*, i.e. a transmitter of force. I will argue that the predicates that permit instrument subjects include ACT and accordingly inherit this constraint. Indeterminacy in the linking system emerges from the fact that both agents and instruments are actors and thus potential subjects. The relevant pragmatic constraint on subject selection is formulated as a derivational construction that permits the licensing

of an instrument as subject only when it is linked to an activated proposition (in a particular way that will be explained in Section 4).

There are several advantages of this analysis. First, as I will demonstrate, it is empirically stronger than previous accounts. For example, it delimits the class of predicates that permits instrument subjects more precisely than previous analyses, such as Reinhart (2002) and Grimm (to appear). Second, it provides a clear account of how constraints on instrument subjects connect to the linking system more generally. Finally, and most importantly, it makes explicit the contributions of lexical semantic and pragmatic constraints to the distribution of instrument subjects. It is the first analysis that I am aware of to provide a formal account of the construction's pragmatic component, though several previous accounts have suggested that information structure plays some role (e.g. Schlesinger 1989, DeLancey 1991, Brousseau 1998, Alexiadou and Schäfer 2006).

I will now touch upon three key aspects of the approach that I am taking in this chapter. The first is my operative definition of the notion *instrument*. I assume the following pre-theoretical definition: an instrument is a physical entity that is controlled by another entity in order to bring about an event.² This definition rules out the class of natural forces (e.g. *The storm damaged the house*), which cannot be controlled. It is, however, a fairly liberal definition, including a wide range of entities that some previous researchers have considered to be instruments, including machines (5a), chemical substances (5b), body parts (5c), animals construed as tools (5d), artifacts (5e), and natural kinds (5f).

- (5) a. John computed the answer with a calculator.
- b. The doctor cured the patient with chamomile.

² This definition excludes instrument-like abstract entities, such as the subject NP in *The new finding revolutionized the theory*. I will leave subjects of this sort as a topic for future research.

- c. John touched the insect with his left thumb.
- d. Mary plowed the field with her oxen.
- e. John opened the door with the key.
- f. Mary broke the vase with the rock.

These subclasses of instruments differ in one key respect: the degree to which the instrument is “potent,” that is, capable of generating force independently of the entity that controls it. Machines, chemical substances, draft animals, and body parts have a certain inherent potency that the agent is able to “unleash” in scenarios like (5a-d). In contrast, simple objects such as keys and rocks have no inherent force; the force that they transmit in scenarios like (5e-f) comes about entirely through the manipulation of the agent. One robust generalization that has appeared many times in the literature is that relatively potent instruments are realized as subjects more freely than less potent instruments (Schlesinger 1989, Levin 1993: 80, Kamp and Rossdeutscher 1994, Brousseau 1998: 110-114, Alexiadou and Schäfer 2006, Grimm to appear). Throughout this chapter, I will focus on the class of instruments that have no inherent force, because they provide the most conservative picture of the distribution of instrument subjects.

The second issue is the approach that I am taking in dealing with grammaticality/acceptability judgments. The literature on instrument subjects is plagued with judgments that conflict with each other or with what is attested in naturalistic data. For example, Schlesinger (1989: 193) judges that example (6a) is acceptable “only in certain contexts,” while Alexiadou and Schäfer (2006: 45) claim that it can never be acceptable. The judgments provided in the literature also tend to be more conservative than what can be found in actual discourse. In addition to *clean*, the verbs *murder* (Van Valin and Wilkins 1996: 310, Grimm to appear: 2) and *write* (Schlesinger 1989: 195) are among the predicates

that have been claimed never to take instrument subjects. Yet all can do so given an appropriate grammatical or discourse context, as the (b) sentences below illustrate.

- (6) a. ?/* The rag cleaned the dishes. (*judgments*: Schlesinger 1989, Alexiadou and Schäfer 2006)
b. My friend called me back and said that the cloth cleaned the chandelier in no time at all.³
- (7) a. * The bullet murdered the president. (*judgment*: Grimm to appear)
b. The chemical composition of these flashburn particles matched the type of bullet that murdered Jill Dando.⁴
- (8) a. * The pen writes a letter. (*judgment*: Schlesinger 1989)
b. This is the pen that wrote love letters to my mother, signed all my report cards, signed the tuition checks for college and everything else.⁵

It is certainly possible that dialectal, ideolectal, or register differences play some role in these discrepancies. I think that much of the problem, though, stems from the fact that instrument-subject sentences have typically been assessed for acceptability in isolation, without a supporting context. This is problematic because, as I will demonstrate, context plays a crucial role in shaping the distribution of instrument subjects. In my own attempt to delimit the class of predicates that permits instrument subjects, I take a liberal approach to grammaticality judgments, taking an instrument-subject sentence to be grammatical if an appropriate context can be constructed for it. In order to isolate the lexical semantic

³ <http://www.microfibersunlimited.com/Customers.htm>, accessed 11/19/2009

⁴ <http://forums.mirror.co.uk/viewtopic.php?t=50342&postdays=0&postorder=asc&start=60>, accessed 11/19/2009

⁵ <http://www.fountainpennetwork.com/forum/index.php?act=Print&client=printer&f=31&t=74773>, accessed 11/19/2009

constraints on instrument subjects, it is essential to factor out the independent constraints imposed by the pragmatic component.⁶

The final point is that the instrument “subject” phenomenon is not actually restricted to subjects. Like agents, instruments can undergo demotion in the passive construction, as shown by the fact that they can be realized as adjunct *by*-phrases.

(9) [Context: *Newspaper headline*]

Killed by the gun she gave as a gift.⁷

(10) Another point to keep in mind is that sterling silver should be cleaned by a cloth.⁸

This demonstrates that what we are investigating when we research instrument “subjects” are the lexical constraints through which instruments are chosen as external arguments (in the present framework, the first element of the ARG-ST list), rather than the syntactic processes through which they are realized as (surface) subjects. While keeping this crucial point in mind, I will continue to use the term *instrument subjects* for convenience.

This chapter begins with a review of previous literature on instrument subjects (Section 2), and then presents an account of the lexical semantic (Section 3) and pragmatic (Section 4) constraints on their distribution in English. Section 5 concludes.

2. Previous Analyses

In this section, I review previous attempts to formulate constraints on the distribution of instrument subjects, with particular focus on research dealing with subject selection in

⁶ As is standard, ungrammatical sentences are marked with an asterisk; sentences that are pragmatically odd are marked with the pound sign. The symbol (#) indicates that the sentence is infelicitous in the absence of appropriate contextual support.

⁷ <http://www.simplysark.info/forum/sark-news/2868-channel-news-killed-gun-she-gave-gift.html>, accessed 11/19/2009

⁸ www.ringsurf.com/online/2976-sterling_silver_rings.html, accessed 11/19/2009

English. We start with accounts that emphasize the semantics of instruments: their inherent properties and the roles they play within events. We then turn to analyses that emphasize the selectional restrictions imposed by predicates. The section concludes with a summary of previous claims that information structure plays a role in the distribution of instrument subjects.

2.1 Semantics of Instruments

2.1.1 Intermediary and Facilitating Instruments

There have been several attempts to use semantic criteria to divide instruments into two subclasses, with only one having the potential for subjecthood. Several researchers have proposed a distinction between *intermediary* and *facilitating* instruments, with intermediary instruments being the only potential subjects (Marantz 1984: 247, Levin 1993: 80, Brousseau 1998, Alexiadou and Schäfer 2006). The precise nature of the distinction, however, varies between accounts. One take on the distinction is that intermediary instruments generate their own force, so that they can be conceived of as acting independently; facilitating instruments, in contrast, are not inherently potent (Levin 1993: 80). This is closely related to Kamp and Rossdeutcher's (1994) distinction between *instrument causers*, which are inherently potent, and *pure instruments*, which are not. The following examples illustrate this take on the intermediary/facilitating distinction.

(11) The crane/*pitchfork loaded the truck. (*judgments*: Levin 1993: 80)

(12) Die Kamille/*Das Skalpell heilte den Patienten. (*judgments*: Kamp and Rossdeutscher 1994: 143)
'The chamomile/*scalpel cured the patient.'

Both cranes and chamomile have inherent potency – mechanical and chemical, respectively – that enables them to act upon other objects with some level of independence. For this reason, they are linked as subjects more readily than entities that lack inherent force, such as pitchforks and scalpels. Schlesinger’s (1989: 190) Naturalness Condition 1 contains a similar observation: “When the event is not instigated by a human agent, or when the agent is unknown or no longer on the scene, the instrument by means of which the action is performed or which is involved in the event may be naturally expressed as the subject.” The generalization is that an instrument’s degree of force-dynamic independence is positively correlated with the ease with which it can be realized as subject.

There is a broad consensus that potency plays a significant role in the acceptability of instrument-subject sentences. In addition to the research discussed above, Grimm (to appear) argues that potency is one of a handful of semantic factors relevant to subject selection. However, to claim that instruments must be potent in order to be potential subjects would be too strong, at least for English. Such a claim would rule out simple inanimate objects that lack chemical potency, and we have already seen that NPs denoting objects such as clubs (1b), keys (2b) pens (3b, 8b), and cloths (6b) can be acceptable subjects. Pitchforks and scalpels, functioning as instruments, may be coded as subjects, too, given an appropriate context: consider (13) and (14) in a courtroom scenario in which it is crucial to identify the specific instruments involved in an event.

(13) Is this the pitchfork that loaded the truck?

(14) Is this the scalpel that made the incision?⁹

⁹ The direct counterpart of (12), *Is this the scalpel that cured the patient?*, is odd for some speakers, presumably because of the relatively trivial role that a scalpel would typically play in curing a patient. However, there are attested examples in which ‘the scalpel’ refers metonymically to the event of undergoing surgery.

(i) It was indeed the injections of hormones and ‘the scalpel’ that cured me of my unease.

The second take on the intermediary/facilitating distinction is more promising. Marantz (1984: 247) conceives of intermediary instruments as intervening participants in an action chain that links the agent to the patient. For example, he paraphrases the action chain underlying (15a) as “Elmer does something to the key; the key does something to the cage.” Because the key is an intermediary instrument, it can be realized as subject, as in (15b). In contrast, facilitating instruments do not serve as intermediaries in an action chain, as shown by the fact that (16a) cannot be paraphrased as “Elmer did something to the magnifying glass; the magnifying glass did something to the book.” Accordingly, the magnifying glass is not an acceptable subject. I believe that this generalization is correct, and in Section 3 I will give my suggestions for refining and formalizing it.

- (15) a. Elmer unlocked the porcupine cage with a key.
- b. A key unlocked the porcupine cage.
- (16) a. Elmer examined the inscription with a magnifying glass.
- b. * The magnifying glass examined the inscription.

Brousseau’s (1998) analysis of the IS alternation in the West African language Fɔ̀ngbè incorporates both the ‘potency’ and ‘intermediary’ models of the distinction between intermediary and facilitating instruments. The basic pattern of the Fɔ̀ngbè alternation is much like English. Some instruments, such as *tú* ‘a rifle’ in (17), can be realized either as adjuncts (introduced by the marker *kpó(dó)* ... *kpó* ‘with’) or as subjects (marked by sentence-initial position), while others, such as *gáfù* ‘a fork’ in (18), can only be linked as adjuncts (Brousseau 1998: 89).

(blogs.news.com.au/heraldsun/...php/.../no_scalpel_could_cure_them/, accessed 11/20/2009)

- (17) a. Asíba hù Kòkú kpó tú kpó.
Asiba kill Koku with rifle with
'Asiba killed Koku with a rifle.'
- b. Tú éís hù Kòkú.
rifle DEM kill Koku
'This rifle killed Koku.'
- (18) a. Asíba d̀ù làn kpó gáfù kpó.
Asiba eat meat with fork with
'Asiba ate the meat with a fork.'
- b. * Gáfù éís d̀ù làn.
fork DEM eat meat
'This fork ate the meat.'

Following Marantz, Brousseau distinguishes intermediary instruments, such as the rifle in (17), which can be realized as subjects, from facilitating instruments, such as the fork in (18), which cannot. Intermediary instruments are conceived as part of a chain of action that accomplishes the change of state or location entailed by the verb. The verb *hù* 'kill' in (17) entails a change of state from life to death, in which the rifle participates. In contrast, the verb *d̀ù* entails consumption of an object, a change of state in which the fork does not directly participate.

Brousseau further divides the class of intermediary instruments according to the property of inherent potency. She argues that intermediary instruments that generate force independently can be freely realized as subjects, while controlled instruments require the presence of a special focus particle. For example, the scenes depicted in (17a) and (19a) can easily be paraphrased with the rifle as subject (17b), but not with the knife (19b). According to Brousseau (1998: 111), this is because the rifle is inherently potent and hence more independent of the controlling agent than the knife. Non-potent intermediary

instruments, such as the axe in (20), can only be linked as subjects in the presence of the emphatic particle *we*.

- (19) a. Asíbá hù Kòkú kpó jìvì kpó.
 Asiba kill Koku with knife with
 ‘Asiba killed Koku with a knife.’
 b. ?? jìvì élò hù Kòkú.
 knife DEM kill Koku
 ‘This knife killed Koku.’
- (20) Àsìyóví ó wè já àtín.
 axe DET EMP cut-into-pieces tree
 ‘This is the axe that cut a tree into pieces.’ *or*
 ‘It’s with the axe that they cut the tree into pieces.’¹⁰

My analysis of the English alternation connects with Brousseau’s work on Fɔ̀ngbè in several respects. I will adopt the idea that the class of potential instrument subjects are those that participate in a chain of action entailed by the verb (cf. Marantz 1984: 247). I will also provide an account of her observation that non-potent instruments require pragmatic support to be linked as subjects, a generalization that holds in English as well. We will return to the pragmatic component of Brousseau’s analysis in Section 2.3.

2.1.2 Semantic Roles Assigned to Instruments

In addition to attempting to delimit the class of instruments that are potential subjects, previous research has also aimed to identify the semantic features that instruments bear when they are realized as subjects. The central question underlying this program is the following: if (active) subjects are typically agents, then how do instruments satisfy the semantic constraints associated with subjecthood? Two main approaches have been

¹⁰ Brousseau offers two French glosses for example (20): “C’est la hache qui a débité un arbre” and “C’est avec la hache qu’on a débité un arbre.” She suggests that the second gloss is a more appropriate translation.

proposed. The first is *underspecification*, in which an instrument subject bears the same semantic role(s) as a canonical subject, such as *agent* (DeLancey 1984, 1991, Schlesinger 1989, Jackendoff 1990, Alexiadou and Schäfer 2006), *cause(τ)* (Brousseau 1998, Alexiadou and Schäfer 2006), *causer/initiator* (Ramchand 2003), or *effector* (Van Valin and Wilkins 1996). Bearing the underspecified semantic role(s) is taken to be a prerequisite for realization as subject. The second approach invokes a *causal chain* containing both the agent and the instrument, with the linking system specifying that any member of the causal chain is a potential subject (Croft 1991, 1994, 1998, Guilfoyle 2000, Wolff et al. 2009, cf. Marantz 1984: 247). The two approaches make distinct predictions with respect to the role of agents in instrument-subject sentences. In the causal chain approach, agents may be present in the semantic representation; in the underspecification approach, they cannot be.

Starting with the underspecification approach, a review of the literature reveals that researchers use terms like *agent* and *causer* in diverse and sometimes idiosyncratic ways. Accordingly, this discussion will be organized with respect to the semantic content of the roles posited, rather than the role labels. There are two types of underspecified semantic roles that have been associated with instrument subjects. One set of analyses holds that instrument subjects are causal, i.e. participants in an event that causes an object to change state or location (DeLancey 1984, 1991, Schlesinger 1989, Brousseau 1998, Ramchand 2003, Alexiadou and Schäfer 2006). One advantage of this proposal is that it offers a unified semantics for (active) subjects: sentient agents, natural forces, and instruments can be accommodated under a single umbrella. However, there is evidence that it does not fully capture the constraints on subject selection in English. First, some predicates with non-

causative semantics permit instrument subjects, e.g. surface contact verbs such as *sweep* and *hit*.¹¹

- (21) [Context: description of exhibit of specimens associated with famous people]
Twig from a broom that swept the room that Mao Tse Tung used as a student at Hunan University from 1917 to 1919.¹²
- (22) The bat that hit the shot was examined by the umpires and determined to be a composite material.¹³

Another problem with this proposal is common to all underspecification accounts: they require that the semantic representations of instrument-subject sentences lack agents. This is because the instrument is assigned the same, unique semantic role that would otherwise be assigned to the agent in a canonical sentence. However, some predicates that allow instrument subjects entail the presence of a volitional, controlling argument, such as *murder* and *write* (cf. Van Valin and Wilkins 1996, who argue that *murder* selects for a volitional agent).¹⁴ As the agent is a necessary component of the semantics of these predicates, the underspecification account does not go through. We will revisit the question of how to represent the agent in instrument-subject sentences in Section 3.

Alexiadou and Schäfer (2006) claim that the causal analysis captures the semantics of some, but not all, instrument subjects. They propose that all instrument subjects are

¹¹ According to Rappaport Hovav and Levin (1998), surface contact verbs are associated with a basic non-causative event schema (the *activity* schema), which can be augmented to include a result state (the *accomplishment* schema). Instrument subjects can appear in both event types, as shown by the fact that they are compatible with temporal *for* phrases (which select activities) and *in* phrases (which select accomplishments).

- (i) This is the broom that swept the floor for hours.
(ii) This is the broom that swept the floor in an hour.

¹² <http://legacy.lclark.edu/dept/gallery/objects/Artists.Specimens06.pdf>, accessed 9/12/2009

¹³ <http://www.funoncapecod.com/blog/?p=32>, accessed 9/12/2009

¹⁴ As the term “automatic writing” illustrates, it appears that certain speakers do allow for events of writing that lack sentient agents.

associated with one of two underspecified semantic roles: *causer* or *agent*. Causers are associated with a causing event, as in DeLancey's and Ramchand's proposals. Their definition of agency is rather obscure: an NP is an *agent* if "(a property of) the NP grounds the coming about of the event" (p. 46). Given this vague definition, it is difficult to identify the predictions that their account makes. However, they make the interesting suggestion that the constraint underlying the distribution of instrument subjects is sensitive to focus. The pragmatic component of their account will be discussed in Section 2.3.

The second type of underspecification account appeals to action, i.e. the transmission of force, rather than causation (Jackendoff 1990, Van Valin and Wilkins 1996). Jackendoff claims that instrument subjects, like agentive subjects, are represented as the first argument of an *action tier*, which encodes relations of physical affectedness (force transmission) that hold between objects. This is similar to Van Valin and Wilkins's role of *effector*, "the dynamic participant doing something in an event" (1996: 289). This proposal is empirically stronger than the causation approach, because it captures the fact that non-causative verbs entailing force transmission permit instrument subjects. In fact, in Section 3 I will argue that the class of predicates that allow them is delimited by an entailment of physical force transmission. However, because Jackendoff's and Van Valin and Wilkins's analyses involve underspecification, they still cannot account for the observation that some predicates that allow instrument subjects entail the presence of an agent in the event.

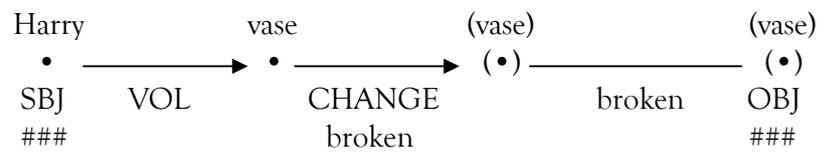
In contrast with the underspecification approach, accounts invoking a causal chain allow (but do not require) the agent to appear in the semantic representation of an instrument-subject sentence. Guilfoyle (2000) and Wolff et al. (2009) argue that the

causing argument of change-of-state and change-of-location verbs is represented as a chain that links the initiator to the patient through a series of causally connected subevents (cf. Marantz 1984: 247). In the case of instrument-subject sentences, the instrument intervenes between the initiating agent and the patient. They suggest that languages differ with respect to how subjects (or external arguments) are selected on the basis of event representations that include a causal chain. They propose that some languages, like Irish (Guilfoyle 2000) and Korean (Wolff et al. 2009), permit only the initiator of the causal chain to be the subject, while other languages, such as English, accept any causal participant. This is why, they claim, English allows instrument subjects while Irish and Korean do not.

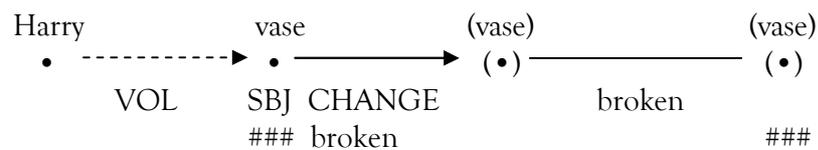
Neither Guilfoyle's nor Wolff et al.'s accounts provide details of how the causal chain approach could be formalized. Croft (1991, 1994, 1998) presents one possible approach. Recall from Chapter 2 that in Croft's model, semantic representations consist of causal chains in which force is transmitted from one participant to another. The verbal profile indicates the portion of the causal chain that is in focus in a particular sentence. The first participant of the profiled portion of the causal chain is linked as the subject, while the final profiled participant is the object. The following examples illustrate the basic components of the causal chain and the verbal profile (Croft 1994: 38).¹⁵ The verb *break* is associated with a causative profile, which includes the initiator of the event (23), and an inchoative profile, which excludes it (24). See further discussion in Chapter 2.

¹⁵ I constructed example (24) on the basis of (23), which appears in Croft (1994: 38). See Chapter 2, Section 3.1.1 for details about Croft's system of representation.

(23) Harry broke the vase.



(24) The vase broke.



Croft (1998: 46) suggests that the IS alternation can be represented in a parallel way. Given an event representation in which the agent acts on the instrument and the instrument on the patient, one verbal profile would highlight the entire causal chain, with the result that the agent is linked as subject, while the other would begin with the instrument. One clear advantage of this approach is that it allows for the agent of instrument-subject sentences to be present in the representation, albeit de-profiled. This elegantly captures two facets of the relationship between agents and instruments that are typically hard to reconcile: that some verbs that allow instrument subjects select for an agent and that, at the same time, the agent is de-emphasized or “offstage” (cf. Schlesinger’s (1989) Naturalness Condition 1).

Croft’s account, though, raises questions about the status of instruments in lexical representations. In order for a verb to have the “instrument subject” profile, which begins with an instrument, the instrument must be part of the causal chain lexically associated with the verb. However, some predicates that appear not to select for instruments – such as *break* in Croft’s examples above – nevertheless permit instrument subjects. My analysis of

the semantics of instrument-subject sentences attempts to preserve Croft’s account of the relationship between agents and instruments, while allowing for the fact that not all instruments are lexically selected.

2.2 Semantics of Predicates

Another strand in the literature places constraints on the distribution of instrument subjects within the lexical semantics of predicates. For instance, Reinhart (2002) proposes that the predicates that permit instrument subjects select for a causing entity that is underspecified for sentience. Her account predicts that (transitive) *break* and *open* should allow instrument subjects because their causing argument need not be sentient, whereas predicates like *write* should forbid instrument subjects because they entail the presence of a sentient agent. As we have already seen, this is too strong. Reinhart’s account faces additional empirical difficulties as well. She argues that the class of predicates that undergo the causative-inchoative alternation (for her, a process of de-causativization) are precisely those that select for an underspecified cause. She thus predicts that the class of predicates that allow instrument subjects and those that undergo decausativization should be coextensive. However, this prediction is not borne out for either English or Hebrew, as the following examples illustrate.¹⁶

- (25) a.(#) et adom katav et ha-mixtav.
 pen red write.3SGM.pst ACC DEF-letter
 ‘A red pen wrote the letter.’
 b. ha-mixtav nixtav (im et adom).
 DEF-letter write.psv.3SGM.pst (with pen red)
 ‘The letter was written (with a red pen).’
 c. * The letter wrote.

¹⁶ Thanks to Yael Fuerst for providing judgments and suggestions on the Hebrew data.

- (26) a. (#) *ala pats'a* et Nancy Kerrigan.
 club injure.3SGF.pst ACC Nancy Kerrigan
 ‘A club injured Nancy Kerrigan.’
- b. Nancy Kerrigan *niftse'a* (* *im ala*).
 Nancy Kerrigan injure.inch.3SGF.pst (with club)
 ‘Nancy Kerrigan “injured” (*with a club).’
- c. * Nancy Kerrigan injured.

As the (a) examples above illustrate, both *katav* ‘write’ and *pats’a* ‘injure’ permit instrument subjects with appropriate contextual support, like their English counterparts. However, they vary with respect to the availability of the causative-inchoative alternation. Both *katav* and its counterpart *write* lack inchoative forms, as illustrated by (25b-c).¹⁷ In contrast, *pats’a* differs from *injure* in that the former, but not the latter, can appear in the inchoative form. This demonstrates that the distribution of instrument subjects is not tied to the availability of the causative-inchoative alternation.

Grimm (to appear) situates the constraints on instrument subjects in the correspondence between the lexical entailments of predicates and the semantic properties of instruments. He argues that verbs entail or fail to entail a small set of event-based properties that constitute agency, loosely related to Dowty’s (1991) proto-agent entailments: *instigation*, if the event is brought about by one of its participants; *sentience*, if one participant is necessarily sentient; *volition*, if one participant acts volitionally; and *persistence*,

¹⁷ In Hebrew, the morphological pattern *niph'al*, illustrated by (25-26b), can receive either an inchoative or a passive interpretation, depending on the lexical properties of the verb (Doron 2003). I use Reinhart and Siloni’s (2003) diagnostic to determine the interpretation of a given *niph'al* form. Because passives, but not inchoatives, contain a controlling agent in the semantic representation, only the former permit instrument adjuncts. The diagnostic demonstrates that *nixtav* (25b) is passive and *niftse'a* (26b) is inchoative, and therefore that *pats’a* ‘injure’ undergoes decausativization but *katav* ‘write’ does not.

if the properties or very existence of a participant persist throughout the event.¹⁸ The properties are arranged in an *agency lattice*, which represents the possible space in which predicates may be situated with respect to the (non)-entailment of agency properties. In Grimm's model, constraints on subject selection emerge from a predicate's position on the lattice. For example, he argues that *murder* has all four entailments: instigation, sentence, volition, and persistence. Accordingly, he predicts that all possible subjects of *murder* should instantiate all four properties. As Grimm points out, this rules out the possibility of instrument subjects, which are necessarily non-sentient and non-volitional. Grimm considers this to be a strength of his analysis. However, as we have seen, some verbs that entail the presence of a volitional agent, including *murder*, do allow instrument subjects. Therefore, the constraints that Grimm proposes are too strong.

Dowty's (1991) theory itself does not explicitly address instrument subjects. However, it makes problematic predictions that are similar to those of Grimm's analysis. Dowty claims that the subject of a transitive sentence denotes the event participant that bears the highest number of proto-agent entailments: volitionality, sentience, causality, movement, and independent existence. Instruments are predicted to appear as subjects when they carry the most proto-agent entailments. In event types that entail the presence of a volitional agent, such as events of murdering and writing, the agent always has more proto-agent entailments than the instrument. Writers and murderers bear four entailments: volitionality, sentience, causality, and independent existence. In contrast, murder weapons and writing implements have at most three: causality, movement, and independent

¹⁸ Grimm proposes a distinction between *existential persistence*, reflecting whether the participant itself exists at the beginning and end of the event, and *qualitative persistence*, whether the participant's properties are constant throughout the event.

existence. Dowty's model thus predicts that agents will invariably be linked as subjects in events of writing and murdering, a prediction that, as we have seen, is not borne out.

Brousseau's (1998) account of the IS alternation in Fõngbè also has a lexical component, on top of the constraints on the semantics of the instrument that were discussed in Section 2.1.1. Recall her generalization that instruments that lack inherent potency can only be realized as subjects when there is pragmatic support in the form of an emphatic particle. She observes that this constraint does not apply to instruments that are lexically selected, which can be linked as subjects even without contextual support. In Section 3, I will show that in contrast with Fõngbè, lexically selected subjects do not have any special access to subject position in English.

2.3 The Effects of Pragmatics

Several researchers have previously observed that pragmatic constraints play a role in the acceptability of instrument-subject sentences. Two distinct pragmatic notions have been claimed to play a role: salience or "topicality" (Schlesinger 1989, Brousseau 1998) and contrast (DeLancey 1991, Alexiadou and Schäfer 2006). Starting with the former, both Schlesinger and Brousseau claim that instrument-subject sentences are more natural when the instrument is relatively salient in the discourse with respect to the agent. Schlesinger (1989: 191) expresses this through his Naturalness Condition 2: "to the extent that attention is drawn to the instrument by means of which an action is performed and away from the instigator of the action, the former will be naturally expressed as the sentence subject." Brousseau (1998: 113) claims that the instrument subject construction is a

topicalization device on par with the passive voice, which obtains only when the instrument is more prominent in the discourse than the agent. In Section 4, I demonstrate that an appeal to the salience or topicality of the instrument is not sufficient to capture the pragmatic constraints on instrument subjects in English. Instead, the instrument subject must be linked to the prior discourse through association with an activated proposition (e.g. Lambrecht 1994, Lambrecht and Michaelis 1998, Dryer 1996, Birner and Ward 1998, Birner et al. 2007).

Turning to analyses involving contrast, DeLancey (1991: 348) claims that sentences like *The key opened the lock* are only acceptable “outside of the popular folklore of linguists” if contrastive stress is involved, e.g. if the key is being contrasted with another possible instrument. Alexiadou and Schäfer (2006: 45) develop this observation further, demonstrating that contrastive stress on any element of the predication – subject, verb, or object – improves the acceptability of an instrument-subject sentence (27). They show that contrastive stress also makes instrument subjects possible in German (28) and suggest that the same is true in Dutch and Greek.

- (27) a. ?(?) The key opened the door. (*judgments*: Alexiadou and Schäfer 2006)
 b. The KEY opened the door.
 c. The key OPENED the door.
 d. The key opened THIS door.
- (28) a. ?(?) Der Schlüssel öffnete die Tür.
 ‘The key opened the door.’
 b. DIESER Schlüssel öffnete die Tür.
 ‘THIS key opened the door.’

Alexiadou and Schäfer also attempt to explain why contrastive focus would influence the instrument subject construction in this way (2006: 45): “We think that focus by precluding alternatives stresses the existence of a non-trivial relation between (a property

of) the subject and the event expressed by the whole VP.” They then suggest that focus allows the instrument to be classified as an *agent* – on their account, the NP “that grounds the coming about of the event” (p. 46), and therefore to be realized as subject. However, they do not explain how contrastive stress on the predicate would have such an effect on the interpretation of the subject.¹⁹

In Section 4, I provide an alternative explanation for the observation that instrument-subject sentences become more acceptable in the presence of contrastive stress. I argue that a certain class of instrument subjects is required to be linked to an activated (open) proposition. Contrastive stress serves to activate an open proposition, and thus renders instrument-subject sentences more felicitous.

3. Lexical Semantic Constraints on Instrument Subjects

This section presents a new account of the lexical semantic constraints on the distribution of instrument subjects in English. It incorporates constraints both on the lexical representations of predicates and on the semantics of instruments, specifically on the role that the instrument plays within an event. The section begins by delimiting the class of predicates that have the potential to license instrument subjects, namely polyvalent predicates that entail that an object is physically affected (Section 3.1). I show how this class of predicates can be defined within a hierarchical model of the lexicon, in which narrower classes of lexemes inherit the features of the broader lexeme class(es) that subsume them. We then turn to the way in which instruments are incorporated into event

¹⁹ Note that this claim runs counter to Lambrecht’s (1995) proposal concerning the relationship between information structure, semantic roles, and grammatical functions. He argues that focus on an agentive subject is typologically marked, and that for this reason, focus expressions are *less* likely to be interpreted as agentive.

structure representations headed by physical affectedness verbs (Section 3.2). I propose that the first argument of verbs of this class is an *action chain* consisting of force-dynamic subevents that ultimately affect an object (cf. Croft 1991, 1994, 1998). An instrument is included in the action chain only if it plays a direct role in affecting the object. Finally, Section 3.3 proposes that physical affectedness verbs inherit a linking constraint from the broader class of lexemes that entail action, specifically that the subject must be an actor. This creates indeterminacy with respect to subject selection for physical affectedness verbs, which permit multiple actors. When both an agent and an instrument appear within the action chain, the lexical semantic component of the linking system permits either to be realized as subject. As I will argue in Section 4, informational constraints on linking have the potential to resolve this indeterminacy.

One brief note before proceeding: this chapter will concentrate on episodic instrument-subject sentences, such as (29a), which report on a specific event. I will set aside generic or “characterizing” instrument-subject sentences, such as (29b), for future work. The contrast between the two sentences below suggests that characterizing instrument-subject sentences are sometimes acceptable in isolation where their episodic counterparts are not. The source of this distinction is an interesting topic for future study.

- (29) a. (#) This pen wrote the ransom note.
b. This pen writes well.

3.1 Delimiting the Class of Predicates

The relevant class of predicates is delimited as follows. In order to have the capacity to license an instrument subject in English, a predicate must entail that one or more

participants physically affect another participant. The predicates that satisfy this requirement fall into three broad classes: verbs entailing a caused change of state, verbs entailing a caused change of location, and verbs of surface contact. Table 1 lists subclasses of predicates that fall into each major category; the subclass labels come from Levin (1993).

Table 1: Classes of Verbs Entailing Caused Physical Affectedness

Superclass	Subclasses (adapted from Levin 1993)
Change of State	Verbs of Cutting (<i>cut</i>), Combining and Attaching (<i>mix</i>), Separating and Disassembling (<i>split</i>), Coloring (<i>paint</i>), Image Creation (<i>paint</i>), Creation and Transformation (<i>build</i>), Ingesting (<i>eat</i>), Bodily Injury (<i>injure</i>), Grooming and Bodily Care (<i>brush</i>), Killing (<i>kill</i>), Destroying (<i>destroy</i>), Change of State (<i>break</i>)
Change of Location	Verbs of Putting (<i>put</i>), Removing (<i>remove</i>), Sending and Carrying (<i>send</i>), Throwing (<i>throw</i>), Motion (<i>go</i>)
Surface Contact	Verbs of Exerting Force (<i>push</i>), Holding (<i>hold</i>), Contact by Impact (<i>hit</i>), Poking (<i>poke</i>), Contact (<i>touch</i>)

An important thing to note is that not all of the subclasses listed in Table 1 actually permit instrument subjects. For example, it is well known that verbs of ingestion (e.g. *eat*, *drink*) forbid them. This is because the distribution of instrument subjects is constrained not only by the lexical entailments of the predicate, but also the way in which instruments can be integrated into the event representation. My claim is that all of the verb classes in Table 1 license an action chain by virtue of their lexical entailments. However, they differ as to whether instruments can be construed as members of the action chain, which is a prerequisite for realization as subject. Verbs of ingestion, for instance, forbid instruments

from appearing in the action chain because of the facilitating (as opposed to intermediary) role that they play in events of ingestion. (The semantic constraints on participants in the action chain will be discussed in Section 3.2.) Table 1, then, illustrates the classes of verbs whose lexical semantic representations are compatible with the licensing of instrument subjects.

We now examine the constraints on lexical entailments in more depth. The required (complex) entailment is that one or more participants exert a force that physically affects another participant. This can be broken down into three distinct requirements: there must be two or more participants, there must be affectedness (force transmission) and force transmission must occur in the physical domain. We will look at each requirement in turn, holding the other two variables constant to show that each has an independent effect.

3.1.1 Requirement 1: Two or More Participants

The most basic requirement is that the predicate must entail the presence of two or more participants. To illustrate this, we can compare two sets of predicates entailing affectedness in the physical domain, one selecting for only one participant and the other for two or more participants. Here, I compare inchoative change-of-state and change-of-location verbs, which select only for an affected entity, to their causative counterparts, which additionally select for the event or participant that induces the change. The latter permit instrument subjects, in some cases with contextual support. The former do not, and indeed forbid instruments altogether. This is illustrated by the change-of-state verb *smash* in (30-31) and the change-of-location verb *slide* in (32-33).

- (30) a. John smashed the vase with the club.
b. (#) The club smashed the vase.
- (31) a. The vase smashed (*with the club).
b. * The club smashed. (*with the same interpretation as 31a*)
- (32) a. John slid the paper across the table with the pen.
b. (#) The pen slid the paper across the table.
- (33) a. The paper slid across the table (*with the pen).
b. *The pen slid across the table. (*with the same interpretation as 33a*)

On the present analysis, inchoative verbs forbid instrument subjects because they do not select for a subevent entailing force transmission that can be decomposed into an action chain. The ban on instrument adjuncts requires further explanation. Reinhart and Siloni (2003) argue that instrument adjuncts occur only when there is a controlling agent in the semantic representation. This explains why instrument subjects cannot co-occur with inchoative verbs, which do not lexically select for an agent. In contrast, passive forms of causative verbs, which are generally assumed to encode the agent as an implicit argument, do allow instrument adjuncts. We can see this by comparing (31a) and (33a) to parallel passive constructions (34).

- (34) a. The vase was smashed with the club.
b. The paper was slid across the table with a pen.

3.1.2 Requirement 2: Affectedness

We now look for evidence that affectedness constrains the distribution of instrument subjects, holding both number of participants and the domain in which the event occurs constant. We thus compare physical affectedness verbs to predicates that relate two (or more) participants in the physical domain but do not entail affectedness. Perception verbs such as *see* and *hear* fit the bill because they relate a perceiver to an object in the physical world without entailing that the object is affected in any way. These intuitions are

supported by Jackendoff's (1990) distributional test for affectedness, according to which *What X did to Y* is felicitous only if X affects Y. The three major classes of physical affectedness verbs that we have seen pass the test: change-of-state verbs (35a), change-of-location verbs (35b) and surface contact verbs (35c). Perceptual verbs, however, do not (35d).

- (35) a. What John did to the insect was dissect it.
- b. What John did to the insect was move it.
- c. What John did to the insect was touch it.
- d. # What John did to the insect was see it.

In contrast with physical affectedness verbs, exemplified here by the surface contact verb *touch*, perception verbs like *see* do not allow instrument subjects. They do, however, permit instrument adjuncts.

- (36) a. John touched the insect with the twig.
- b. (#) The twig touched the insect.
- (37) a. John saw the insect with the magnifying glass.
- b. * The magnifying glass saw the insect.

In my analysis, the ban on instrument subjects emerges from the fact that perception verbs do not contain action chains, because they do not entail affectedness.

3.1.3 Requirement 3: Physical Domain

The third essential component of the semantics of predicates that permit instrument subjects is that they denote events that take place in the physical domain. This means that the object must be physically affected; it must either undergo a physical change of state or location or be the recipient of force transmitted via surface contact. To test this claim, I compare verbs entailing physical affectedness to verbs entailing affectedness in another domain, such as the domain of information (cf. Pustejovsky 1995, 2006, Pustejovsky and

Bouillon 1995, Asher and Pustejovsky 2005). One member of the latter class is *revise*, which entails a caused change of state in an informational object. As the following examples demonstrate, *revise* does not license an instrument argument, either as an adjunct or as a subject (38a-b).

- (38) a. John revised the theory (*with a red pen).²⁰
b. * The red pen revised the theory.

In line with the previous two subsections, my claim is that *revise* disallows instrument subjects because only verbs entailing affectedness in the physical domain select for an action chain.

3.1.4 Representing Physical Affectedness

The rest of this subsection illustrates how the lexical entailments of physical affectedness verbs are represented, as well as how they fit into the lexical semantic system as a whole. The question guiding this discussion is how to account for the unique selectional properties of these predicates – specifically their ability to take instrument subjects – on the basis of their semantic representations. Generally speaking, there are three main approaches to accounting for selectional variability across predicates within a language. First, selectional variability can be grounded in the lexical entailments associated with specific participants, as in Dowty (1991), Davis (2001), and Grimm (to appear). Second, it can be grounded in the predicate’s event representation as a whole, rather than in the properties of individual participants, as in Jackendoff (1990), Brousseau (1998), and the

²⁰ Some speakers find the instrument adjunct in (38a) degraded but possible. To the extent that it is acceptable, it “coerces” a physical interpretation of *revise* in which physical changes are made to a document, cf. *edit*. Thus instruments are acceptable as subjects exactly when *revise* takes on the semantics of a physical affectedness verb.

work of Levin and Rappaport Hovav (1995, Rappaport Hovav and Levin 1998, Levin 2009). As a last resort, differences in selectional restrictions can simply be stipulated.

In the present analysis, the capacity to combine with an instrument subject emerges from the predicate's event representation as a whole, rather than the lexical entailments of specific participants (e.g. sentience, volitionality). One supporting piece of evidence is the existence of instrument-subject sentences in which the predicate entails the presence of a sentient agent (e.g. *murder*, *write*). As we have seen, accounts that ground selectional constraints in the lexical entailments of participants predict that agents should be invariably linked as subjects (Dowty 1991, Grimm to appear), counter what we see in the data. Because the present account delimits the class of predicates through constraints on event representations as a whole, it makes the right predictions: polyvalent physical affectedness verbs may combine with instrument subjects, regardless of whether they license an agent.

The event representations that I propose are based on those of Levin and Rappaport Hovav (1995, Rappaport Hovav and Levin 1998, Levin 2009), and are also influenced by the work of Croft (1991, 1994, 1998), Jackendoff (1990), and Pustejovsky (1995, 2006), among others. Following Levin and Rappaport Hovav, I assume that the event representation of a verb has two components: an *event schema* that relates the obligatory participants through a small vocabulary of primitive predicates and a *root* representing the idiosyncratic aspects of the verb's meaning that distinguish it from others with the same event schema. Levin (2009: 8) argues that roots are divided into a small set of ontological categories, including manner (*wipe*), result state (*dry*), and thing (*saddle*). To illustrate, the

verbs *jog*, *run*, *creak* and *whistle* share an event schema in which a single participant performs an action, but differ with respect to the root, which is incorporated into the event representation as a modifier expressing manner (39a). Thus, *jog* denotes an individual acting in a particular manner, specifically jogging (39b).

- (39) a. [X ACT <MANNER>]
 b. [X ACT <JOG>]

I also adopt Levin and Rappaport Hovav's inventory of primitive predicates that serve as the basis of event schemas. The predicates relevant to this analysis are ACT, BECOME, and CAUSE. ACT denotes emission of force; subevents in which two participants are related by ACT entail an asymmetric transmission of force from the first participant to the second (cf. Croft 1991, 1994, 1998, as well as the "action tier" of Jackendoff 1990). BECOME indicates a change in a participant. Finally, CAUSE is used to relate two subevents in which the first plays a critical role in bringing about the second. Similar, though not identical, inventories of primitive predicates can be found in the work of Jackendoff (1990) and Brousseau (1998).

With these preliminaries in place, I now introduce the event schemas associated with verbs of affectedness. The representations shown here are very similar to those proposed by Levin and Rappaport Hovav (Rappaport Hovav and Levin 1998, Levin 2009).²¹ The basic

²¹ Levin and Rappaport Hovav's version of (40b) differs from mine in that the affected participant is not represented as part of the causing subevent. This is shown in (i) below.

(i) [X ACT [CAUSE [BECOME [Y <STATE>]]]]

Because I will be claiming that causative affectedness verbs inherit the semantics of their non-causative counterparts, it is crucial for my purposes that the action subevent is represented identically in both cases.

In addition, Rappaport Hovav and Levin (1998) distinguish between core participants of the event schema and participants selected by the root, the affected argument of (40a) being an example of the latter. This distinction is not crucial to my analysis, so I set it aside.

structure common to all verbs of this class is represented in (40a), which represents one participant acting on another. Causative affectedness verbs have an additional meaning component, which is represented in (40b): the subevent entailing action is construed as causing a change in the affected participant. (40b) is the event schema associated with causative verbs that entail a change of state or location, such as *kill* and *throw*. Surface contact verbs such as *hit* have the simpler event schema illustrated in (40a), without a caused change of state or location.²²

- (40) a. [X ACT Y]
 b. [X ACT Y [CAUSE [BECOME [Y <STATE>]]]]

We now turn to the issue of how to distinguish physical affectedness verbs from the larger class of predicates entailing affectedness. My proposal is that the roots of predicates specify ontological domain in addition to ontological category.²³ For example, the event representation of *wipe* not only indicates that its root functions as a manner modifying the event, but also that it falls within the physical domain. Thus (physical) surface contact verbs and causative verbs entailing physical change can be represented in the following way, where P indicates the physical domain.

- (41) a. [X ACT Y <MANNER: P>]
 b. [X ACT Y [CAUSE [BECOME [Y <STATE: P>]]]]

²² Rappaport Hovav and Levin (1998) argue that the event schemas of surface contact verbs like *hit* can be augmented to incorporate a result state or location, e.g. *He hit the ball across the room*.

²³ Verbs that have the capacity to encode physical events as well as abstract events, such as *grasp* (i-ii), have multiple domain specifications.

- (i) John grasped the pencil.
 (ii) John grasped the idea.

Note that this is a distinct phenomenon from verbs like *edit*, which encode events that transpire in multiple domains simultaneously.

The way that I am using ontological domain information in event representations is significantly different from the proposals of Pustejovsky and colleagues (Pustejovsky 1995, 2006, Pustejovsky and Bouillon 1995, Asher and Pustejovsky 2005). They demonstrate that predicates may place restrictions on the ontological categories of their arguments. This can be seen through the phenomenon of *coercion*, in which the type lexically associated with an argument shifts to satisfy the restrictions placed by a predicate.²⁴ Consider how the complements of *read* are interpreted in the following examples.

- (42) a. John read the book.
 b. John read the wall.
 c. John read the theory.

Pustejovsky and colleagues propose that *read* selects for an object that is typed as belonging to both the informational and physical domains. The object *book* in (42a) satisfies this constraint, as books are physical objects that bear informational content. In contrast, walls are typically conceptualized as purely physical objects, and thus (42b) must be coerced to an interpretation in which the wall carries information. Likewise, theories are typically informational objects, but (42c) implies that the theory has a physical manifestation.

In my analysis, the place of ontological domain information extends beyond selectional restrictions to entire event representations. This is crucial because the class of predicates that permit instrument subjects are those that entail affectedness in the physical domain, rather than those that select for a complement that is a physical object. For example, the verb *read*, which selects for a complement with a physical manifestation but does not entail that it is affected, forbids instrument subjects.

²⁴ According to Asher and Pustejovsky (2005: 8), the ontological type conventionally associated with an object depends on “commonsense metaphysical intuitions,” but can be shifted by discourse context as well as by selectional restrictions.

3.1.5 The Lexicon as a Multiple Inheritance Hierarchy

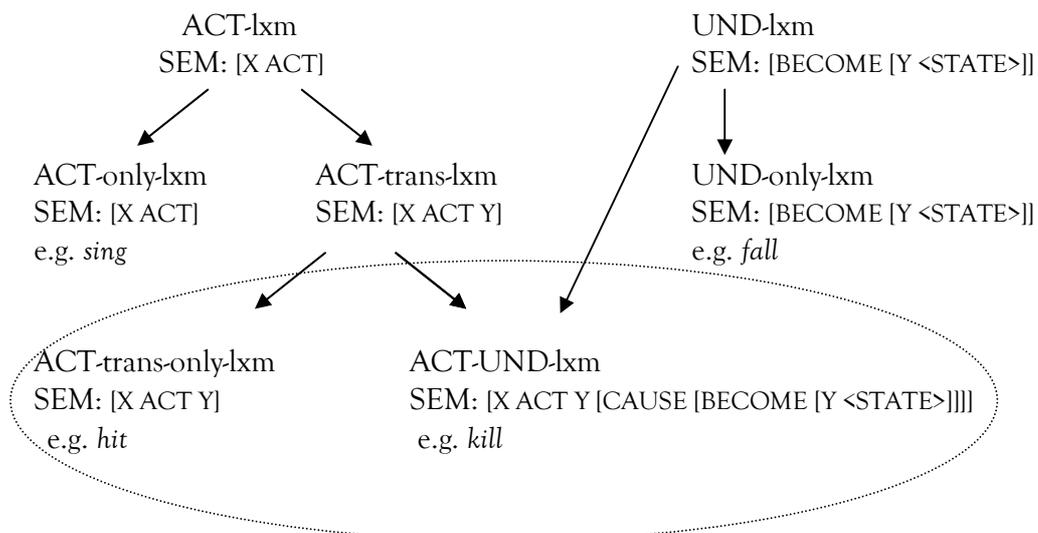
This subsection situates the proposed representations of physical affectedness verbs in a model of how the lexicon is organized. I assume that the basic structure of the lexicon is hierarchical, meaning that the feature specifications of relatively coarse-grained classes of lexemes are inherited by the relatively fine-grained lexeme classes that they subsume. Lexeme classes may inherit features from more than one superclass, a phenomenon called *multiple inheritance*. This model of the lexicon was introduced in Chapter 2. Here, I will focus on how the model can capture constraints on the distribution of instrument subjects, as well as the more general phenomenon of selectional variability across predicates. Davis and Koenig have proposed a finely articulated account of English argument realization situated within a hierarchical model of the lexicon (Davis and Koenig 2000, Davis 2001, Koenig and Davis 2003). My approach differs from theirs in one crucial respect: Davis and Koenig take patterns of argument linking to be grounded in the lexical entailments associated with participants, while I have argued for an approach grounded in event structure representations.

I propose that the lexical entries of verbs are organized along three dimensions: the event schema, the ontological category associated with the root (e.g. manner (*wipe*), thing (to *saddle*)), and the ontological domain associated with the root (e.g. physical (*misplace*), informational (*understand*)). Verbs inherit components of their event representations from all three dimensions, resulting in a relatively fine-grained classification of verbs. Each dimension has its own internal structure, with two dimensions – event schema and root ontological domain – exhibiting multiple inheritance. To illustrate how the system works, I

will start with an overview of the internal structure of each dimension, and then turn to how the dimensions are integrated via multiple inheritance.

Figure 1 illustrates a portion of the hierarchy of lexemes, organized along the dimension of event schema. At the top of the hierarchy are two coarse-grained lexical classes, those entailing action (ACT-lxm) and those entailing change in an entity (UND-lxm, abbreviating *undergoer*). The partial representations shown here contain only semantic information, but full lexical representations also bear feature specifications for phonology, morphology, argument structure, syntax, and pragmatics (see Chapter 2 for more discussion).²⁵

Figure 1. Partial Hierarchy of Lexemes, Organized by Event Schema



It is important to note that Figure 1 represents generalizations over classes of actual lexemes, of type *lxm*. The source of these generalizations are lexeme class constructions (type *lex*), which encode restrictions on the feature structures of classes of lexemes. These

²⁵ As discussed in Chapter 2, the event structure approach adopted here differs from Sag (2010), who uses frames to represent semantic structure.

are discussed in Chapter 2. For example, lexemes belonging to the class *ACT-lxm* are shaped by the construction in (43).

$$(43) \quad \text{ACT-lex} \rightarrow \left(\begin{array}{l} \text{ACT-lxm} \\ \dots \\ \text{SEM: [X ACT]} \\ \dots \end{array} \right)$$

The arrows in Figure 1 represent inheritance. This relation compels a lexeme class to carry all the features of the class(es) that subsume it. The daughter class may contain additional features above and beyond its inheritance, so that it represents a more specific instance of the class(es) it inherits from. Consider the subclasses of *ACT-lxm*: *ACT-only-lxm*, which contains all and only the features of the mother class, and *ACT-trans-lxm*, which builds on the inherited semantics by adding another participant. Similarly, *ACT-trans-lxm* has two subclasses: *ACT-trans-only-lxm* carries its exact semantics, while *ACT-UND-lxm* has an augmented event representation including a caused result state. The complex event representation associated with *ACT-UND-lxm* comes about through multiple inheritance, in which the action event inherited from *ACT-trans-lxm* and the change-of-state event inherited from *UND-lxm* are combined via the relation *CAUSE*.

Returning to the IS alternation, I have argued that only polyvalent physical affectedness verbs have the capacity to combine with instrument subjects. The hierarchy of event schemas proposed here allows us to distinguish this class of predicates from those that do not permit instrument subjects. The predicates that do license instrument subjects are members of the class *ACT-trans-lxm*, including its subclasses of surface contact verbs (*ACT-trans-only-lxm*, e.g. *hit*) and verbs entailing change of state and location (*ACT-UND-lxm*,

e.g. *kill*). These subclasses are indicated by a circle in Figure 1. All predicates that fall outside this class forbid instrument subjects.

The class of predicates permitting instrument subjects is further constrained by the semantics of the root. This too can be elegantly captured in a multiple inheritance model. I have proposed that the semantics of the root is organized along two dimensions: ontological category and ontological domain. Ontological category seems not to play a role in the IS alternation, so I set it aside here. Instead, I focus on ontological domain, which is crucial to delimiting the class of predicates that permit instrument subjects.

Figure 2 illustrates a fragment of the lexicon, organized both by event schema and the ontological domain of the root. Focusing first on the top right corner of the hierarchy, we see two main classes of lexemes organized by domain: P-lxm, whose roots entail activity in the physical domain, and I-lxm, whose semantics are situated in the informational domain. They are the mothers of three subclasses: P-only-lxm, I-only-lxm, and PI-lxm, which denote events situated simultaneously in both the physical and informational domains. Ontological domain is represented as a superscript to the event schema, which in the top right portion of the diagram is unspecified.

Figure 2. Partial Hierarchy of Lexemes, Organized by Event Schema and Ontological Domain

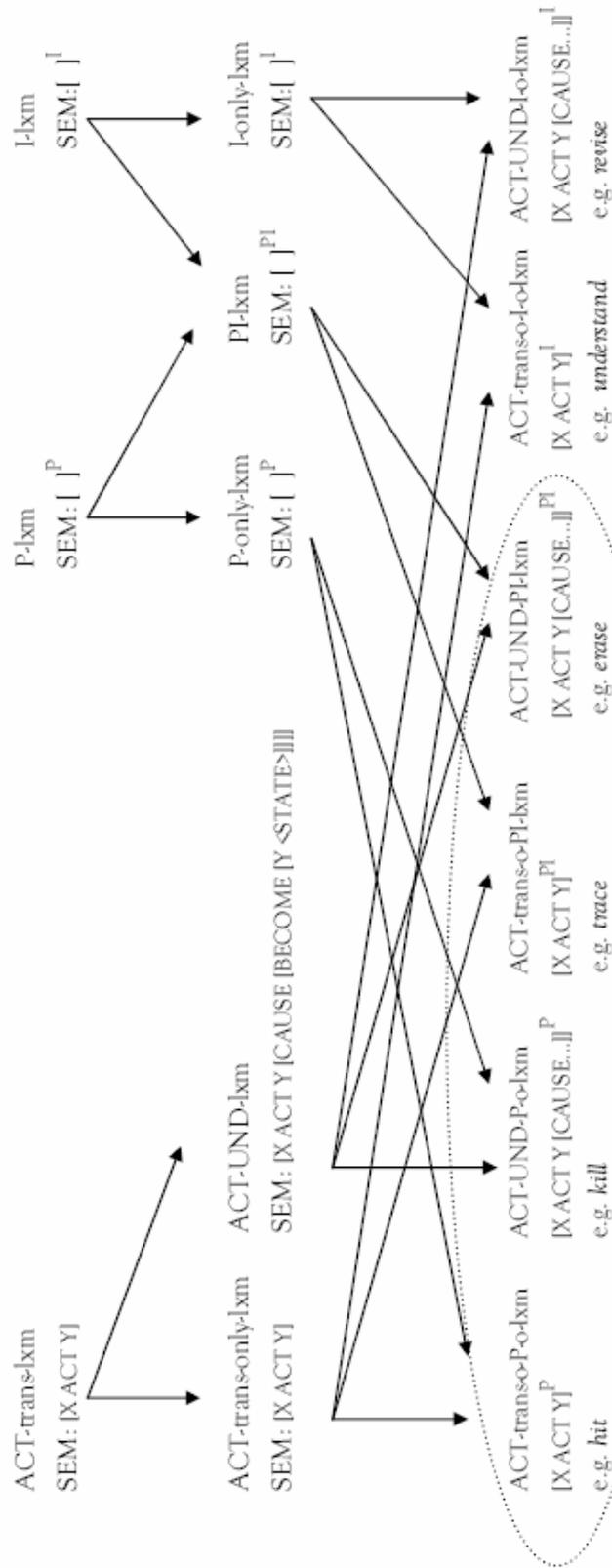


Figure 2 as a whole illustrates the classes of lexemes formed by multiple inheritance from the ontological domain hierarchy and a small portion of the event schema hierarchy – specifically the daughters of ACT-trans-lxm, the event schema compatible with instrument subjects. Crossing the two dimensions results in six subclasses of verbs. ACT-trans-only-P-only-lxm entails simple affectedness in the physical domain, i.e. surface contact verbs like *hit*. ACT-UND-P-only-lxm entails a caused change of state in the physical domain, e.g. *kill*. Similarly, ACT-trans-only-PI-lxm represents affectedness without change in the physical and informational domains; a likely member of this class is *trace*. Its causative counterpart, ACT-UND-PI-lxm, includes verbs like *erase*, which entail physical and informational change in an object. Finally, ACT-trans-only-I-only-lxm entails “surface contact” in the informational domain, e.g. *understand*, while ACT-UND-I-only-lxm requires change in an informational object, e.g. *revise*.

As I have argued, predicates that license instrument subjects entail affectedness in the physical domain. In the multiple inheritance model, these are the predicates that inherit semantic features both from ACT-trans-lxm and P-lxm, indicated by a circle in Figure 2. Note that this includes verbs entailing affectedness in both in the physical and informational domains, i.e. ACT-trans-only-PI-lxm (*trace*) and ACT-UND-PI-lxm (*erase*). The following naturally-occurring examples demonstrate that the PI classes permit instrument subjects.

- (44) a. The principal claim to novelty concerned the parallel-motion type amplifying mechanism, which ensured that *the pencil that traced the diagram* moved in a straight line.²⁶

²⁶ www.archivingindustry.com/Indicator/internalspring.htm, accessed 11/19/2009

- b. If a single word is unused or unnecessary to your poem, delete it, delete the application that deleted it, destroy *the pencil that erased it ...*²⁷

We have now seen that the class of predicates that allow instrument subjects can be delimited in a multiple inheritance model of the lexicon. In the next section, we turn to the special semantics associated with this class.

3.2 Action Chains and the Semantics of Instrument Actors

I propose that a predicate's ability to license an instrument subject is rooted in the presence of an *action chain* in its event representation. An action chain is a sequence of subevents involving force transmission that ultimately results in the affectedness of an object. Similar constructs play central roles in the theories of event representation proposed by Jackendoff (1990) and Croft (1991, 1994, 1998). In my analysis the contribution of action chains is relatively circumscribed. Specifically, I propose that they are unique to verbs entailing affectedness. This appears to be a consequence of how events are conceptualized. In events that do not involve affectedness there is no "result" that can be construed as the work of multiple participants.

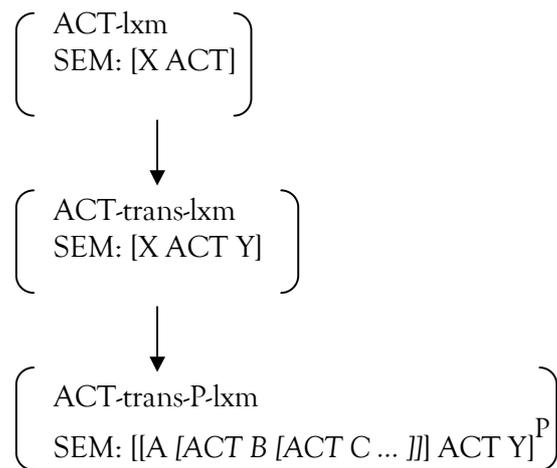
This subsection begins with a proposal for the representation of the action chain, then turns to how it constrains the semantics of instrument actors. Finally, it addresses issues of selection, illustrating how instrument subjects surface when the predicate selects for an agent, or when it does *not* select for an instrument.

²⁷ www.500vs365.blogspot.com/2009/08/improve-your-poetry.html, accessed 11/18/2009

3.2.1 Representing the Action Chain

In the present account, action chains are unique to the semantics of polyvalent physical affectedness verbs: the class ACT-trans-P-lxm. Action chains are incorporated into the event representations that these verbs inherit from coarser-grained verb classes. This is illustrated in Figure 3. As we have already seen, ACT-trans-lxm inherits the event representation of ACT-lxm and augments it with an argument representing the affected entity. The semantics of ACT-trans-lxm is in turn passed down to ACT-trans-P-lxm, which builds up the event structure further by decomposing the first argument into an action chain.

Figure 3. ACT-trans-P-lxm: Inheritance of Semantics



Because the action chain as a whole is the first argument of ACT, it is conceptualized as acting upon the affected entity. The chain itself consists of a sequence of subevents that link participants via the predicate ACT, denoting the asymmetric transmission of force. All elements of the action chain except the first participant A appear in italics, indicating that they are optional; the ellipses following the third participant C indicate that the chain may continue. Thus, an action chain may be as short as a single participant, and in principle it

may be indefinitely long. However, in practice most action chains will contain at most two participants due to the semantic constraints placed on actors, to which we turn now.

3.2.2 Constraints on Actors

In order for an event participant to appear in the action chain, it must play a particular role in how the event unfolds: it must transmit force that directly contributes to affecting an object in the way specified by the verb's event representation. This excludes two classes of instruments from the action chain: instruments that do not transmit force, and instruments that transmit force that does not contribute to the entailed change in the object. The distinction proposed here roughly corresponds to some takes on the distinction between intermediary and facilitating instruments (e.g. Marantz 1984, Brousseau 1998). However, the present analysis grounds the distinction between classes of instruments in the properties of event representations, and therefore makes clearer predictions.

First, instrument actors must transmit force. Take for example the verb *copy*, which on one interpretation entails a physical and informational change in an object, as in *John copied the physics formula*. Example (45a) illustrates two types of instruments that participate in events of copying: instruments like *a red pen*, which transmit force that induces physical and informational change, and those that do not transmit force at all, such as *a magnifying glass*.²⁸ Both types can be realized as adjuncts (45a), but only the former can be linked as subjects (45b), because presence in the action chain is a prerequisite for realization as subject.

²⁸ Magnifying glasses do not transmit force in events of copying. They may, however, transmit force in other event types, such as events of burning: *John burned the insect with the magnifying glass*.

- (45) a. John copied the formula with a red pen/a magnifying glass.
b. This is the red pen/* the magnifying glass that copied the formula.

Second, the force that an instrument actor transmits must directly contribute to the entailed event of affectedness. For example, verbs of ingestion such as *eat* and *drink* entail that an object is consumed. Accordingly, any member of an action chain associated with these verbs must exert force that directly contributes to the event of consumption. Instruments such as *the straw* in (46) do not contribute to the entailed change of state, but rather to a non-entailed change of location in which the liquid moves into the consumer's mouth. Because instruments of this sort are not part of the action chain, they cannot be linked as subjects.

- (46) a. John drank the Thai iced tea with the straw.
b. * The straw drank the Thai iced tea.

Why is it that verbs like *eat* and *drink* do not permit instrument actors at all? I propose that this is a consequence of how events of ingestion are conceptualized – namely, that it is difficult to conceive of an instrument that could intervene between the consumer and the consumed in the event of consumption. The present analysis differs from previous accounts (e.g. Reinhart 2002, Grimm to appear) in that it ties the non-occurrence of ingestion verbs with instrument subjects to global properties of events of consumption, rather than to selectional constraints placed on subjects.

3.2.3 Instruments, Agents, and Selection

This subsection takes a closer look at issues of selection, an area in which I will argue that the present account makes significant improvements over previous analyses. We start with two related observations. First, as we have already seen, some verbs that combine with

instrument subjects entail the presence of a sentient agent, e.g. *murder* and *write*. Others, such as *hit* and *kill*, do not. Likewise, some verbs with the capacity to license instrument subjects lexically select for a particular type of instrument. One example is *type*, which requires a typewriter or computer; in (47a), the presence of an instrument of this sort is implicitly understood. In contrast, verbs like *destroy* do not entail the presence of an instrument, and thus sentences like (47b) can be interpreted without an implicit instrument.

- (47) a. John typed the memo
b. John destroyed the memo.

These observations are problematic for several previous accounts, in which the possibility of licensing an instrument subject hinges either on the absence of a selected agent or on the presence of a selected instrument. As Section 2 pointed out, the latter is a problem for Croft's (1998) analysis, in which instrument subjects emerge from lexically represented verbal profiles that begin with the instrument. On the flip side, many previous accounts build in the assumption that there is no agent in the semantic representation of instrument-subject sentences (Schlesinger 1989, Jackendoff 1990, DeLancey 1991, Van Valin and Wilkins 1996, Reinhart 2002, Alexiadou and Schäfer 2006, Grimm to appear).

In contrast, the action-chain approach advocated here allows a predicate with an appropriate event representation to license an instrument subject regardless of whether it selects for an instrument or an agent. As I will argue in section 3.3, the lexical semantic system allows any participant that appears in the action chain to be realized as subject, whether or not it is semantically selected. Verbs like *type* (48), which select for an instrument, and *murder*, which select for an agent (49), simply come with restrictions on the

forecast, I will argue that extra contextual support is needed to link a non-initial member of the action chain as a subject. Thus, it requires more pragmatic support to realize an instrument as subject when it follows an agent in the action chain than when it is the initiator of the event.

Another question that my analysis raises is why the semantically represented agents of instrument-subject sentences are not available to act as controllers. The following example illustrates the problem. The well-known contrast between (51a) and (51b) indicates that the implicit agent of a passive construction is available to serve as the controller of a rationale clause; however, implicit agent control is generally impossible in sentences headed by inchoative verbs (Manzini 1983). This has led some researchers to claim that the semantic representation of passive sentences includes an agent, while inchoative sentences do not (e.g. Mauner and Tanenhaus 1995). Example (51c) demonstrates that instrument-subject sentences pattern with inchoatives in disallowing agent control of rationale clauses.

- (51) a. The ship was sunk in order to collect the insurance money.
b. # The ship sank in order to collect the insurance money.
c. # This is the bullet that killed John in order to collect the insurance money.

It is important to explain why agents are unable to control rationale clauses if they are indeed part of the semantics of instrument-subject sentences like (51c). One possibility is as follows. I assume that the default controller of a rationale clause is the “deep” subject - in the present theoretical framework, the first NP on the ARG-ST list. In sentences with instrument subjects, the instrument is the first member of the ARG-ST list. As such, the instrument serves as the controller of rationale clauses, with semantically anomalous

results. The agent appears within the action chain, but is not linked to the ARG-ST list, and therefore cannot serve as a controller.

3.3 Linking Actors as Subjects

The previous subsections illustrated how the class of predicates that permit instrument subjects can be delimited by virtue of a special feature of their semantics, the presence of an action chain. This subsection provides an account of how the lexical semantic system maps onto argument structure. I propose that the class of polyvalent physical affectedness verbs, ACT-trans-P-lxm, inherits a simple linking constraint common to all verbs that entail action: the first member of the ARG-ST list is the first argument of ACT, which typically results in an actor being realized as subject (cf. Jackendoff 1990: 268-269, Davis 2001: 175).

This is illustrated by Figure 4.

Figure 4. ACT-trans-P-lxm: Inheritance of Argument Linking

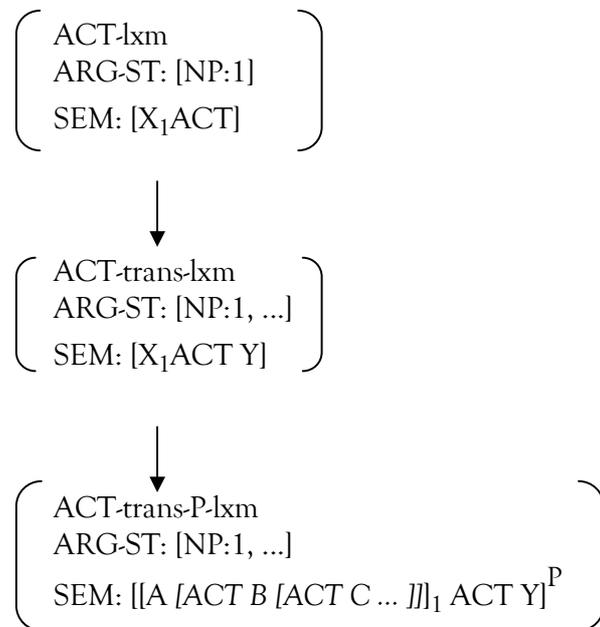


Figure 4 illustrates that the linking between event structure and argument structure, indicated by co-indexation, is passed down unaltered from class to class. In all classes, the first argument of the ARG-ST list is co-indexed with the first argument of ACT in the semantic representation. However, the semantics of *ACT-trans-P-lxm* causes it to have a distinct subject selection pattern. Because the first semantic argument of *ACT-trans-P-lxm* is an action chain, any element of the chain can be linked as subject. This results in indeterminacy when both an agent and an instrument appear in the action chain; in this case, both are candidates for realization as subject. English verbs classified as *ACT-trans-P-lxm* are thus associated with the lexeme class construction below.

English *ACT-trans-P-lex* Lexeme Class Construction

$$(52) \quad \textit{ACT-trans-P-lex} \rightarrow \left(\begin{array}{l} \textit{ACT-trans-P-lxm} \\ \dots \\ \text{SEM: } [[\text{A } [\text{ACT } B \text{ } [\text{ACT } C \text{ } \dots]]]_1 \text{ ACT } Y]^P \\ \text{ARG-ST: } [\text{NP:1, } \dots] \\ \dots \end{array} \right)$$

This approach also allows us to account for languages in which (active) subjects of polyvalent physical affectedness verbs must correspond to the initiators of the action chain, such as Irish (Guilfoyle 2000) and Korean (Wolff et al. 2009). In these languages, as in English, *ACT-trans-P* lexemes inherit the linking constraint originating with *ACT-lxm*, that subjects must be actors. However, they differ from English in further specifying argument linking: the first ARG-ST member must be linked to the initial actor. This is shown below.

Korean, Irish *ACT-trans-Plex* Lexeme Class Construction

$$(53) \text{ ACT-trans-Plex} \rightarrow \left(\begin{array}{l} \text{ACT-trans-Plexm} \\ \dots \\ \text{SEM: } [[A_1 [\text{ACT B} [\text{ACT C} \dots]]] \text{ACT Y}]^P \\ \text{ARG-ST: } [\text{NP:1}, \dots] \\ \dots \end{array} \right)$$

4. Pragmatic Constraints on Instrument Subjects

Several researchers have previously suggested that pragmatic considerations play a role in the acceptability of instrument-subject sentences (Schlesinger 1989, DeLancey 1991, Brousseau 1998, Alexiadou and Schäfer 2006). This section provides an in-depth exploration of these pragmatic effects and proposes an analysis that captures them. We start with the observation, noted by DeLancey (1991: 348), that most textbook cases of instrument-subject sentences are pragmatically odd in the absence of supporting context. Consider the two versions of the dialogue below.

- (54) Mother: What happened today?
 Babysitter: Henry broke the crystal vase with a baseball bat!
- (55) Mother: What happened today?
 Babysitter: # A/The baseball bat broke the vase!

Suppose a mother, upon returning home from work, asks the babysitter what happened during the day. The babysitter's response in (54), which links the agent of the event as subject and the instrument as an adjunct, is perfectly acceptable. In contrast, it is infelicitous for the babysitter to report the same event using an instrument-subject sentence, as shown by (55). The contrast stems from the fact that the babysitter's responses

in (54-55) function as *event-reporting* sentences: they introduce an event as a whole, with no part of it being previously activated in the discourse (Lambrecht 1994: 124).²⁹ Instrument-subject sentences like the response in (55) are odd in event-reporting contexts because, as I will argue, they *must* be associated with an activated proposition in order to be acceptable in context. This constraint applies only to instruments that are non-initiating participants in the action chain; that is, instrument that are controlled by agents.

This section begins by laying down the evidence that this class of instrument subjects must be associated with an activated proposition (Section 4.1). I then demonstrate how this constraint can be formalized through a derivational construction (Section 4.2). This account predicts that the effects of pragmatic constraints should interact with the semantics of the action chain; Section 4.3 shows that this prediction is borne out.

4.1 Instrument Subjects and Activated Propositions

Before presenting the evidence that instrument subjects denoting non-initiating actors are associated with activated propositions, I will first demonstrate that the notion of topicality is not sufficient to capture the pragmatic constraints on their distribution, contra Brousseau (1998: 113). Consider the example below. As we have already seen, (56a) is odd in isolation. Establishing the baseball bat as a discourse topic, as in (56b), has little, if any, effect on the acceptability of the instrument subject. However, when the instrument receives contrastive focus, as in (56c), the sentence is much improved. This demonstrates

²⁹ Lambrecht (1994: 124) notes that only intransitive sentences can be formally marked as event-reporting (i.e. *thetic*) in English. However, the sentences in (54-55) satisfy the informational constraints on event-reporting sentences.

that topicality is not at the root of the relationship between information structure and the acceptability of instrument-subject sentences.

- (56) a. # The baseball bat broke the crystal vase.
b. My baseball bat is silver and about two feet long. My dad bought it for me for my birthday. # It broke the crystal vase yesterday.
c. THIS baseball bat broke the crystal vase, not THAT one.

There are three strands of evidence indicating that instrument subjects must be linked to a proposition that is activated in the discourse. First, instrument subjects improve in acceptability when they appear in grammatical constructions that activate an open proposition, such as relative clauses, clefts, and bare (prosodic) focus. Second, they are relatively acceptable as answers to questions, where they are associated with an open proposition. Finally, instrument-subject sentences are felicitous in conversational contexts that activate relevant propositions by virtue of the interlocutors' goals (e.g., the implicit questions they are trying to answer), rather than through some element of the linguistic context.

4.1.1 Pragmatically Marked Constructions

One way in which the pragmatic constraints on instrument subjects can be satisfied is through non-canonical constructions that are conventionally associated with open propositions, such as relative clauses (Lambrecht 1994: 51-56) and *it*-clefts (Delin 1995, Dryer 1996, Lambrecht 2001). The following examples show that instruments that are odd as the subjects of syntactically simple sentences (at least in isolation) significantly improve in the pragmatically marked constructions.

- (57) a. (#) This bullet killed Joyce Alexander.

- b. The revolver which the police took from the appellant was loaded, with the exception of one chamber from which a shot had been discharged, and it was this bullet that killed Joyce Alexander.
- (58) a. (#) The pen wrote love letters to my mother, signed all my report cards, signed the tuition checks for college and everything else.
- b. This is the pen that wrote love letters to my mother, signed all my report cards, signed the tuition checks for college and everything else.

It-clefts activate an open proposition consisting of the content of the subordinate clause combined with the variable that is filled in by the clefted constituent. In (57b), the open proposition is *X killed Joyce Alexander*. The instrument subject resolves the variable *X*. Similarly, the open proposition associated with the relative clause in (58b) is *X wrote love letters to my mother, signed all my report cards, signed the tuition checks for college and everything else*.

In the *it*-cleft and relative clause sentences above, the instrument subject serves as the focus, in the sense of Lambrecht (1994): the portion of a pragmatically structured proposition that is asserted but not pragmatically presupposed (i.e. part of the open proposition). However, an instrument subject need not be a focus in order to be acceptable; it is sufficient for the instrument to be linked to *any* element of the open proposition. This is observed by Alexiadou and Schäfer (2006: 45), who demonstrate that bare focus on any element of an instrument-subject sentence boosts acceptability. They illustrate this through bare focus on the subject (59b), verb (59c), and object (59d). I add two related observations: that instrument-subject sentences improve both when they contain contrastive topics, which resolve open propositions with two or more variables,

one of which is topical (59e; see discussion in Chapter 1),³⁰ and under *verum focus* (59f), where focus is on the truth or falsity of a fully specified proposition.

- (59) a. The key opened the door.
b. THIS key opened the door.
c. The key OPENED the door.
d. The key opened THIS door.
e. THIS key opened THIS door; THAT key opened THAT door.
f. The key DIDN'T open the door.

Except in the case of *verum focus*, I assume that bare focus is associated with an open proposition consisting of the non-focused (background) content combined with the variable to be filled in by the focused element. The open propositions associated with (59b-e) are, respectively: *X opened the door*, *This key did X to the door*, *The key opened X*, and *X opened Y*. The instrument subject may be (part of) the focus, as in (59b) and (59e), but it may also be linked to the background, as in (59c) and (59d). The instrument is also linked to the background in *verum focus*, which applies to a full proposition.

In this discussion, I have been careful to claim that *it*-clefts, relative clauses, and bare focus activate (open) propositions. In doing so, I have set aside the question of whether these propositions need to be *previously* activated in the discourse, as well as whether they must be associated with existential presuppositions (e.g. whether the open proposition *X opened the door* implies a belief that something opened the door). This is because previous research has demonstrated that the answers to these questions vary across constructions. For example, Prince (1978) demonstrates that *wh*-clefts require a previously activated open proposition, while *it*-clefts do not. Similarly, Dryer (1996) argues that *it*-clefts are associated with an existential presupposition, while bare focus is not. Whatever the ways in which

³⁰ Lambrecht and Michaelis (1998) provide an alternative account of double-accent sentences such as (59e), in which the first accent serves to ratify a topic. See further discussion in Chapter 2.

these constructions differ, they share the property of (re)-activating a proposition when they are uttered.

4.1.2 Overt Questions

The second way in which an instrument subject can be linked to an activated proposition is through an overt question, which introduces an open proposition to the discourse. In (60-61) below, A's questions introduce the following open propositions, respectively: *Mary was killed by X means* and *The chandelier was cleaned by X means*.

- (60) A: How was Mary killed?
B: A bullet killed her.
- (61) A: How did the chandelier finally get clean?
B: This rag cleaned it.

In B's responses, the instrument subject serves as the focus, filling in the variable of the activated open proposition. These sentences are felicitous in context, despite being odd in isolation.

4.1.3 Conversational Context and Implicit Questions

The pragmatic constraints on instrument subjects can also be satisfied through association with an open proposition activated by an implicit question. According to Roberts (1996, 2004), the structure of discourse is pervasively shaped by questions, many of them implicit. As we will see, implicit questions may emerge through the interaction of linguistic content with the non-linguistic context, or they may be generated on the basis of the non-linguistic context alone. They range from very general to quite specific. Roberts claims that one goal of conversation in general is to accumulate and share information about the world, i.e. to answer the question "What is the way the world is?". In specific conversations,

interlocutors aim to accomplish part of this goal by answering a set of questions with narrower scope, called the “questions under discussion.”

I assume that implicit questions, like their overt counterparts, activate open propositions that can serve as contextual support for the licensing of instrument subjects.³¹ In the scenario in (62), the context of conversation as well as A’s initial utterance enables B to infer a question, namely “How does one enter the lounge?”. The question activates an open proposition: informally, *One enters the lounge by X means*. The instrument subject in the second sentence of B’s response is associated with the means variable of the open proposition, which (along with the presence of bare focus) makes it relatively felicitous.

- (62) [Context: B approaches A, who is standing outside the student lounge looking puzzled.]
A: I can’t get in.
B: They changed the locks. THAT key opens the door now.

In (62), linguistic context plays a key role in enabling interlocutors to infer the implicit question. There are also cases in which implicit questions are formed entirely through the interlocutors’ knowledge of the goals associated with certain types of discourse. Consider, for example, conversations that occur in a museum setting. One conversational goal common to museum discourses is to come to an understanding of the significance of the artifacts or works of art on display. The context itself thus generates a very general implicit question, something along the lines of “What is the significance of each artifact?”. This question becomes more narrow when interlocutors focus on a particular artifact. If, for example, they come across a broom displayed as an artifact, the implicit question that is likely to be generated is “What was done with this broom?”, with the associated open

³¹ Roberts herself assumes a different approach to the semantics and pragmatics of questions; see Roberts (1996) for more details.

proposition *X was done with this broom*. As example (63) illustrates, instrument subjects are far more acceptable when they have contextual support of this kind than when they appear in isolation.

- (63) a. (#) A broom swept the room.
b. [Context: *description of exhibit of specimens associated with famous people*]
Twig from a broom that swept the room that Mao Tse Tung used as a student at Hunan University from 1917 to 1919.³²

Two more discourse genres that facilitate the licensing of instrument subjects are advertising and product reviews. Both genres are associated with a general implicit question along the lines of “What is the function of this product or service?”. As in the museum case, interlocutors generate a specific question when they are introduced to a particular product. This is illustrated by (64b), which originally appeared in a list of testimonials for a microfiber cloth designed for cleaning. Again, the instrument-subject sentence is more felicitous when it is associated with an open proposition than when it appears in isolation.

- (64) a. (#) The cloth cleaned the chandelier.
b. My friend called me back and said that the cloth cleaned the chandelier in no time at all.³³

What the discourse genres that facilitate instrument subjects share are implicit questions that deal with the ways in which inanimate objects function. In other words, they are contexts in which instrumentality is under discussion.

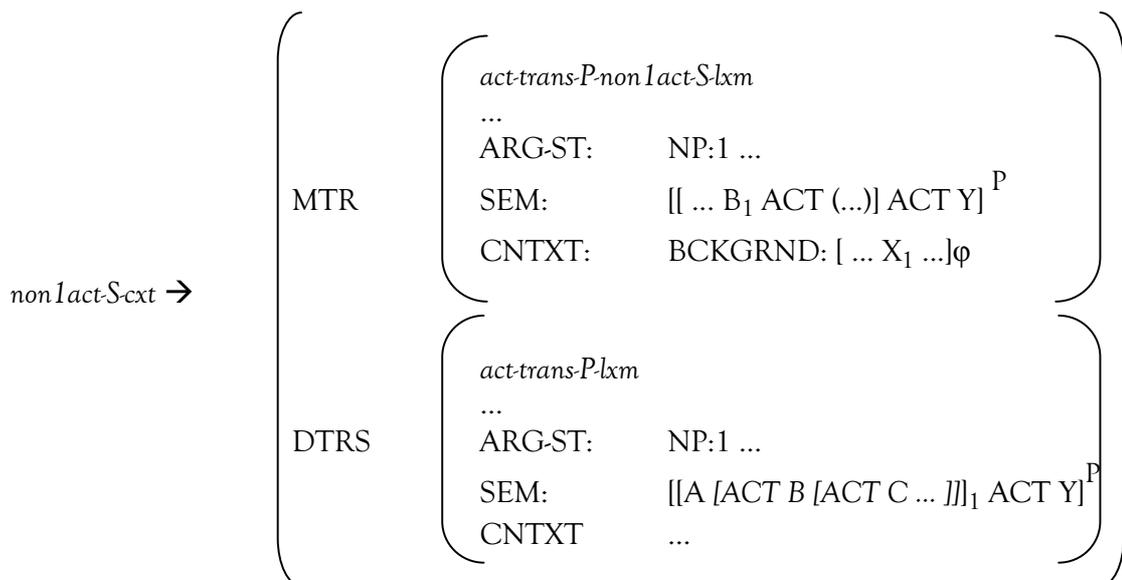
³² <http://legacy.lclark.edu/dept/gallery/objects/Artists.Specimens06.pdf>, retrieved 9/12/2009

³³ <http://www.microfibersunlimited.com/Customers.htm>, accessed 11/19/2009

4.2 Formalizing the Effects of Pragmatics

Figure 5 illustrates the construction that encodes the pragmatic constraints on the distribution of instrument subjects. I call it the Non-Initial Actor – Subject construction, abbreviated as *non1act-Sext*. Like all combinatoric constructions in SBCG, it is a means of composing a well-formed sign (the mother, or MTR) on the basis of one or more existing signs (the daughter(s), or DTRS). In this construction, the daughter sign is a lexeme belonging to the class ACT-trans-P-lxm, a polyvalent verb entailing physical affectedness. I have proposed that members of this class are associated with a particular event representation (SEM) and mapping between semantics and argument structure (ARG-ST), specifically that the first NP of ARG-ST is associated with an action chain. The mother sign, also a lexeme, contains a more complex link that holds between semantics, argument structure, and pragmatics (CNTXT). The first NP in ARG-ST is associated both with a non-initial actor in SEM and with an element of an activated proposition in CNTXT.

Figure 5. The Non-Initial Actor – Subject Construction



The non-initial status of the linked actor in the mother sign is indicated by ellipses at the beginning of the action chain, indicating an unspecified number of preceding subevents.³⁴

The activated proposition appears as part of the BCKGRND, a subcomponent of CNTXT representing a set of propositions that “correspond to the set of utterance felicity conditions, which any part of the utterance sign may in principle contribute to” (Sag 2010: 16).

The construction is derivational, in the morphological sense, composing lexemes from other lexemes. There are several reasons why capturing pragmatic constraints on instrument subjects through a derivational construction may not initially seem intuitive, but there are reasons why it is advantageous. First, the licensing of an instrument subject has no morphological reflections, at least in English. However, this is not a strong challenge to the derivational approach. Many argument structure alternations, perhaps most notably the causative-inchoative alternation, lack morphological reflections in English and yet are widely treated within the lexical system (e.g. Jackendoff 1990, Levin and Rappaport Hovav 1995).

Second, one could argue that the composition of a verb with its subject is a phrasal phenomenon, rather than a lexical phenomenon, and thus would be better addressed through a phrasal construction. This is reasonable, but as I pointed out in Section 1, what actually alternates in the instrument “subject” alternation is which element of the semantic

³⁴ The parenthetical ellipses at the end of the action chain indicate the optional presence of additional subevents that follow the linked actor.

representation is linked to the first NP on the ARG-ST list.³⁵ This element may not actually be realized as subject if the ARG-ST list is subsequently altered by another derivational construction, such as passivization. I assume that in the English passive construction, the first element of the daughter lexeme's ARG-ST list is demoted in the mother lexeme, becoming the controller of the optionally expressed *by*-phrase (Bresnan 2001, Sag 2010). If we treat the instrument subject construction as derivational, applying before the passive, it accounts for the fact that instruments can control *by*-phrases in the passive.

Finally, it might seem odd to include constraints on the context of utterance within the representations of lexemes. However, this is required elsewhere in the system in order to account for other aspects of pragmatics that are lexically encoded, such as conventional implicature and lexical presupposition triggers (at least on some accounts). What is new about the construction in Figure 5 is that it integrates pragmatic constraints with lexical semantic constraints in the mapping to argument structure.

4.3 Discourse Context and the Semantics of the Action Chain

The construction introduced above requires that a non-initial actor be associated with a salient proposition in order to be realized as subject. This predicts that pragmatic constraints apply only to instruments that are preceded by agents within the action chain. As I will demonstrate, this prediction is borne out in both English and Fongbè (Brousseau 1998) and is consistent with the data that Alexiadou and Schäfer (2006) provide from German, Dutch, and Greek.

³⁵ Recall from Chapter 2 that ARG-ST is unique to the representation of lexical signs; phrasal signs do not have ARG-ST features.

We start by re-introducing a diagnostic that distinguishes instruments that are controlled by agents from those that are not. As discussed in Section 3.1, Reinhart and Siloni (2003) claim that instruments can only appear as *with*-adjuncts when there is a controlling agent in the semantic representation. This diagnostic distinguishes passive sentences (65a), which are taken to have an implicit agent, from inchoative sentences (65b), which seem not to represent an agent.

- (65) a. The vase was smashed with the club.
b. * The vase smashed with the club.

We can use the same diagnostic to distinguish between *controlled* and *uncontrolled instruments*. In order for an instrument to be felicitously linked as a *with*-adjunct, it must be fully controlled by an agent. Any force that the instrument generates must be construed as coming about through the manipulation of the agent. The (a) sentences below are acceptable because the instrument is readily interpreted as being controlled. In contrast, the (b) sentences, though possible, are relatively odd because the instrument is construed as uncontrolled, generating and transmitting force independently of the agent.

- (66) a. Henry smashed the vase with a baseball bat.
b. (?) Henry smashed the vase with a flying baseball bat.
(67) a. Henry moved the dirt with a shovel.
b. Henry moved the dirt with a crane.

The *with*-adjunct diagnostic thus enables us to determine whether an instrument is compatible with agent control in a particular event. Significantly, uncontrolled instruments are more easily realized as subjects than instruments that are under agent control. This is illustrated below. Though both types of instruments are possible subjects, controlled

instruments require more contextual support than uncontrolled instruments, which are acceptable in isolation.

- (68) a. (#) The baseball bat smashed the vase.
b. The flying baseball bat smashed the vase.
(69) a. (#) The shovel moved the dirt.
b. The crane moved the dirt.

As Alexiadou and Schäfer demonstrate (2006: 44), a similar pattern is attested in a range of other languages, including Dutch (70), German (71) and Greek (72). In the following examples, uncontrolled instruments are acceptable subjects where their controlled counterparts are degraded or impossible.

- (70) a. (?) De steen heft de ruit gebroken.
the stone has the pane broken
'The stone broke the pane.'
b. De vliegende steen heeft de ruit gebroken.
the flying stone has the pane broken
'The flying stone broke the pane.'
(71) a. Der Kran hob die Kiste hoch.
'The crane picked the crate up.'
b. * Die Gabel hob die Kartoffel hoch.
'The fork picked the potato up.'
(72) to tsekuri espase to parathiro #(peftondas).
'The axe broke the window (by falling).'

As discussed in Section 2, several researchers have given informal explanations of this pattern. The controlled/uncontrolled distinction that I have proposed closely corresponds to one take on the distinction between intermediary and facilitating instruments. In this view, intermediary instruments independently generate force and facilitating instruments do not, with only the former being readily realized as subjects (Levin 1993, Kamp and Rossdeutscher 1994). It also recalls Schlesinger's Naturalness Condition 1.

The present account builds on these generalizations by providing them with a representational basis. I assume that the representation of a sentence with an uncontrolled instrument subject lacks an agent, because instruments of this class are incompatible with agent control. The instrument is represented as the first (and only) member of the action chain. For this reason, it is immune to the pragmatic constraints that hold of non-initial actors. In contrast, sentences with controlled instrument subjects represent the agent as the initial actor, and thus must be associated with a salient proposition. The following examples demonstrate that the distribution of the two instrument types is conditioned by pragmatics in English. Uncontrolled instruments are acceptable as subjects of event-reporting sentences, while controlled instruments are not.

- (73) A: What happened today?
 B: A flying baseball bat broke the crystal vase!
- (74) A: What happened today?
 B: # A baseball bat broke the crystal vase!

The data from Fɔ̀ngbè that Brousseau (1998) discusses can be accounted for in the same way. Recall her observation that relatively autonomous instruments are generally acceptable as subjects (75b), whereas controlled instrument subjects are infelicitous in the absence of an emphatic particle that provides pragmatic support: note the contrast between (75c) and (76).

- (75) a. Asíbà hù Kòkú k̀pó tú/ jìvì k̀pó.
 Asiba kill Koku with rifle/knife with
 'Asiba killed Koku with a rifle/knife.'
- b. Tú éíṣ hù Kòkú.
 rifle DEM kill Koku
 'This rifle killed Koku.'

c. ?? jìvì élò hù Kòkú.
 knife DEM kill Koku
 ‘This knife killed Koku.’

(76) Àsìyóví ó wè já àtín.
 axe DET EMP cut-into-pieces tree
 ‘This is the axe that cut a tree into pieces.’ *or*
 ‘It’s with the axe that they cut the tree into pieces.’

This is consistent with the present claim that uncontrolled instruments, as initial actors, are free of the pragmatic constraints associated with controlled instruments.

5. Conclusion

This chapter has presented a case that information structure constrains the distribution of instrument subjects by resolving underspecified input from the lexical semantic linking system. Underspecification arises through inheritance when the ACT-trans-P-lxm class, which permits multiple actors, inherits the general linking constraint stating that subjects must be actors. Resolution is implemented through a derivational construction which links non-initiating actors as subjects when they are linked to an activated proposition.

This analysis has significant implications for our understanding of the relationship between information structure, lexical semantics, and syntax. The resolution construction presented in this chapter, along with the direct licensing construction proposed in Chapters 3 and 4, can be seen as two concrete elements of the widely assumed but often elusive link between subjecthood and information structure. This will be discussed further in Chapter 6.

Chapter 6

Conclusions and Future Directions

1. Introduction

This dissertation has presented a novel approach to capturing the effects of information structure on argument realization. In addition, it has demonstrated that information structure plays a significant role in the selection of subjects in English. Chapter 1 introduced the interface model of argument realization, in which the lexical semantic and informational components jointly determine the selection of subjects and objects. It also presented the two components of the interface model that are the focus of this dissertation: direct licensing, in which a grammatical constituent is licensed to fulfill an informational function, and resolution, in which information structure mediates between lexical semantic constraints and surface patterns of argument linking. Chapter 2 sketched the necessary background for this project. It began by surveying previous approaches to argument realization, and then introduced the assumptions that I make about the nature of the informational component. Finally, it presented a brief overview of the theoretical framework that I adopt, Sign-Based Construction Grammar.

The following three chapters applied the interface model to a range of subject selection phenomena. In Chapter 3, I introduced the Topical Exclamative (TE) construction and demonstrated that the main-clause subject conventionally functions as a topic, though it is

not semantically selected. This is shown through several strands of evidence, the two most important being that subjects must be interpreted specifically and that the proposition expressed by the sentence must be construed as relevant to the subject referent. Both of these constraints reflect core features of topicality. Chapter 4 extends the analysis to Copy Raising (CR), another environment in which the matrix subject functions as a topic expression. Subject licensing in TE and CR is accomplished through a direct licensing construction that alters the combinatorial requirements of the main predicate so that it comes to select a topical subject. I also show that the topic-marking function of TE and CR is at the root of a wide range of additional pragmatic constraints that hold of each construction.

Chapter 5 turns to the phenomenon of resolution by information structure. Previous analyses of instrument-subject sentences have noted that information structure influences the construction's acceptability, but none have attempted to formalize this observation. I propose that the effects of information structure come in the form of a construction that states that a (non-initiating) instrument can be linked as subject when it is associated with an activated proposition. This construction resolves underspecification in the mapping between lexical semantic and syntax, which is represented through a hierarchical model of the lexicon structured by inheritance relations.

To summarize, this dissertation has proposed two ways in which to bring information structure into contact with argument realization, and illustrated that their effects are evident in English subject selection phenomena. It has also generated a wealth of questions for future research. I survey several particularly interesting questions in this chapter.

Section 2 revisits the relationship between information structure and argument realization, situating the constructional model advocated here within a broader context. In Section 3, I offer some preliminary remarks on the typological status of the interface model, particularly its relationship to previous classifications of languages according to the nature of the interface between information structure and syntax. In Section 4, I consider various ways in which the present model might be constrained. Finally, Section 5 explores several possible extensions of the interface model to other argument realization phenomena.

2. Linking Argument Realization to Information Structure

Chapter 1 of this dissertation began with DuBois's (2003: 34) observation that while there is clearly some relationship between information structure and the selection of subjects and objects, its properties have proved rather mysterious. DuBois himself argues that grammatical relations are shaped by information status (1987, 2003), a claim that has surfaced in varying forms in the literature. Like all constituents, subjects and objects are affected by the general tendency to place given information before new information (Firbas 1964, 1966, Prince 1981a, Horn 1986, Birner and Ward 1998). In a language such as English with relatively strict SVO word order, this typically results in discourse-activated subjects.

In recent years, a growing number of researchers have argued that the connection between information structure and argument realization goes beyond the probabilistic correlation with information status. Aissen (1999) proposes that constraints linking discourse prominence to grammatical functions underlie the passive alternation. Lambrecht (1995) argues that patterns of argument realization in English demonstrate that

there is a three-way interaction between lexical semantics, information structure, and grammatical relations. Goldberg (2004) provides an overview of the ways in which pragmatics and argument structure interact, many of which go beyond a simple correlation with information status.

The interface model advocated here provides a framework in which to characterize the relationship between information structure and argument realization. It consists of a body of constructions that change a predicate's combinatorial potential along with its informational requirements. The connection between information structure and argument realization can thus be seen as the composite of (1) the body of constructions that pair a pattern of argument linking with specific informational content and (2) the probabilistic connection between grammatical relations and information structure.

Therefore, the direct licensing and resolution constructions that I have proposed here provide only a partial picture of the relationship between subjecthood and information structure in English. Recall the claim I made in Chapter 1 that subjecthood is itself a generalization over the ways in which phrases are formally marked as prominent in a language. In SBCG, the subject is typically the first member of the argument structure (ARG-ST) and valence (VAL) lists, as well as the external argument (XARG). The representational component of the link between subjecthood and information structure, then, is the body of constructions that pair a particular information structure with a change to the XARG and/or the first members of the ARG-ST and VAL lists. The Topic Licensing construction that underlies TE and CR changes the features of the first VAL member (by removing the requirement that it be expletive *it*) and co-indexes this

constituent with the informational relation of topicality. The Non-Initial Actor – Subject construction that licenses instrument subjects specifies that a non-initiating actor can be linked to the first ARG-ST member if it is associated with an element of an activated proposition. These are presumably just two of many constructions that alter the formal correlates of subjecthood in response to informational constraints. Other candidate constructions include the passive, which alters the composition of the ARG-ST list (Sag 2010) and obeys informational constraints (Tomlin 1985, Thompson 1987, Aissen 1999). Another possibility is the null instantiation of the subject in “diary-drop” examples like (1), which requires that the null-instantiated referent be highly activated in the discourse (Haegeman 1990, Haegeman and Ihsane 2001, Reiman 1994, Scott 2004). This could be implemented through a construction that removes a member from the valence list (Michaelis 2009), with concomitant informational constraints.

(1) Ø Got up. Ø Saw that I didn’t have any milk. Ø Went to the store.

A full picture of the informational component of subjecthood in English would most likely involve all of these constructions as well as others. It would also require an account of how the constructions are related to each other. This could be formalized through inheritance relations, which structure the set of combinatorial constructions as well as lexeme class constructions (Michaelis 2009, Sag 2010, submitted, cf. Goldberg 1995, Michaelis and Lambrecht 1996a, Kay 2002). This in turn would require a better understanding of how the constructions themselves are constrained, e.g. whether there are any restrictions on the ways in which information structure can influence grammatical relations. We will return to this issue in Section 4.

3. A Typological Perspective

It has long been acknowledged that the effects of information structure on syntax are circumscribed in English, relative to many other languages. Li and Thompson (1976) propose that languages can be classified with respect to two dimensions: *subject-prominence* (SP) and *topic-prominence* (TP). In SP languages, including English and other Indo-European languages, a grammatical subject/predicate distinction forms the backbone of clause organization, while in TP languages, such as Mandarin, the basic clause structure is topic-comment.¹ According to Li and Thompson, this accounts for a wide range of differences between the two classes of languages. Because the grammatical subject is central in SP languages, they tend to have expletive subjects, indicating that the subject function must be filled even in the absence of a meaningful subject phrase. They also make relatively frequent use of the passive voice, which marks atypical subject choice. TP languages are less likely to have expletive subjects or a passive alternation. On the other hand, topic expressions play a more central role in grammatical phenomena in TP languages than in SP languages. This is evident in co-reference across clauses, which tends to be controlled by topics in TP languages but by subjects in SP languages. Another characteristic feature of TP languages is the widespread nature of Multiple Subject Constructions (MSCs), in which the first subject is a topic with no selectional ties to any predicate within the sentences. SP languages, in contrast, are said to lack MSCs.²

¹ The typology also contains languages that are both SP and TP, such as Korean and Japanese, and languages that are neither, such as Tagalog, but I will focus here on the central distinction between SP and TP languages.

² It is important to distinguish between true MSCs, in which multiple NPs exhibit grammatical properties of subjecthood, and structures with a detached initial NP that is not a grammatical subject (cf. Salmon 2008: 10). English has the latter (Rodman 1974, Lambrecht 1994) but not the former.

Li and Thompson's typology classifies languages according to the extent to which the subject function is grammaticalized, i.e. independent of information structure. Classifying English as a SP language thus amounts to a claim that the interface between information structure and syntax is restricted. It also makes the prediction that languages like English should have a relatively limited capacity for direct licensing. According to Li and Thompson, grammatical subjects (other than expletives) "must stand in a selectional relation with some predicate in the sentence" (p. 462). This rules out informational licensing of grammatical subjects, which are essential to sentence structure in SP languages.

Chapters 3 and 4 of this dissertation demonstrated that information structure has the capacity to license grammatical subjects in English. This observation is consistent with Li and Thompson's model, which conceives of the SP/TP distinction as a scale rather than two absolute values. Like Mandarin, English permits subjects to be licensed entirely on the basis of information structure. However, in English the direct licensing of topical subjects occurs only in extraposition constructions, while in Mandarin topic-licensing is available across constructions, including those with semantically-selected subjects. This suggests that languages vary along a continuum with respect to the extent to which information structure influences argument realization.

Previous typological accounts of other components of the interface between syntax and information structure also involve variation along a continuum. Huang claims that languages range from "hot" to "cool" with respect to the availability of null instantiation of arguments; "hot" languages such as English are the most restrictive (see Chapter 2, Section 3.2.4 for further discussion). In addition, languages vary with respect to the relationship

between information structure and word order, which is less direct in English than in many other languages. This is the case even in comparison with some closely related languages such as German, in which topicalization is far more widespread than in English (Huang 1984, Hawkins 1985). Moving further afield, Lambrecht (1994) and Ladd (1996) note that Romance languages such as French, Italian, and Romanian use syntax to mark many of the informational contrasts that English codes prosodically. The relationship between information structure and syntax is even closer in so-called “discourse-configurational” languages such as Finnish (Vilkuna 1989) and Hungarian (É. Kiss 1987), in which informational constructs such as topic and focus are associated with particular syntactic positions and thus underlie a wide variety of syntactic phenomena.

The wide variety of typological classification systems that have been proposed – subject-prominent vs. topic-prominent, configurational vs. discourse-configurational, hot vs. cool – raises the question of how the different dimensions interact. One hypothesis is that the relationship between information structure and syntax remains constant across dimensions. On this view, a language in which information structure is closely reflected in the syntax would be predicted to be topic-prominent, discourse-configurational, and cool. A competing hypothesis is that the dimensions of the interface should diverge within a language, with information structure coded to different degrees by different components of the syntax. Previous research has found intra-language variability of this sort across levels of representation. For example, Lambrecht (1994) and Ladd (1996) observe that the syntax of Romance languages more closely reflects information structure than that of English,

whereas English prosody is more expressive than that of Romance. The question is whether comparable trade-offs exist within the syntactic system.

Hawkins (1985, Chapter 2) argues in favor of trade-offs across syntactic domains. He seeks to tie together two observations: that German has freer word order than English and that English allows a wider range of arguments to appear as subjects. For example, instruments are more freely realized as instruments in English than in German (cf. Alexiadou and Schäfer 2006). Hawkins argues that the relative freedom of subject selection in English compensates for the fact that information structure often cannot be expressed through word order.³ This is because there is potential for a closer relationship between subjecthood and its informational correlates if subject selection is less semantically constrained. Though Hawkins compares only English and German, the proposal could easily be extended, given that English is known to have greater freedom in subject choice compared to many other languages (Van Voorst 1996, Guilfoyle 2000, Alexiadou and Schäfer 2006, Wolff et al. 2009).

Thus, Hawkins departs from the typological classification systems that we have seen so far in proposing that English has a relatively close relationship between information structure and syntax in the domain of argument selection. On first glance, this may seem to be at odds with Li and Thompson's (1976) typology, which predicts that English should have a low capacity for direct licensing by information structure. In fact, the two proposals are potentially compatible because each addresses a distinct aspect of argument selection: Li and Thompson's model deals with direct licensing, Hawkins's with resolution. It would

³ Hawkins's claim about German is the reverse: relatively free word order can be seen as compensating for strict limitations on subject selection.

be possible for a language to have little direct licensing with widespread effects of resolution (or vice versa). This is what the two models, taken together, suggest for English. It would require a more detailed picture of the relationship between information structure and argument realization in order to gauge whether this prediction is accurate.

In addition, this raises the question of whether the two components of interface model typically are equally strong within languages, or whether trade-offs of the type Hawkins observes are common. This is an interesting direction for future research.

4. Constraining the Interface Model

Another goal for future research is to identify constraints on the interface model, which determine how information structure can influence argument realization, in English and cross-linguistically. I will briefly illustrate the types of constraints that I have in mind. I have demonstrated that English allows direct licensing in the non-thematic subject positions of extraposition structures. In contrast, topic-prominent languages such as Mandarin and Japanese permit direct licensing in a wide range of environments. In the present framework, this distributional difference corresponds to a formal distinction. As I have argued, the construction that licenses topical subjects in extraposition structures *changes* the features of the first VAL member (by removing the requirement that it be an expletive), while direct licensing in Multiple Subject Constructions (MSCs) and indirect passives presumably *adds* a phrase to the valence list. Thus, we might hypothesize that in English, information structure can motivate a change in the features of a valence member, but not the addition of a new valent. This hypothesis should be tested on a larger body of argument realization phenomena. I suggest some possibilities in Section 5.

It is also essential to investigate whether there are constraints on the types of informational objects that can influence argument realization in specific ways. For example, it seems reasonable to hypothesize that direct licensing must involve an informational relation, rather than another component of information structure, such as information status. The relation *topic* underlies the cases of direct licensing that I have studied in this dissertation. This corresponds to the widespread claim that subjects in MSCs and indirect passives are licensed to serve as topics. I have yet to find any evidence for direct licensing as a consequence of constraints on information status. It would be premature to draw conclusions at this point, but this is clearly an important question for future research.

5. Extending the Interface Model

There are many additional directions that one could take in expanding the interface model. In this section I touch upon a few possibilities. First, it would be interesting to investigate whether there are informational effects on object selection in English that are comparable to the subject selection phenomena that we have seen. If the hypothesis that direct licensing in English is only possible in extraposition structures is correct, we would expect little direct licensing of objects, given the relative rarity of expletive objects. Still, there is one possible case that comes to mind: “free datives” such as (2) below (examples from Horn 2008: 170), which are not selected by the verb but denote participants that are highly relevant to the event it describes (cf. Webelhuth and Dannenberg 2006; see further discussion in Chapter 2, Section 3.2.3). More research is needed to determine whether free datives in English and other languages can be captured through direct licensing by information structure, as suggested (in rather different terms) by Shibatani (1994).

- (2) a. When I was a young girl, I had me a cowboy.
b. I'm gonna grab me a freight train.

Another possible direction is to determine whether the English dative alternation can be analyzed as an instance of resolution. As discussed in Chapter 2 (Section 3.2.2), many researchers have previously suggested that information structure conditions the dative alternation (Green 1974, Erteschik-Shir 1979, Thompson 1990, Goldberg 1995, 2004, Basilico 1998, Polinsky 1998, Arnold et al. 2000, Wasow 2002, Ruppenhofer 2004). However, it is an open question whether the informational constraints on the alternation should be expressed in probabilistic or representational terms.

More research is also needed in order to identify the inventory of informational constructs that have the capacity to underlie direct licensing. As mentioned above, there is considerable cross-linguistic evidence that subjects can be licensed to serve as topics. In addition, I have found evidence that in the Stage Subject construction, illustrated in (3), the subject is licensed to serve as a *stage topic* (cf. Erteschik-Shir 1997), the spatiotemporal setting in which a state of affairs takes place.

- (3) a. Recent years have witnessed a revival of interest in Marxism.
b. New Canaan has seen a rise in the cost of real estate.

This raises the question of whether comparable effects can be found for focus – and if not, why not.

Another important goal is to devise a means of capturing the commonalities across constructions in which information structure influences argument realization. One specific question that comes to mind is whether there is a formal relationship between the constructions that underlie the licensing of topics, which (arguably) include topicalization,

left dislocation, the unlinked topic construction, prolepsis, and Topic Licensing itself, the licensing of a referential subject in an extraposition structure. Identifying the inheritance relations that structure the set of “information structure constructions” would also go a long way towards the goal of constraining the interface model.

Finally, we might test the predictions that the interface model makes for the processing of argument realization phenomena during sentence comprehension. Exactly how this could be done is a question for future research, but I will offer some preliminary ideas here. If, as I have claimed, the licensing of the subject in TE, CR, and the instrument subject construction is pragmatically constrained, we would expect that the processing of these structures would be facilitated if the appropriate context is provided. Of course, contextual support facilitates the processing of *any* sentence, not just those that are pragmatically marked. Crucially, then, the interface model would predict that context would facilitate “information structure constructions” more strongly than their informationally neutral counterparts.

6. Conclusion

The interface between information structure and argument realization constitutes a relatively new and little-studied area of research. For this reason, this dissertation has raised far more question than it has answered. However, it has demonstrated that information structure has a direct and significant effect on subject selection in English, which has long been considered a primarily semantic phenomenon. In addition, it has proposed a formal framework that I have argued is suitable for analyzing a wide range of argument realization

phenomena, and that I hope will prove useful on the many possible paths for future research.

Bibliography

- Abbott, Barbara. 2003. Specificity and referentiality. Paper presented at ESSLLI XV, Vienna. <https://www.msu.edu/~abbottb/specandref.pdf>
- Achard, Michel. 2000. The distribution of French raising constructions. *Proceedings of the Berkeley Linguistics Society* 26: 1-15.
- Aissen, Judith. 1999. Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory* 17: 673-711.
- Akmajian, Adrian. 1972. Getting tough. *Linguistic Inquiry* 3: 373-377.
- Akmajian, Adrian. 1979. *Aspects of the Grammar of Focus in English*. New York: Garland.
- Alexiadou, Artemis and Florian Schäfer. 2006. Instrument subjects are agents or causers. *Proceedings of the West Coast Conference on Formal Linguistics* 25: 40-48.
- Anderson, Lloyd. 1986. Evidentials, paths of change, and mental maps: Typologically regular asymmetries. In W. Chafe and J. Nichols, eds., *Evidentiality: The Linguistic Coding of Epistemology*, 273-312. Norwood: Ablex.
- Anscombe, G.E.M. 1957. *Intention*. Oxford: Basil Blackwell.
- Ariel, Mira. 1990. *Accessing Noun-Phrase Antecedents*. New York: Routledge.
- Arnold, Jennifer, Anthony Losongco, Thomas Wasow and Ryan Ginstrom. 2000. Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76: 28-55.
- Asher, Nicholas. 2006. Things and their aspects. *Philosophical Issues* 16: 1-23.
- Asher, Nicholas and James Pustejovsky. 2005. Word meaning and commonsense metaphysics. www.semanticsarchive.net
- Asudeh, Ash. 2004. *Resumption as Resource Management*. Ph.D. Dissertation, Stanford University.
- Asudeh, Ash and Ida Toivonen. 2005. Copy raising and its consequences for perceptual reports. In A. Zaenen, J. Simpson, T. Holloway King, J. Maling, and C. Manning, eds., *Architectures, Rules, and Preferences: Variations on Themes by Joan W. Bresnan*, 49-67. Stanford, CA: CSLI.
- Asudeh, Ash and Ida Toivonen. 2006. Expletives and the syntax and semantics of copy raising. In M. Butt and T. Holloway King, eds., *Proceedings of the LFG06 Conference*. Stanford, CA: CSLI.
- Asudeh, Ash, and Ida Toivonen. 2009. Copy raising and perception. <http://http-server.carleton.ca/~asudeh/pdf/asudeh-toivonen09.pdf>

- Authier, J.-Marc and Lisa Reed. 1992. On the syntactic status of French affected datives. *The Linguistic Review* 9: 295-311.
- Babby, Leonard H. 1980. *Existential Sentences and Negation in Russian*. Ann Arbor, MI: Karoma.
- Baker, Mark C. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.
- Baker, Mark C. 1997. Thematic roles and syntactic structure. In L. Haegeman, ed., *Elements of Grammar: Handbook of Generative Syntax*, 73-137. Dordrecht: Kluwer.
- Basilico, David. 1998. Object position and predication forms. *Natural Language and Linguistic Theory* 16: 541-595.
- Basilico, David. 2003. The topic of small clauses. *Linguistic Inquiry* 34: 1-35.
- Beaver, David I. and Brady Z. Clark. 2008. *Sense and Sensitivity: How Focus Determines Meaning*. Malden, MA: Wiley-Blackwell.
- Belletti, Adriana and Luigi Rizzi. 1988. Psych-verbs and theta-theory. *Natural Language and Linguistic Theory* 6: 291-352.
- Berman, Ruth A. 1982. Dative marking of the affectee role. *Hebrew Annual Review* 6: 35-59.
- Birner, Betty J., Jeffrey P. Kaplan, and Gregory Ward. 2007. Functional compositionality and the interaction of discourse constraints. *Language* 83: 317-343.
- Birner, Betty J. and Gregory Ward. 1998. *Information Status and Noncanonical Word Order in English*. Amsterdam: John Benjamins.
- Bobaljik, Jonathan D. and Dianne Jonas. 1996. Subject positions and the roles of TP. *Linguistic Inquiry* 27: 195-236.
- Bolinger, Dwight. 1961. Syntactic blends and other matters. *Language* 37: 366-381.
- Bolinger, Dwight. 1973. Ambient *it* is meaningful too. *Journal of Linguistics* 9: 261-270.
- Borer, Hagit and Yosef Grodzinsky. 1986. Syntactic cliticisation and lexical cliticisation: The case of Hebrew dative clitics. In H. Borer, ed., *The Syntax of Pronominal Clitics*, vol. 19 of *Syntax and Semantics*, 175-217. New York: Academic Press.
- Borkin, Ann. 1973. To be and not to be. *Proceedings of the Chicago Linguistic Society* 9: 44-56.
- Bouchard, Denis. 1995. *The Semantics of Syntax: A Minimalist Approach to Grammar*. Chicago: University of Chicago Press.
- Bowdle, Brian F. and Dedre Gentner. 1997. Informativity and asymmetry in comparisons. *Cognitive Psychology* 34: 244-286.
- Brenier, Jason M. and Laura A. Michaelis. 2005. Optimization via syntactic amalgam: Syntax-prosody mismatch and copula doubling. *Corpus Linguistics and Linguistic Theory* 1: 45-88.

- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Malden, MA: Blackwell.
- Bresnan, Joan and Jonni Kanerva. 1989. Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20: 1-50.
- Brousseau, Anne-Marie. 1998. *Réalisations Argumentales et Classes du Verbes en Fongbè*. Paris: Peeters.
- Büring, Daniel. 2003. On D-trees, beans, and B-accents. *Linguistics and Philosophy* 26: 511-545.
- Büring, Daniel. 2006. Focus projection and default prominence. In V. Molnár and S. Winkler, eds., *The Architecture of Focus*, 321-346. Berlin/New York: De Gruyter.
- Carlson, Gregory N. 1977. *Reference to Kinds in English*. Ph.D. Dissertation, University of Massachusetts-Amherst.
- Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Li, ed., 25-55.
- Chen, Ping. 1996. Pragmatic interpretations of structural topics and relativization in Chinese. *Journal of Pragmatics* 26: 389-406.
- Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6: 338-405.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1981. *Lectures on Government and Binding Theory*. Dordrecht: Foris.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chung, Sandra. 1978. *Case Marking and Grammatical Relations in Polynesian*. Austin, TX: University of Texas Press.
- Cohen, Ariel and Nomi Erteschik-Shir. 2002. Topic, focus, and the interpretation of bare plurals. *Natural Language Semantics* 10: 125-145.
- Comrie, Bernard. 1986. Reflections on subject and object control, *Journal of Semantics* 4: 47-65.
- Coppock, Elizabeth. 2005. Alignment in syntactic blending. Ms., Stanford University.
- Coppock, Elizabeth. 2009. Parallel grammatical encoding in syntactic production: Evidence from syntactic blends. *Language and Cognitive Processes* 25: 28-39.
- Cresti, Diana. 1995. *Indefinite Topics*. Ph.D. Dissertation, MIT.
- Croft, William. 1991. *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: University of Chicago Press.
- Croft, William. 1994. The semantics of subjecthood. In M. Yaguello, ed., *Subjecthood and Subjectivity: The Status of the Subject in Linguistic Theory*. Paris: Orphrys.

- Croft, William. 1998. Event structure in argument linking. In M. Butt and W. Geuder, eds., *The Projection of Arguments: Lexical and Compositional Factors*, 1-43. Stanford, CA: CSLI.
- Croft, William, and D. Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- Cruse, D. Alan. 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.
- Culicover, Peter and Ray Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford University Press.
- Davies, William. 2005. Madurese prolepsis and its implications for a typology of raising. *Language* 81: 645-665.
- Davis, Anthony R. 2001. *Linking by Types in the Hierarchical Lexicon*. Stanford, CA: CSLI.
- Davis, Anthony R. and Jean-Pierre Koenig. 2000. Linking as constraints on word classes in a hierarchical lexicon. *Language* 76: 56-91.
- Davison, Alice. 1984. Syntactic markedness and the definition of sentence topic. *Language* 60: 797-846.
- DeLancey, Scott. 1984. Notes on agentivity and causation. *Studies in Language* 8: 181-213.
- DeLancey, Scott. 1991. Event construal and case role assignment. *Proceedings of the Berkeley Linguistics Society* 17: 338-353.
- Delin, Judy. 1995. Presupposition and shared knowledge in it-clefts. *Language and Cognitive Processes* 10: 97-120.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.
- Dik, Simon 1978. *Functional Grammar*. Amsterdam: North-Holland.
- Dik, Simon, Maria E. Hoffman, Jan R. de Jong, Sie Ing Djiang, Harry Stroomer, and Lourens de Vries. 1981. On the typology of focus phenomena. In T. Hoekstra, H. van der Hulst, and M. Moortgat, eds., *Perspectives on Functional Grammar*, 41-74. Dordrecht: Foris.
- Donnellan, Keith S. 1966. Reference and definite descriptions. *The Philosophical Review* 75: 281-304.
- Doron, Edit. 2003. Agency and voice: The semantics of the Semitic templates. *Natural Language Semantics* 11: 1-67.
- Doron, Edit and Caroline Heycock. 1999. Filling and licensing multiple specifiers. In D. Adger, S. Pintzuk, B. Plunkett, and G. Tsoulas, eds., *Specifiers: Minimalist Approaches*, 69-89. Oxford: Oxford University Press.
- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67: 547-619.

- Drubig, Hans Bernard and Wolfram Schaffer. 2001. Focus constructions. In M. Haspelmath, E. König, W. Östereicher and W. Raible, eds., *Language Universals and Language Typology: An International Handbook*, 1079-1104. Berlin: de Gruyter.
- Dryer, Matthew. 1996. Focus, pragmatic presupposition, and activated propositions. *Journal of Pragmatics* 26: 475-523.
- DuBois, John W. 1987. The discourse basis of ergativity. *Language* 63: 805-855.
- DuBois, John W. 2003. Argument structure: Grammar in use. In Du Bois et al. 2003, 11-60.
- DuBois, John W., Lorraine E. Kumpf, and William J. Ashby. 2003. *Preferred Argument Structure: Grammar as Architecture for Function*. Amsterdam: Benjamins.
- Elliot, Dale E. 1974. Toward a grammar of exclamations. *Foundations of Language* 11: 231-246.
- Endriss, Cornelia. 2009. *Quantificational Topics: A Scopal Treatment of Exceptional Wide Scope Phenomena*. Dordrecht: Springer.
- Endriss, Cornelia and Stefan Hinterwimmer. 2007. Direct and indirect aboutness topics. In C. Féry, G. Fanselow and M. Krifka, eds., *Interdisciplinary Studies of Information Structure* 6: 83-96.
- Engdahl, Elisabeth and Enric Vallduví. 1994. Information packaging and grammar architecture: A constraint-based approach. In E. Engdahl, ed., *Integrating Information Structure into Constraint-Based and Categorical Approaches*, 39-79. Amsterdam: ILLC, dYANA-2 Report R.1.3.B.
- Engdahl, Elisabeth and Enric Vallduví. 1996. Information Packaging in HPSG. In C. Grover and E. Vallduví, eds., *Edinburgh Working Papers in Cognitive Science, Vol 12: Studies in HPSG*, 1-32. Scotland: Centre for Cognitive Science, University of Edinburgh.
- Erteschik-Shir, Nomi. 1979. Discourse constraints on dative movement. In T. Givón, ed., *Discourse and Syntax*, vol. 12 of *Syntax and Semantics*, 441-467. New York: Academic Press.
- Erteschik-Shir, Nomi. 2007. *Information Structure: The Syntax-Discourse Interface*. Oxford: Oxford University Press.
- Faller, Martina. 2006. Evidentiality above and below speech acts. To appear in C. Paradis and L. Egberg, eds., *Functions of Language: Evidentiality*. <http://personalpages.manchester.ac.uk/staff/Martina.T.Faller/>
- Fillmore, Charles. 1968. The case for Case. In E. Bach and R. T. Harms, eds., *Universals in Linguistic Theory*, 1-68. New York: Holt, Rinehart and Winston.
- Fillmore, Charles. 1999. Inversion and constructional inheritance. In G. Webelhuth, J.-P. Koenig, and A. Kathol, eds., *Lexical and Constructional Aspects of Linguistic Explanation*, 113-128. Stanford, CA: CSLI.

- Fillmore, Charles, Paul Kay, and Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64: 501-538.
- Firbas, Jan. 1964. On defining the theme in functional sentence analysis. *Travaux Linguistiques de Prague* 1: 267-280.
- Firbas, Jan. 1966. Non-thematic subjects in contemporary English. *Travaux Linguistiques de Prague* 2: 239-256.
- Frazier, Lyn, Charles Clifton, Keith Rayner, Patricia Deevy, Sungryong Koh, and Markus Bader. 2005. Interface problems: Structural constraints on interpretation? *Journal of Psycholinguistic Research* 34: 201-231.
- Fujii, Tomohiro. 2005. Cycle, linearization of chains, and multiple case checking. In S. Blaho, L. Vicente, and E. Schoorlemmer, eds., *Proceedings of ConSOLE XIII*: 39-65.
- Gati, Itamar and Amos Tversky. 1984. Weighting common and distinctive features in perceptual and conceptual judgments. *Cognitive Psychology* 16: 341-370.
- Gisborne, Nikolas. 1998. The attributory structure, evidential meaning, and the semantics of sound-class English verbs. *UCL Working Papers in Linguistics* 10: 389-414.
- Gisborne, Nikolas and Jasper Holmes. 2007. A history of English evidential verbs of appearance. *English Language and Linguistics* 11: 1-29.
- Givón, Talmy. 1983. *Topic Continuity in Discourse*. Amsterdam/Philadelphia: John Benjamins.
- Givón, Talmy. 1984. *Syntax: A Functional-Typological Approach*. Amsterdam/Philadelphia: John Benjamins.
- Gleitman, Lila, Henry Gleitman, Carol Miller and Ruth Ostrin. 1996. Similar, and similar concepts. *Cognition* 58: 321-376.
- Godfrey, John, Edward Holliman and Jane McDaniel. 1992. SWITCHBOARD: Telephone speech corpus for research and development. *International Conference on Acoustics, Speech, and Signal Processing* 92: 517-520.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Goldberg, Adele E. 2000. Patient arguments of causative verbs can be omitted: The role of information structure in argument distribution. *Language Sciences* 23: 503-524.
- Goldberg, Adele E. 2002. Surface generalizations: An alternative to alternations. *Cognitive Linguistics* 13: 327-356.
- Goldberg, Adele E. 2004. Pragmatics and argument structure. In L. Horn and G. Ward, eds., *Handbook of Pragmatics*, 427-441. Oxford: Blackwell.
- Goldberg, Adele E. 2005. Constructions, lexical semantics, and the correspondence principle: Accounting for generalizations and subregularities in the realization of

- arguments. In N. Erteschik-Shir and T. Rapoport, eds., *The Syntax of Aspect*, 215-236. Oxford: Oxford University Press.
- Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Green, Georgia. 1974. *Semantics and Syntactic Regularity*. Bloomington: Indiana University Press.
- Gregory, Michelle L. and Laura A. Michaelis. 2001. Topicalization and left dislocation: A functional opposition revisited. *Journal of Pragmatics* 33: 1665-1706.
- Grimm, Scott. To appear. The bounds of subjecthood: Evidence from instruments. *Proceedings of the Berkeley Linguistics Society* 33.
- Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10: 279-326.
- Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, MA: MIT Press.
- Grosz, Barbara and Yael Ziv. 1998. Centering, global focus, and right dislocation. In M. Walker, A. Joshi and E. Prince, eds., *Centering Theory in Discourse*, 293-307. Oxford: Clarendon Press.
- Gruber, Jeffrey S. 1965. *Studies in Lexical Relations*. Ph.D. Dissertation, MIT.
- Gruber, Jeffrey S. 1976. *Lexical Structures in Syntax and Semantics*. Amsterdam: North-Holland.
- Guilfoyle, Eithne. 2000. Tense and N-features in Irish. In E. Guilfoyle and A. Carnie, eds., *The Syntax of Verb-Initial Languages*. Oxford: Oxford University Press.
- Gundel, Jeannette. 1985. 'Shared knowledge' and topicality. *Journal of Pragmatics* 9: 83-107.
- Haegeman, Liliane. 1990. Understood subjects in English diaries. *Multilingua* 9: 157-199.
- Haegeman, Liliane and Tabea Ihsane. 2001. Adult null subjects in the non-*pro*-drop languages: Two diary dialects. *Language Acquisition* 9: 329-346.
- Halliday, Michael. 1967. Notes on transitivity and theme in English, Part 2. *Journal of Linguistics* 3: 199-244.
- Hawkins, John. 1985. *A Comparative Typology of German and English: Unifying the Contrasts*. Austin, TX: University of Texas Press.
- Hedberg, Nancy. 2003. The prosody of contrastive topic and focus in spoken English. *Proceedings of the Workshop on Information Structure in Context*, 14-52. Stuttgart: Institut für Maschinelle Sprachverarbeitung.
- Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. Dissertation, University of Massachusetts-Amherst.
- Heycock, Caroline. 1994. *Layers of Predication*. New York: Garland.

- Heycock, Caroline. 2007. Japanese *wa*, *ga*, and information structure. In S. Miyagawa and M. Saito, eds., *Handbook of Japanese Linguistics*, 54-83. Oxford : Oxford University Press
- Heycock, Caroline and Edit Doron. 2003. Categorical subjects. *Gengo Kenkyu* 123: 95-135.
- Heycock, Caroline and Anthony Kroch. 1999. Pseudocleft connectivity: Implications for the LF interface. *Linguistic Inquiry* 30: 365-397.
- Heycock, Caroline and Anthony Kroch. 2002. Topic, focus and syntactic representation. *Proceedings of the West Coast Conference on Formal Linguistics* 21: 141-165.
- Horn, Laurence R. 1981a. A pragmatic approach to certain ambiguities. *Linguistics and Philosophy* 4: 321-358.
- Horn, Laurence R. 1981b. Exhaustiveness and the semantics of clefts. *Proceedings of the North East Linguistics Society* 11: 125-142.
- Horn, Laurence R. 1986. Presupposition, theme and variations. *Proceedings of the Chicago Linguistic Society* 22; *Papers from the Parasession on Pragmatics and Grammatical Theory*: 168-92.
- Horn, Laurence R. 1989. *A Natural History of Negation*. Chicago: University of Chicago Press.
- Horn, Laurence R. 2008. "I love me some him": The landscape of non-argument datives. In O. Bonami and P. Cabredo Hofherr, eds., *Empirical Issues in Syntax and Semantics* 7: 169-192.
- Hornby, Peter A. 1972. The psychological subject and predicate. *Cognitive Psychology* 3: 632-642.
- Huang, C.-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15: 531-574.
- Ifantidou, Elly. 2001. *Evidentials and Relevance*. Amsterdam: John Benjamins.
- Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- Jackendoff, Ray. 1978. Grammar as evidence for conceptual structure. In M. Halle, J. Bresnan, and G. Miller, eds., *Linguistic Theory and Psychological Reality*, 201-228. Cambridge, MA: MIT Press.
- Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge, MA: MIT Press.
- Jackendoff, Ray and Peter Culicover. 2003. The semantic basis of control. *Language* 79: 517-556.
- Jacobs, Joachim. 2001. Dimensions of topic-comment. *Linguistics* 39: 641-681.
- Joseph, Brian. 1976. Raising in Modern Greek: A copying process? In J. Hankamer and J. Aissen, eds., *Harvard Studies in Syntax and Semantics, Volume 2*, 241-281. Cambridge, MA: Harvard University Press.

- Kamp, Hans. 1975. Two theories about adjectives. In E. Keenan, ed., *Formal Semantics of Natural Language*, 123-155. Cambridge: Cambridge University Press.
- Kamp, Hans and Uwe Reyle. 1993. *From Discourse to Logic: Introduction to Model-theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht: Kluwer.
- Kamp, Hans and Antje Rossdeutscher. 1994. Remarks on lexical structure and DRS construction. *Theoretical Linguistics* 20: 97-164.
- Kay, Paul. 2002. English subjectless tagged sentences. *Language* 78: 453-481.
- Kay, Paul. 2005. Argument structure constructions and the argument-adjunct distinction. In M. Fried and H. Boas, eds., *Grammatical Constructions: Back to the Roots*, 71-98. Amsterdam: John Benjamins.
- Keenan, Edward L. 1976. Towards a universal definition of "subject." In Li, ed., 303-333.
- É. Kiss, Katalin. 1987. *Configurationality in Hungarian*. Dordrecht: Reidel.
- É. Kiss, Katalin. 1998. Identificational focus versus informational focus. *Language* 74: 245-273.
- Koenig, Jean-Pierre and Anthony Davis. 2003. Semantically transparent linking in HPSG. In S. Müller, ed., *Proceedings of the HPSG03 Conference*. Stanford, CA: CSLI.
- Kratzer, Angelika. 1995. Stage-level and individual-level predicates. In G.N. Carlson and F.J. Pelletier, eds., *The Generic Book*, 125-175. Chicago: University of Chicago Press.
- Krifka, Manfred. 2001. For a structured meaning account of questions and answers. In C. Féry and W. Sternefeld, eds., *Audiatur Vox Sapientia: A Festschrift for Arnim von Stechow*, 287-319. Berlin: Akademie Verlag.
- Krifka, Manfred. 2006. Basic notions of information structure. In C. Féry and M. Krifka, eds., *Interdisciplinary Studies of Information Structure* 6: 13-55.
- Krifka, Manfred, Cornelia Endriss, Peter Staudacher and Stefan Hinterwimmer. 2006. Quantification and topicality. Presented at *Quantifikation und Informationsstruktur*, Potsdam.
- Krifka, Manfred, Francis J. Pelletier, Gregory N. Carlson, Alice ter Meuleun, Godehard Link, and Gennaro Chierchia. 1995. Genericity: An introduction. In G.N. Carlson and F.J. Pelletier, eds., *The Generic Book*, 1-124. Chicago: University of Chicago Press.
- Kripke, Saul. 1977. Speaker's reference and semantic reference. *Midwest Studies in Philosophy* 2: 255-276.
- Kuno, Susumu. 1972. Functional sentence perspective: A case study from Japanese and English. *Linguistic Inquiry* 3: 269-336.
- Kuno, Susumu. 1973. *The Structure of the Japanese Language*. Cambridge, MA: MIT Press.
- Kuroda, S.-Y. 1972. The categorical judgment and thethetic judgment. *Foundations of Language* 9: 153-185.

- Kuroda, S.-Y. 1992. *Japanese Syntax and Semantics*. Dordrecht: Kluwer.
- Laca, Brenda. 1990. Generic objects: Some more pieces of the puzzle. *Lingua* 8: 25-46.
- Ladd, Robert. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
- Ladusaw, William. 1994. Thetic and categorial, stage and individual, weak and strong. *Proceedings of Semantics And Linguistic Theory* 4: 220-229.
- Lakoff, Robin. 1970. Another non-source for comparatives. *Linguistic Inquiry* 1: 128-129.
- Lambert, Silke. 2008. Pragmatic and stylistic relevance of free datives in German. Ms., University at Buffalo.
- Lambrecht, Knud. 1988. There was a farmer had a dog: Syntactic amalgams revisited. *Proceedings of the Berkeley Linguistics Society* 14: 319-339.
- Lambrecht, Knud. 1994. *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.
- Lambrecht, Knud. 1995. The pragmatics of case: On the relationship between semantic, grammatical, and pragmatic roles in English and French. In M. Shibatani and S. Thompson, eds., *Essays in Semantics and Pragmatics in Honor of Charles Fillmore*, 145-190. Amsterdam: John Benjamins.
- Lambrecht, Knud. 2000. When subjects behave like objects. *Studies in Language* 24: 611-682.
- Lambrecht, Knud. 2001. A framework for the analysis of cleft constructions. *Linguistics* 39: 463-516.
- Lambrecht, Knud and Laura A. Michaelis. 1998. Sentence accent in information questions: Default and projection. *Linguistics and Philosophy* 21: 477-544.
- Landau, Idan. 1999. Possessor raising and the structure of VP. *Lingua* 107: 1-37.
- Landau, Idan. 2009. Against broad subjects in Hebrew. *Lingua* 119: 89-101.
- Landman, Fred. 1989. Groups II. *Linguistics and Philosophy* 12: 723-744.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar: Theoretical Preliminaries*. Stanford, CA: Stanford University Press.
- Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar: Descriptive Application*. Stanford, CA: Stanford University Press.
- Langacker, Ronald W. 1995. Raising and transparency. *Language* 71: 1-62.
- Lappin, Shalom. 1983. Theta-roles and NP-movement. *Proceedings of the North East Linguistic Society* 13: 121-128.
- Lappin, Shalom. 1984. Predication and raising. *Proceedings of the North East Linguistic Society* 14: 236-252.

- Lasersohn, Peter. 2005. Context dependence, disagreement and predicates of personal taste. *Linguistics and Philosophy* 28: 643-686.
- Lee, Chungmin. 1999. Contrastive topic: A locus of the interface. In K. Turner, ed., *The Semantics/Pragmatics Interface from Different Points of View 1*, 317-341. London: Elsevier.
- Lee, Chungmin. 2004. Contrastive topic and/or contrastive focus. In B. McClure, ed., *Japanese/Korean Linguistics 12*. Stanford, CA: CSLI.
- Lee, Chungmin. 2006. Contrastive topic and focus in discourse. In K. von Stechow and K. Turner, eds., *Where Semantics Meets Pragmatics*, 381-420. London: Elsevier.
- Legendre, Géraldine, William Raymond, and Paul Smolensky. An Optimality-Theoretic typology of case and grammatical voice systems. *Proceedings of the Berkeley Linguistics Society* 19: 464-478.
- Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.
- Levin, Beth. 2009. Where do verb classes come from? Presentation given at *Verb Typologies Revisited*, Ghent, February 3-5 2009.
- Levin, Beth and Malka Rappaport Hovav. 1995. *Unaccusativity: At the Syntax-Lexical Semantics Interface*. Cambridge, MA: MIT Press.
- Levin, Beth and Malka Rappaport Hovav. 2005. *Argument Realization*. Cambridge: Cambridge University Press.
- Li, Charles N. (ed.). 1976. *Subject and Topic*. New York: Academic Press.
- Li, Charles N. and Sandra A. Thompson. 1976. Subject and topic: a new typology of language. In Li, ed., 457-489.
- Linzen, Tal. 2008. Hebrew possessive datives: An affectedness approach. Ms., Tel Aviv University.
- Ludlow, Peter and Stephen Neale. 1991. Indefinite descriptions: In defense of Russell. *Linguistics and Philosophy* 14: 171-202.
- Mack, Jennifer. To appear. Licensing by information structure. *Proceedings of the Berkeley Linguistics Society* 33.
- Mack, Jennifer and Yael Fuerst. To appear. English optional expletives and the pragmatics of judgments. *Proceedings of the Chicago Linguistic Society* 45.
- Maling, Joan. 1983. Transitive adjectives: A case of categorial reanalysis. In F. Heny and B. Richards, eds., *Linguistic Categories: Auxiliaries and Related Puzzles*, 253-289. Dordrecht: Foris.
- Manning, Christopher D. and Ivan A. Sag. 1999. Dissociations between argument structure and grammatical relations. In G. Webelhuth, J.-P. Koenig, and A. Kathol, eds., *Lexical and Constructional Aspects of Linguistic Explanation*, 63-77. Stanford, CA: CSLI.

- Manzini, Rita. 1983. On control and control theory. *Linguistic Inquiry* 19: 335-391.
- Marantz, Alec. 1984. *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.
- Matushansky, Ora. 2002. Tipping the scales: The syntax of scalarity in the complement of *seem*. *Syntax* 5: 219-276.
- Maurer, Gail and Michael K. Tanenhaus. 1995. Implicit arguments in sentence processing. *Journal of Memory and Language* 34: 357-382.
- McCloskey, James. 1997. Subjecthood and subject positions. In L. Haegeman, ed., *Elements of Grammar*, 197-235. Dordrecht: Kluwer.
- McCloskey, James and Peter Sells. 1988. Control and A-chains in Modern Irish. *Natural Language and Linguistic Theory* 16: 149-189.
- McNally, Louise. 1998. Stativity and theticity. In S. Rothstein, ed., *Events in Grammar*, 293-307. Dordrecht: Kluwer.
- Michaelis, Laura A. 2010. Making the case for Construction Grammar. In H. Boas and I. Sag, eds., *Sign-Based Construction Grammar*. Stanford, CA: CSLI Publications.
- Michaelis, Laura A. 2009. Sign-Based Construction Grammar. In B. Heine and H. Narrog, eds., *The Oxford Handbook of Linguistic Analysis*, 155-176. Oxford: Oxford University Press.
- Michaelis, Laura A. 2001. Exclamative constructions. In M. Haspelmath, E. König, W. Österreicher and W. Raible, eds., *Language Universals and Language Typology: An International Handbook*, 1038-1050. Berlin: De Gruyter.
- Michaelis, Laura A. and Hartwell S. Francis. 2007. Lexical subjects and the conflation strategy. In N. Hedberg and R. Zacharski, eds., *Topics in the Grammar-Pragmatics Interface: Papers in Honor of Jeanette K. Gundel*, 19-48. Amsterdam: John Benjamins.
- Michaelis, Laura A. and Knud Lambrecht. 1996a. Toward a construction-based theory of language function: The case of Nominal Extraposition. *Language* 72: 215-247.
- Michaelis, Laura A. and Knud Lambrecht. 1996b. The exclamative sentence type in English. In A. Goldberg, ed., *Conceptual Structure, Discourse and Language*, 375-389. Stanford, CA: CSLI.
- Mikkelsen, Line. 2004. Specificational subjects. www.semanticsarchive.net.
- Miller, Philip H. 2001. Discourse constraints on (non)extraposition from subject in English. *Linguistics* 39(4): 683-701.
- Mithun, Marianne. 1986. Evidential diachrony in Northern Iroquoian. In W. Chafe and J. Nichols, eds., *Evidentiality: The Linguistic Coding of Epistemology*, 89-112. Norwood: Ablex.
- Mithun, Marianne. 1991. The role of motivation in the emergence of grammatical categories: The grammaticization of subjects. In E.C. Traugott and B. Heine, eds., *Approaches to Grammaticalization, Vol. 2*, 159-184. Amsterdam: John Benjamins.

- Moore, John. 1998. Turkish copy-raising and A-chain locality. *Natural Language and Linguistic Theory* 16: 149-189.
- Newman, John. 1982. *The Semantics of Raising Constructions*. Ph.D. Dissertation, University of California-San Diego.
- Oshima, David. 2006. Adversity and Korean/Japanese passives: A constructional analogy. *Journal of East Asian Linguistics* 15: 137-166.
- Otsuka, Tatsuo. 2006. On pragmatically licensed extra-thematic arguments in Japanese. *Bulletin of the Faculty of Education, Chiba University II: Humanities and Social Sciences* 47: 84-89.
- Pan, Haihua and Jianhua Hu. 2008. A semantic-pragmatic account of (dangling) topics in Mandarin Chinese. *Journal of Pragmatics* 40: 1966-1981.
- Partee, Barbara H. 1971. On the requirement that transformations preserve meaning. In C. Fillmore and D. Langendoen, eds., *Studies in Linguistic Semantics*, 1-22. New York: Holt, Rinehart and Winston.
- Partee, Barbara H. 1992. Topic, focus, and quantification. In A. Wyner and S. Moore, eds., *Proceedings of Semantics And Linguistic Theory* 1: 257-280.
- Partee, Barbara H. 1995. Lexical semantics and compositionality. In L. Gleitman and M. Liberman, eds., *An Invitation to Cognitive Science I: Language*, 311-360. Cambridge, MA: MIT Press.
- Perlmutter, David M. 1980. Relational grammar. In E. A. Moravcsik and J. R. Wirth, eds., *Current Approaches to Syntax*, vol. 13 of *Syntax and Semantics*, 195-229. New York: Academic Press.
- Perlmutter, David M. and Paul M. Postal. 1984. The 1-Advancement Exclusiveness Law. In D. M. Perlmutter and C. Rosen, eds., *Studies in Relational Grammar* 2, 81-125. Chicago: University of Chicago Press.
- Pesetsky, David. 1987. Binding problems with experiencer verbs. *Linguistic Inquiry* 18: 126-140.
- Pesetsky, David. 1995. *Zero Syntax: Experiencers and Cascades*. Cambridge, MA: MIT Press.
- Polinsky, Maria. 1998. A non-syntactic account of some asymmetries in the double object construction. In J.-P. Koenig, ed., *Discourse and Cognition: Bridging the Gap*, 403-422. Stanford, CA: CSLI.
- Pollard, Carl and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Stanford, CA: CSLI.
- Portner, Paul and Katsihiko Yabushita. 1998. The semantics and pragmatics of topic phrases. *Linguistics and Philosophy* 21: 117-157.
- Portner, Paul and Katsihiko Yabushita. 2001. Specific indefinites and the information structure theory of topics. *Journal of Semantics* 18: 271-297.

- Portner, Paul and Raffaella Zanuttini. 2004. The semantics of nominal exclamatives. In R. Elugardo and R.J. Stainton, eds., *Ellipsis and Nonsentential Speech*, 57-67. Dordrecht: Kluwer.
- Postal, Paul. 1974. *On Raising*. Cambridge, MA: MIT Press.
- Postal, Paul and John R. Ross. 1971. Tough movement si, tough deletion no! *Linguistic Inquiry* 2: 544-546.
- Potsdam, Eric, and Jeffrey T. Runner. 2001. Richard returns: Copy raising and its implications. *Proceedings of the Chicago Linguistic Society* 37: 453-468.
- Prince, Ellen f. 1978. A comparison of *wh*-clefts and *it*-clefts in discourse. *Language* 54: 883-906.
- Prince, Ellen F. 1981a. Toward a taxonomy of given/new information. In P. Cole, ed., *Radical Pragmatics*, 223-254. New York: Academic Press.
- Prince, Ellen F. 1981b. Topicalization, focus-movement, and Yiddish-movement: A pragmatic differentiation. *Proceedings of the Berkeley Linguistics Society* 7: 249-265.
- Prince, Ellen F. 1981c. On the inferencing of indefinite *this* NPs. In A. Joshi, B. Webber and I. Sag, eds., *Elements of Discourse Understanding*, 231-250. Cambridge: Cambridge University Press.
- Prince, Ellen F. 1986. On the syntactic marking of presupposed open propositions. *Proceedings of the Chicago Linguistic Society* 22; *Papers from the Parasession on Pragmatics and Grammatical Theory*: 208-222.
- Prince, Ellen. 1992. The ZPG letter: Subjects, definites, and information status. In S. Thompson and W. Mann, eds., *Discourse Description: Diverse Analyses of a Fundraising Text*, 295-325. Philadelphia: John Benjamins.
- Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.
- Pustejovsky, James. 2006. Type theory and lexical decomposition. www.semanticsarchive.net
- Pustejovsky, James and Pierette Bouillon. 1995. Aspectual coercion and logical polysemy. *Journal of Semantics* 12: 133-162.
- Ramchand, Gillian. 2003. First phase syntax. Ms., University of Oxford.
- Rappaport Hovav, Malka and Beth Levin. 1998. Building verb meanings. In M. Butt and W. Geuder, eds., *The Projection of Arguments: Lexical and Compositional Factors*, 97-134. Stanford, CA: CSLI.
- Reiman, Patricia. 2004. Subjectless sentences in English. *UTA Working Papers in Linguistics* 1: 141-152.
- Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27: 53-94.
- Reinhart, Tanya. 2002. The theta system: An overview. *Theoretical Linguistics* 28: 229-290.

- Reinhart, Tanya and Tal Siloni. 2003. The lexicon-syntax parameter: Reflexivization and other arity operations. *Linguistic Inquiry* 36: 389-436.
- Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. In J.-H. Yoon and A. Kathol, eds., *Ohio State University Working Papers in Linguistics* 49. 1998 revision: <http://www.ling.ohio-state.edu/~croberts/infostr.pdf>
- Roberts, Craige. 2004. Context in dynamic interpretation. In L.R. Horn and G. Ward, eds., *The Handbook of Pragmatics*, 197-220. Malden, MA: Blackwell.
- Rodman, Robert. 1974. On left dislocation. *Papers in Linguistics* 7: 437-466.
- Rogers, Andy. 1971. Three kinds of physical perception verbs. *Proceedings of the Chicago Linguistic Society* 7: 206-222.
- Rogers, Andy. 1972. Another look at flip perception verbs. *Proceedings of the Chicago Linguistic Society* 8: 303-315.
- Rogers, Andy. 1973. *Physical Perception Verbs in English: A Study in Lexical Relatedness*. Ph.D. dissertation, UCLA.
- Rogers, Andy. 1974. A Transderivational Constraint on Richard? *Proceedings of the Chicago Linguistic Society* 10: 551-558.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1: 75-116.
- Rosenbaum, Peter. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: MIT Press.
- Rosengren, Inger. 1997. The thematic/categorical distinction revisited once more. *Linguistics* 35: 439-479.
- Ross, J.R. 1982. Pronoun deleting processes in German. Paper presented at the Annual Meeting of the Linguistic Society of America, San Diego, CA.
- Ruppenhofer, Josef. 2004. *The Interaction of Valence and Information Structure*. Ph.D. Dissertation, University of California-Berkeley.
- Sag, Ivan A. Submitted. English filler-gap constructions. <http://lingo.stanford.edu/sag/publications.html>
- Sag, Ivan A. 2010. Sign-Based Construction Grammar: An informal synopsis. In H. Boas and I. A. Sag, eds., *Sign-Based Construction Grammar*. Stanford, CA: CSLI Publications.
- Sag, Ivan A., Emily Bender, and Thomas Wasow. 2003. *Syntactic Theory: A Formal Introduction, Second Edition*. Stanford, CA: CSLI Publications.
- Sag, Ivan A., and Carl Pollard. 1991. An integrated theory of complement control. *Language* 67: 63-113.
- Salmon, William N. 2008. *Dislocations, Context, and Composition: Or, Double Subjects in Brazilian Portuguese*. Ph.D. Dissertation, Yale University.

- Sasse, Hans-Jurgen. 1987. The thetic/categorical distinction revisited. *Linguistics* 25: 511-580.
- Schlesinger, Izchak M. 1989. Instruments as agents: On the nature of semantic relations. *Journal of Linguistics* 25: 189-210.
- Schmerling, Susan F. 1978. Synonymy judgments as syntactic evidence. In P. Cole, ed., *Pragmatics*, vol. 9 of *Syntax and Semantics*, 299-314. New York: Academic Press.
- Scott, Katherine. 2004. *Diary Drop in English: A Syntactic and Pragmatic Analysis*. M.A. Thesis, University College London.
- Selkirk, Elisabeth O. 1984. *Phonology and Syntax*. Cambridge, MA; MIT Press.
- Shi, Dingxu. 2000. Topic and topic-comment constructions in Mandarin Chinese. *Language* 76: 383-408.
- Shibatani, Masayoshi. 1994. An integrational approach to possessor raising, ethical datives, and adversative passives. *Proceedings of the Berkeley Linguistics Society* 20: 461-486.
- Stalnaker, R. 1978. Assertion. In P. Cole, ed., *Pragmatics*, vol. 9 of *Syntax and Semantics*, 315-332. New York: Academic Press.
- Stephenson, Tamina. 2007. *Towards a Theory of Subjective Meaning*. Ph.D. Dissertation, MIT.
- Strawson, P.F. 1964. Identifying reference and truth values. In D. Steinberg and L. Jakobovits, eds., *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*, 86-99. Cambridge: Cambridge University Press.
- Szabó, Zoltan. 2003. On qualification. *Philosophical Perspectives 17: Language and Philosophical Linguistics*, 409- 38.
- Szabolcsi, Anna. 1981. Compositionality in focus. *Folia Linguistica Societatis Linguisticae Europaeae* 15: 141-162.
- Talmy, Leonard. 1976. Semantic causative types. In M. Shibatani, ed., *The Grammar of Causative Constructions*, 43-116. New York: Academic Press.
- Talmy, Leonard. 1978. Figure and ground in complex sentences. In J.H. Greenberg, ed., *Universals of Human Language, Vol. 4: Syntax*, 625-649. Stanford, CA: Stanford University Press.
- Talmy, Leonard. 1983. How language structures space. In H. L. Pick and L. P. Acredolo, eds., *Spatial Orientation*, 225-282. New York: Plenum Press.
- Talmy, Leonard. 1988. Force dynamics in language and cognition. *Cognitive Science* 12: 49-100.
- Talmy, Leonard. 2000. *Toward a Cognitive Semantics*. Cambridge, MA: MIT Press.
- Thompson, Sandra. 1987. The passive in English: A discourse perspective. In R. Channon and L. Shockley, eds., *In Honor of Ilse Lehiste*, 497-511. Dordrecht: Foris.

- Thompson, Sandra. 1990. Information flow and 'dative shift' in English. In J. Edmondson, K. Feagan, and P. Mühlhäusler, eds., *Development and Diversity: Linguistic Variation Across Time and Space*, 239-253. Dallas: SIL.
- Thrasher, Richard. 1977. *One Way to Say More by Saying Less: A Study of So-Called Subjectless Sentences*. Tokyo: The Eihosha Ltd.
- Tomlin, Russell. 1985. Interaction of subject, theme, and agent. In J. Worth, ed., *Beyond the Sentence: Discourse and Sentential Form*, 61-80. Ann Arbor, MI: Karoma.
- Tversky, Amos. 1977. Features of similarity. *Psychological Review* 84: 327-350.
- Tversky, Amos and Itamar Gati. 1978. Studies of similarity. In E. Rosch and B. Lloyd, eds., *Cognition and Categorization*, 79-98. Hillsdale, NJ: Erlbaum.
- Ura, Hiroyuki. 1998. Checking, economy, and copy-raising in Igbo. *Linguistic Analysis* 28: 67-88.
- Vallduvi, Enric. 1992. *The Informational Component*. New York: Garland.
- Vallduvi, Enric and Maria Vilkuna. 1998. On rheme and kontrast. In P. Culicover and L. McNally, eds., *The Limits of Syntax*, vol. 29 of *Syntax and Semantics*, 79-106. New York: Academic Press.
- Van Oosten, Jeanne. 1986. *The Nature of Subjects, Topics and Agents: A Cognitive Explanation*. Ph.D. Dissertation, University of California-Berkeley.
- Van Valin, Robert D., Jr. and David P. Wilkins. 1996. The case for "effector": Case roles, agents and agency revisited. In M. Shibatani and S. Thompson, eds., *Grammatical Constructions: Their Form and Meaning*, 289-322. Oxford: Oxford University Press.
- Van Valin, Robert D., Jr. and Randy J. LaPolla. 1997. *Syntax: Structure, Meaning, and Function*. Cambridge: Cambridge University Press.
- Van Voorst, Jan. 1996. Some systematic differences between the Dutch, French and English transitive construction. *Language Sciences* 18: 227-245.
- Vilkuna, Maria. 1989. *Free Word Order in Finnish*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Vilkuna, Maria. 1994. Discourse configurationality in Finnish. In K.É. Kiss, ed., *Discourse Configurational Languages*, 244-268. Oxford: Oxford University Press.
- von Stechow, Arnim. 1990. Current issues in the theory of focus. In A. von Stechow and D. Wunderlich, eds., *Semantik: Ein Internationales Handbuch Der Zeitgenössischen Forschung*, 804-825. Berlin: De Gruyter.

- Ward, Gregory. 1985. *The Semantics and Pragmatics of Preposing*. Ph.D. Dissertation, University of Pennsylvania.
- Ward, Gregory and Betty J. Birner. 1996. On the discourse function of rightward movement in English. In A. Goldberg, ed., *Conceptual Structure, Discourse and Language*, 463-479. Stanford, CA: CSLI.
- Wasow, Thomas. 2002. *Postverbal Behavior*. Stanford, CA: CSLI.
- Webelhuth, Gert and Clare J. Dannenberg. Southern American English personal datives: The theoretical significance of dialectal variation. *American Speech* 81: 31-55.
- Wechsler, Stephen. 1995. *The Semantic Basis of Argument Structure*. Stanford, CA: CSLI.
- Wolff, Phillip, Ga-Hyun Jeon, and Yu Li. 2009. Causers in English, Korean, and Chinese and the individuation of events. *Language and Cognition* 1-2: 165-194.
- Xu, Liejiong. 2001. The topic-prominence parameter. In H. Pan, ed., *Studies in Chinese Linguistics II*, 209-234. Hong Kong: The Linguistic Society of Hong Kong.
- Xu, Liejiong and D.T. Langendoen. 1985. Topic structures in Chinese. *Language* 61: 1-27.
- Zanuttini, Raffaella and Paul Portner. 2003. Exclamative clauses: At the syntax-semantics interface. *Language* 79: 39-81.
- Zimmerman, Malte. 2006. Contrastive focus. In C. Féry, M. Fanselow, and M. Krifka, eds., *Interdisciplinary Studies of Information Structure* 6: 147-159.