Understanding the *Self*: The Distribution of Anaphora within Prepositional Phrases

by

Jenny Simone Lederer

B.A. (University of California, Santa Cruz) 1999
M.A. (University of California, Berkeley) 2003

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Linguistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Eve Sweetser, Chair
Professor Lynn Nichols
Professor Line Mikkelsen
Professor Milton Azevedo

Spring 2009

Understanding the *Self*: The Distribution Anaphora within Prepositional Phrases

by Jenny Simone Lederer

Abstract

Understanding the *Self*: The Distribution of Anaphora within Prepositional Phrases

by

Jenny Simone Lederer

Doctor of Philosophy in Linguistics

University of California, Berkeley

Professor Eve Sweetser, Chair

This thesis investigates the distribution of reflexive and nonreflexive pronouns in the prepositional phrase, concluding that multiple semantic factors play a role in the appearance of one pronoun over the other. The distributional trends in English and Spanish are explained by referencing the crucial role space plays in grammar, and the resulting implications for Binding Theory are discussed. The motivating forces for the corpus distribution are based on perceived directionality and location of the denoted event with respect to the body of the event's protagonist. The patterns found in the corpus data are attributed to a range of factors that play a role in the semantic specifications and associations of the pronouns themselves.

First, it is argued that the high rate of reflexive pronouns in events that are metaphorically located in the body is due to the reflexive pronoun's close semantic association with the concept of self, a metaphorical body-internal entity. Second, it is argued that the reflexive pronoun is used to signal either an event which is performed on the body (in the referent's peri-personal space) or directed toward the body. Cases

of these types are explained by a schematic, semantic parallelism between syntactically complex reflexive events and syntactically simple reflexive events. In both cases, the reflexive pronoun signals a contrastive element. In syntactically complex cases, the PP examples (e.g. *John pushed the box toward himself*), and syntactically simple cases, those with basic clause structure (e.g. *John kicked himself*), the reflexive is used to signal that the direction of the event is counter to the direction of expectation, thus explaining why certain reflexive events (e.g. bathe, or pull something toward you) do not have to, and most often, do not occur with the reflexive pronoun.

Supplemental to the discussion of spatial parameters on the distribution of anaphora, this thesis supports the claim that a higher occurrence of the reflexive pronoun is associated with character (as opposed to narrator/speaker) viewpoint. A link between literal space and metaphorical perspective is made through a binary set of participant roles: Source and Target. These schematized, abstracted roles often conflate two dimensions present in most events: energy transfer and metaphorical perspective, and this association serves as an explanation into to why the reflexive is connected not only to certain spatial configurations but also to metaphorical viewpoint.

Contents

List of Tables

Acknowledgements

"Slow and steady wins the race", and in my case, finishes a dissertation. This project began six years ago in the spring of 2003, when I attended a graduate course on anaphora with Lynn Nichols. It (along with me) has endured six years of teaching, several unrelated projects, and the birth of two children. During some relatively long periods "on hold", when my attention turned to other parts of my life, it was never forgotten, and in some ways, has benefited from the aging process.

I have truly loved the process of writing this thesis along with my time in the Linguistics Department at UC Berkeley. In fact, my graduate experience has always been about the process and not the final product. I can honestly think of no better way to spend the last eight years of my life. I feel so lucky to enjoy what I do, and be able to actually work doing what I love. I have many friends and teachers to thank for my experience. These people include my committee members Eve Sweetser, Lynn Nichols, and Line Mikkelsen, other mentors at previous moments in my graduate career, Andrew Garrett, John Ohala, and George Lakoff, along with members of my student cohort, the staff of the Linguistics Department, namely Belén Flores, and the other friends I've made at Cal since 2001. I also thank the Linguistics Department and the Psychology Department at UC Santa Cruz for introducing me to the study of language and cognition. I still vividly remember my first sleepy morning in Syntax I with Jim McCloskey, back in 1997, where I sat in the last desk of the row thinking to myself, *what the hell is this?* Twelve years later, I can now actually answer that question with some kind of authority…

On a personal note, I thank my *marido*, Rafael Matos-Galí, my parents Karen Alaria and Nick Lederer, my late grandfather Walther Lederer, his wife Evelyn, my uncle Tom Lederer, and my closest friends from past and present. Without all of their support and encouragement throughout my education, I would never have been able to achieve this dissertation.

Finally, and as humbly as possible, *I* would like to thank *me*. I say this not just because the topic of this dissertation is binding and anaphora, not because this sentence violates binding principles, but because I am, in fact, proud of my own endurance and accomplishment. I hope that two very special little people in my life one day read this dissertation and realize how intellectual pursuits are challenging, and fun, and make you feel good about yourself.

> The abstract system of rules and principles constructed by theorists seldom emerge organically from the sensitive, fine-grained description of fully representative data (hence their mortality rate is high and their lifetime often tragically short). In brief, all the glory attaches to general principles and abstract theory; careful attention to the minutiae of language data is left for those without the insight and imagination to be good theorists.
>
> *Langacker, R. W, (2002:262), Concept, Image and Symbol: The Cognitive Basis of Grammar.*

Chapter 1

Anaphora and grammar

1 Introduction

The goal of this dissertation is to illuminate the reasons as to why reflexive pronouns such as *himself* and nonreflexive pronouns such as *him* vary in their distribution within prepositional phrases. The phenomenon is exemplified in the following data, which come from previous research on the topic:[1]

(1.1) *Corporal Crump$_i$ pinned the medal beside him$_i$ / *himself$_i$ (on the wall).* (Wechsler 1997: 15 #39a)

(1.2) *Corporal Crump$_i$ pinned the medal onto *him$_i$ / himself$_i$.* (Wechsler 1997: 15 #39b)

(1.3) *Krag$_i$ the robot placed a sandwich in front of him$_i$/?*himself$_i$.* (Faltz 1977: 106 #3.20)

(1.4) *Krag$_i$ the robot unscrewed a panel in his abdomen and place a sandwich inside himself$_i$/?him$_i$.* (Faltz 1977: 106 #3.21)

(1.5) *John$_i$ pulled the blanket over himself$_i$/ him$_i$.* (Kuno 1987: 66 #9.18a,b)

[1] Grammaticality judgments are provided by original (cited) author. Examples of my own are judged for grammaticality by me. In this case, Faltz uses the marking ?* to question the grammaticality of the reflexive in this example. In all other cases, cited author-generated examples and my own are evaluated for grammaticality along the following scale: *=ungrammatical, ?= awkward, ??=very awkward, ???=extremely awkward, but not conclusively ungrammatical.

In this dissertation we are concerned with answering questions such as the following:

- Are these author-generated examples representative of a larger set of English data?
- Why does the nonreflexive occur with *beside* in 1.1 while the reflexive occurs with *onto* in 1.2?
- Why would the action of placing something in front of the body (in 1.3) as opposed to inside the body (in 1.4) sound more natural with the nonreflexive pronoun?
- Why do both pronoun types occur in 1.5 and what is the difference in meaning between the use of one over the other?
- What do these examples tell us about anaphoric distribution in other syntactic contexts?
- And, finally, how can these examples inform our theory of anaphora, specifically, and our theory of grammar, in general?

The essence of the story to be told in this thesis is that a detailed exploration of anaphoric distribution in the PP offers a new perspective from which we can explain anaphoric distribution in general.

In the quote that introduces this dissertation, Ron Langacker facetiously refers linguists who are interested in the details of language data as bad theorists. The take-home message of the quote is not that one should ignore the less frequent data in order to be a good theorist, but rather the contrary, the less frequent data often provide insight into the inner workings of the more frequent data. In the same vein, Fauconnier (1997: 33) argues that complex data, often labeled as "peripheral", "rare", or "special", must be understood in order to understand the general "well-behaved" cases of the particular phenomenon. He likens this methodological approach to approaches taken in modern chemistry in which the observation of rare elements is used to theorize about the structure of common elements. He explains, "Statistically infrequent examples

produced in statistically infrequent circumstances are apt to shed light on the basic mechanisms of everyday thinking and talking (p.33)."

Within the literature on anaphora, the PP domain, along with other syntactically peripheral domains such as *by*-phrases and emphatic insertion (*John, himself, wants to go*), is certainly treated as peripheral. Longer analyses on the topic dedicate only small sections of the analysis, if any, to this domain. However, in actual frequency of usage, the reflexive pronoun, used to indicate a reflexive event in a simple syntactic clause structure (e.g. *John kicked himself*), accounts for only 10-15% of all its uses in both child and adult English (Stojanovic 2003: 316). That is, the reflexive is used much more often in emphatic expressions, *by*-phrases, and in the PP together than in a syntactically simple reflexive structure. The fact that the reflexive pronoun exists in multiple syntactic configurations warrants the possibility that the semantic factors that influence its appearance in these structures also play a role it its appearance in basic clause structure.

Based on English and Spanish corpus evidence, I will argue that a detailed understanding of the spatial relations among sentence participants is crucial in order to model the relevant grammatical principles, which account for the real-world examples represented in the corpus data. The detailed parallelism of this phenomenon in the two languages point to linguistic universals, not based upon idiosyncratic grammatical rules, but rather universals grounded in human cognition and the interpretation of dynamic spatial relationships. The findings of this thesis produce important consequences for the theory of how pronouns are produced and understood in English, Spanish, and language in general.

In Chapter 2, I introduce a brief history of existing theoretical models of anaphora. In verbal object position, English reflexive and nonreflexive pronouns exhibit complementary distribution as shown in 1.6:

(1.6) *Peter$_j$ said that John$_i$ loves himself$_{i/*j}$/him$_{*i/j}$.*

This grammatical property has been modeled syntactically based on the structural position the pronoun's referent holds in relation to the pronoun itself within a specified syntactic domain (as done by a large number of researchers including Chomsky (1995) and Pollard and Sag (1992)). The structural model of anaphoric distribution, however, becomes problematic when faced with a syntactic environment in which both pronouns appear to be grammatical as in the data in 1.1-1.5, in which both pronouns show up in the PP. In 1.5, for example, either pronoun is grammatical, and thus apparently in violation of the conditions of Binding Theory exhibited in 1.6. Examples such as 1.1-1.5 have lead researchers like Hestvik (1991) and Reinhart and Reuland (1993) to propose that the prepositional phrase itself potentially constitutes a minimal domain, shielding the pronoun from the binding conditions imposed on it in verbal object position. We will delve into these differing proposals in detail throughout Chapter 2 and finish the chapter by presenting challenges for these models. These challenges include the fact that existing syntactic models do not explain what linguistic factors dictate the choice of pronoun for an individual example, and these models raise concerns like the fact that they must ignore semantic differences among the prepositions and treat all spatial terms as part of one syntactic category.

In Chapter 3, I introduce coded corpus data and discuss the methodology used to build the corpus. The corpus data allows us to look at distributional trends across a large body of evidence to make more informed conclusions about how the pronouns are used in the PP, and the corpus data itself offsets concerns one may have about the evidential validity of author-generated data.

In Chapter 4, based on a collection of over 10,000 examples in the *British National Corpus* and over 3,000 examples of Spanish data taken from the online corpus *Real Academia Española*, I show that the distribution of reflexive versus nonreflexive pronoun is sensitive to the spatial semantics denoted by the individual preposition and its interaction with the described event. I will claim that several spatial factors reliably influence anaphoric distribution: directionality, containment, and proximity to the body. The reflexive pronoun has a higher rate of occurrence when the performed action is directed toward the referent as in 1.7, when the performed action is metaphorically predicated inside the referent as in 1.8, and when the action occurs close to the referent's body as in 1.9:

(1.7) a. *He$_i$ could have stuck pins into himself$_i$ and it would have taken ten seconds for his body to complain.* (FSP 2109)

b. *La anciana$_i$ atrajo hacia sí$_i$ al hijo que idolatraba …*(*CREA*)

The old woman$_i$ pulled the son that she idolized toward herself$_i$ …

(1.8) a. *For a moment Tuan Ti Fo$_i$ closed his eyes, PRO$_i$ seeking that inner stillness deep within himself$_i$, his lips forming the chen yen -- the `true words' - of the mantra.* (GUG 2486)

b. *Y (pro$_i$) sentía dentro de sí$_i$ una sublevación creciente …(CREA)*
And he$_i$ felt a growing revolt inside himself$_i$ …

(1.9) *He$_i$ spends all his time hiding, cloaking and padding things around himself$_i$.* (CH1 7617)

In Chapter 5, I supplement the spatial analysis with the examination of various factors that link the semantics of viewpoint to the distribution of anaphora. Based on previous

research showing that grammatically encoded viewpoint functions as a factor in pronoun choice, I search for viewpoint indicators in the PP data, concluding that corpus evidence supports a motivating role for viewpoint within the PP data.

In Chapter 6, the syntactic repercussion of our finding in Chapters 4 and 5 are explored. In this final chapter we make the case against the explanatory power of Binding Theory. I hypothesize that the incorporation of internalized spatial relations into the grammar of language along with a deeper understanding of the lexical properties of pronouns explain the unique dual occurrence of both reflexive and nonreflexive in peripheral syntactic domains and hints at the viability of a non-syntactic analysis of the core anaphoric distribution shown in 1.6.

Our goal is to give a new perspective on an old topic. By showing statistical tends in corpus data and basing our arguments upon these trends, I hope to provide a convincing argument for the necessity of cognitive principles such as metaphor and schematization in grammar. Because speakers are limited by a finite set of linguistic tools, intended detailed spatial relations are often open to interpretation by the listener; I propose that pronoun choice is actually one more clue the listener has to accurately simulate what she hears. Our story places spatial relationships at the core of grammar, and I conclude that the distribution of anaphora in the PP cannot be understood from an examination of syntactic structure alone.

Chapter 2

The syntax of anaphora in PPs

2.1 Introduction

The motivation for a large-scale corpus analysis of anaphoric distribution in the PP comes from the lack of consensus on how to syntactically model this data. For most syntacticians, the distribution of anaphora in "core" (in the sense of van Hoek 1997) data is clearly syntactic. They argue this data is most easily modeled by syntactic principles. Some researchers have taken this basic assumption and concluded that the distribution of anaphora in the PP is also motivated by syntactic principles. It will be shown in this chapter that the principles of binding at work in the core data can be partially extended to the PP data; however, among those who take this approach, there is disagreement on exactly what syntactic mechanism is responsible for the PP distribution. Furthering the disagreement is a lack of consensus among researchers over grammaticality judgments within the PP data. Here we will evaluate these analyses and pose several questions that remain unanswered by the syntactic model. For purposes of exposition, I will review a standard account of binding within the core data before moving on to a review of syntactic research specific to the PP.

2.2 General overview of binding

Since the introduction of generative syntax, the distribution of English anaphora has been a central issue in the syntactic literature (Safir 2004: 1). Not only has the question of anaphora been the focus of a wide array of an extensive scholarly literature, but the patterning of anaphora in English has also been theoretically linked to other large syntactic issues such as English structures involving 'control' and 'movement' (see

Hornstein 1999 and Kayne 2002 for a full account of these connections). The following examples illustrate the complementary distribution of English anaphoric elements:

(2.1) a. *John$_i$ criticized himself$_{i/*j}$.*

b. *John$_i$ criticized him$_{j/*i}$.*

The majority of the literature on anaphora refers to the anaphoric element, *himself*, in 2.1a as an 'anaphor' and the anaphoric element in 2.1b, *him*, as a 'pronoun'. Built into the term 'anaphor' (as opposed to 'pronoun') is an assumption that that lexical item is in some way referentially dependent on an antecedent, dependent in a way that a pronoun is not (see explanation below). In most cases, in the syntactic literature this identity dependency is based on a structural, syntactic relationship. Because, in this thesis, we want to avoid built-in assumptions about identity assignment, I will avoid using the 'anaphor/pronoun' distinction and adopt a distinction with its origins in the semantic notion of reflexivity. We will refer to both *himself* and *him* as pronouns, with the former qualifying as *reflexive* and the later *nonreflexive* by default. According to Safir (2004: 185), a form is pronominal if "it consists only of a bundle of grammatical features with no idiosyncratic lexical or semantic content …" Later we will take issue with the idea that reflexive pronouns completely lack inherent semantic content, nevertheless, for descriptive purposes this definition of pronoun justifies our use of the term to describe the anaphoric elements we are interested in here without assuming a particular theory of identity assignment.

The data in 2.1 exhibit *complementarity*, complementarity refers to the following phenomenon: in direct object position, the reflexive pronoun *himself* can only be interpreted as having the same semantic value as its clause-mate coargument *John*;

whereas the nonreflexive pronoun *him* can only be interpreted as finding its reference outside of the sentence; the antecedent of *him* cannot be the coargument *John*. The following example shows that once the nonreflexive pronoun is separated from its antecedent by a clause boundary, the sentence becomes grammatical:

(2.2) *John$_i$ said $_{CP}$[that Mary liked him$_{i/j}$].*

However, a reflexive pronoun in the same position as the nonreflexive above is ungrammatical:

(2.3) **John$_i$ said $_{CP}$[that Mary liked himself$_i$].*

The traditional syntactic approach to these data is represented in Binding Theory first presented in Chomsky (1981) based on Reinhart's (1976) c-command analysis. According to Binding Theory, the complementary distribution of reflexive and nonreflexive is argued to be based on structural properties of the sentence and the position in which the pronoun occurs. These basic structural principles that guide the appearance of the reflexive versus nonreflexive pronoun in English are as follows in 2.4[2] (taken from Safir 2004: 9):

(2.4) a. Principle A: An anaphor (reflexive pronoun) must be bound in Domain D (some locally defined domain).

b. Principle B: A pronoun (nonreflexive) must be free in Domain D (some locally defined domain).

[2] The general binding conditions, outlined in other sources, include reciprocals in Principle A, but we will not be looking at the patterning of reciprocals.

Binding is defined as follows (also taken from Safir 2004: 9):

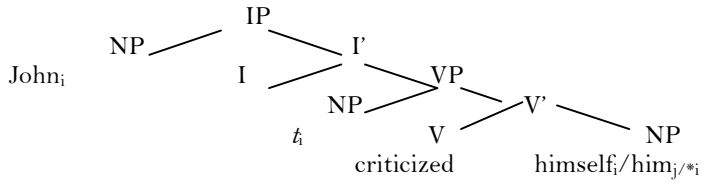
(2.5) *Binding*: X binds Y if X c-commands Y and X and Y are co indexed. If X is not bound it is free.

The structural notion of *c-command* is defined as follows (also taken from Safir 2004: 9):

(2.6) *C-command*: X c-commands Y if the first branching node dominating X also dominates Y.

The following tree structure illustrates these principles:

(2.7)



Within the minimal domain (the VP in this case)[3], the first branching node that dominates the NP *t* (the trace of *John*) is the VP node. This node, VP, also dominates the NP object node which carries the two pronouns. Thus the Specifier NP, containing the trace, *t*, c-commands the object NP position. In this case the trace of *John*, *t*, is co-indexed with the reflexive pronoun, *himself*; thus, *t* binds the reflexive pronoun, and this configuration is grammatical according to Principle A. If the nonreflexive pronoun is co-indexed with the trace of *John*, *t*, then the trace binds the nonreflexive, and this configuration is, thus, ungrammatical according to Principle B, since a nonreflexive pronoun must be free in the local domain (VP). The only grammatical use of the nonreflexive in this position is one in which the nonreflexive is co-indexed with an

[3] Chomsky (1995: 102) states, in describing binding relations in sentences such as 2.7, "Under the hypothesis that subjects are base-generated internal to VP, the VP will be the GC (Governing Category, the 'local domain'), with the trace of the subject (which has itself moved to the [Spec,IP]) serving as the binder." For this tree the binding conditions would also predict the appropriate pronoun if the IP were taken to be the minimal domain.

argument outside the VP or not co-indexed with any NP at all (this would be the case in 2.1b).

Versions of this basic binding theoretical account of the distribution of the core anaphora data have become the standard in the syntactic literature, and within syntactic theory, few if any researchers question the basic principle that anaphoric distribution is based on some type of command relationship (c-command within Chomskyan syntax; o-command[4] in Head Driven Phrase Structure Grammar (Pollard and Sag 1992)). Complementary distribution of anaphoric devices is a common assumption.

2.3 Domain D

In the examples presented thus far, the VP has served as the relevant domain for binding, Domain D. Chomsky (1995) analyzes other examples in which the local domain for Conditions A and B is the NP such as in 2.8:

(2.8) *John$_i$ likes $_{NP}$[Bill's$_j$ stories about himself$_{j/*i}$]*.

In 2.8 the NP, *Bill's stories about himself*, is the local domain; *himself* cannot refer to an antecedent outside of the NP because there is a potential antecedent, *Bill*, in this case, inside the NP. However, in some cases a reflexive pronoun resides within an NP of similar syntactic structure with no local binder as in 2.9:

(2.9) *John$_i$ likes $_{NP}$[stories about himself$_{i/*j}$]*.

[4] O-command assumes hierarchy a scale of obliqueness relations in which the less oblique argument commands the more oblique argument.

When this configuration occurs, the reflexive pronoun must find its antecedent in the subject (specifier) position of the VP not the NP as in 2.8. Example 2.9 shows that the definition of the binding domain must be flexible enough to allow for varying structures to serve as the relevant domain. The modified domain rule incorporates this distinction and generalizes from NP and VP to *complete functional complex* (CFC). A CFC is "a projection containing all grammatical functions compatible with its head (Chomsky 1995: 102)." The rule for finding the appropriate domain for binding is as follows:

(2.10) The Governing Category (the local domain) for α is the minimal CFC that contains α and a governor of α and in which α's binding condition could, in principle, be satisfied.
(Chomsky 1995: 102)

In 2.8, the CFC is the NP, *Bill's stories about himself*, since the Determiner position is filled with *Bill*, who serves as the reflexive's binder. In 2.9, the NP lacks the filled D position. The NP alone does not include a governor, thus the CFC must be the VP, the domain in which *John*, the binder, binds the reflexive pronoun.

Chomsky (1995) does not discuss in detail examples in which the CFC is the PP. Perhaps he purposefully excludes this domain from his demonstration of the binding principles since this is one domain in which syntactic complementarity has been shown to break down. However, many other syntactic theorists do specifically address this domain, so to this literature we now turn.

2.4 Binding in PPs from the syntactician's perspective

When applied to the distribution of anaphora in prepositional phrases, the standard Binding Theory fails. The following six examples (simplified versions of the general trends found in the English corpus data[5]) demonstrate the problem:

(2.11) a. *John$_i$ kept his true feelings within himself$_i$.*

b. *John$_i$ felt the anger within himself$_i$.*

c. *John$_i$ looked for answers within himself$_i$.*

(2.12) a. *John$_i$ put the book next to him$_i$.*

b. *John$_i$ offered the chair next to him$_i$.*

c. *John$_i$ looked for the scissors next to him$_i$.*

These six examples represent three different syntactic configurations. The (a) examples are ditransitive; the verb selects for both an NP object and a locative PP. The PPs in the (a) examples are, thus, syntactic complements (see discussion of Hestvik 1991 below). The (b) examples, are argued to involve PP adjuncts adjoined to the NP object; in 2.11b, one interpretation is that *within himself* identifies the location of the anger not the location of the feeling and in 2.12b, *next to him* unequivocally identifies a particular chair in the room, not the location of the offering event. In the (c) examples the PPs are adjuncts too, eliminating them from the sentence does not make the VP incomplete, nor does it change the semantics of the verb. In this case they are not adjoined within the NP. The exact node of adjunction is debatable, and we will discuss this point below; for now we will simply conclude that the PPs in the (c) examples need to be treated

[5] See Chapter 4 for details of the data.

differently from the PPs from the (a) examples. For purposes of clarity and ease of reference, we offer the following bracketing to represent the differences just described:

(2.13) a. $_{IP}[_{NP}[$ John$_i$$]$ $_{VP}[$ kept $_{NP}[$ his true feelings$]$ $_{PP}[$ within himself$_i$$]]]$

b. $_{IP}$ $[_{NP}[$ John$_i$$]$ $_{VP}[$ felt $_{NP}[$ the anger $_{PP}[$ within himself$_i$$]]]]$.

c. $_{IP}$ $[_{NP}[$ John$_i$$]$ $_{VP}[$ looked for $_{NP}[$ answers$]]$ $_{PP}[$ within himself$_i$$]]$.

(2.14) a. $_{IP}[_{NP}[$ John$_i$$]$ $_{VP}[$ put $_{NP}[$ the book$]$ $_{PP}[$ next to him$_i$$]]]$

b. $_{IP}$ $[_{NP}[$ John$_i$$]$ $_{VP}[$ offered $_{NP}[$ the chair $_{PP}[$ next to him$_i$$]]]]$.

c. $_{IP}$ $[_{NP}[$ John$_i$$]$ $_{VP}[$ looked for $_{NP}[$ the scissors$]]$ $_{PP}[$ next to him$_i$$]]$.

For purposes of exposition, we can assume that *John* is based-generated within the VP; thus, the specifier of the VP in each of these sentences is *John*. Because the PP in 2.14c is attached outside the VP, the nonreflexive in this example is free within its local domain, the IP. There is no violation of the binding principles. However *John* binds the pronoun located in the PP in 2.13a,b and 2.14a,b if the VP is considered the local domain for binding. According to Condition B, the nonreflexive pronoun must be free within its local domain, thus examples 2.14a and 2.14b should be ungrammatical with the reading that the pronoun and *John* are coreferent. Following Tenny (2003:1), we will call the nonreflexive in this position a 'short-distance' pronoun[6]. The data in 2.14 suggest, then, that the local domain for binding when the short-distance pronoun resides in a PP should not be the VP, but rather a more local domain such as the PP itself. This is in fact the analysis proposed in Hestvik 1991, Reinhart and Reuland 1993, Büring 2005, and Safir 2004. If the PP forms its own domain, then the short-distance

[6] Tenny is most-likely referencing the quasi-opposite phenomenon of 'long-distance' reflexives in the creation of this term.

pronouns in 2.14 would not be bound within the local domain and these sentences would not be in violation of Condition B.

There is an unfortunate consequence to adopting this analytic route. If the PP serves as the local domain for binding in 2.13a and 2.13b, the reflexive pronoun has no local binder unless there is some reason to treat these PPs differently from 2.14a and 2.14b. An unbound reflexive pronoun is in violation of Condition A in the standard Binding Theory. With no local binder, Chomsky's (1995) proposal would say that the CFC in these cases must be the next domain up the tree, the VP. But again, if the VP is adopted as the relevant domain for binding, the clearly grammatical examples in 2.14a and 2.14b are ruled out. This paradox signals that pronouns in prepositional phrases require either a revision to the standard Binding Theory or some type of amendment to it.

Two of the four proposals reviewed below amend the standard Binding Theory: both make a distinction between CFC governors and subjects, arguing that the subject position is the relevant binder. After making this difference explicit, they argue that reflexive and nonreflexive pronouns gain their semantic value through different procedural mechanisms. These two proposals advocate syntactic explanations for pronominal distribution in the PP. The second two of the four proposals ultimately abandon the notion that the distribution can be explained by syntactic mechanisms alone. Instead they concede that examples for which either pronoun is grammatical warrant an explanation from discourse parameters, which function supplemental to the syntax.

2.4.1 Theories of differential reference strategies between pronoun types

Working within the framework of Chomskyan Binding Theory, Hestvik (1991) alters the standard assumption that the binding domain for nonreflexive pronouns should include subjects (in the sense of the Spec IP position). He characterizes his data in three distinct groups: PPs which modify intransitive verbs as in 2.15, PPs which serve as complements of ditransitives in 2.16, and adjunct PPs as in 2.17. In the first group, Hestvik characterized the verbs as intransitive:

(2.15) a. *John$_i$ looked around him$_i$/himself$_i$.* (Hestvik 1991: 463,#10a)

b. *John$_i$ glanced behind/around him$_i$/himself$_i$.* (Hestvik 1991: 463,#10b)

c. *In the tunnel, John$_i$ searched above him$_i$/himself$_i$ and below him$_i$/himself$_i$ for an opening.* (Hestvik 1991: 463,#10c)

His ditransitive group in 2.16 below closely resembles the previous ditransitive data in 2.13a, 2.14a. These are examples in which the verb selects for a PP:

(2.16) a. *John$_i$ piled newspapers in front of him$_i$/himself$_i$.* (Hestvik 1991: 463,#11a)

b. *The host$_i$ placed Mary next to him$_i$/himself$_i$.* (Hestvik 1991: 463,#11b)

c. *John$_i$ pulled the blanket over him$_i$/himself$_i$.* (Hestvik 1991: 463,#11c)

d. *The boy$_i$ pushed the girl away from him$_i$/himself$_i$.* (Hestvik 1991: 463,#11d)

e. *John$_i$ put the picture behind him$_i$/himself$_i$.* (Hestvik 1991: 463,#11e)

f. *John$_i$ put the sword down in front of him$_i$/himself$_i$.* (Hestvik 1991: 463,#11f)

g. *The children$_i$ drew circles around them$_i$/themselves$_i$.* (Hestvik 1991: 463,#11g)

Hestvik's data with PP adjuncts in 2.17 resembles the data we discussed in 2.13c and 2.14c. In these cases the verb does not select for a PP:

(2.17) a. ??*John$_i$ found a dollar bill in front of himself$_i$.* (Hestvik 1991: 464,#16a)

b. ??*John$_i$ left Mary behind himself$_i$.* (Hestvik 1991: 464,#16b)

In Hestvik's analysis the PP, as a CFC (Complete Functional Complex), serves as a local domain. Thus, the minimal domain, which includes the pronoun and its governor, in the examples in 2.15, 2.16, and 2.17 is the PP; however the minimal domain which includes the pronoun, its governor, and a subject is IP. In example sets 2.15 and 2.16, the nonreflexive pronoun is co-indexed with the subject, but it is not necessarily bound by it; the two words share an identity –they share the same referent, but their semantic values are filled independently. If the pronoun is nonreflexive, then, based on this characterization of the CFC, it may be free, even if it is co-indexed with the subject of the sentence, so it does not violate Condition B. In these cases the nonreflexive pronoun exhibits independent coreference. For example in 2.16a, the nonreflexive *him* is free of a local binder in the PP. It is not free within the IP, but since the PP is adjoined within the IP, the nonreflexive is licensed. On the other hand, if the pronoun is reflexive, it "must meet its binding requirement in the minimal CFC containing a BT-compatible indexing (p.463)," which is IP. So, the reflexive *himself* in 2.16a is licensed because it is co-indexed with the subject *John* at the IP level. The reflexive, thus, gets its semantic value directly from the co-indexed binder. This "subjectless" Binding Theory relativized to Conditions A and B, therefore, predicts the noncomplementary distribution in 2.15 and 2.16.

Hestvik's differential treatment of the binding domain for reflexives and nonreflexives is analogous to Chomsky's (1995) treatment of pronominal distribution in complex NPs outlined in Section 2.3, examples 2.8 and 2.9[7]. In these cases the reflexive pronoun in a complex NP finds its binder at the level of the IP when there is no local

[7] This approach was suggested by Huang (1983) for NPs.

binder present. His analysis of binding in the PP mirrors Chomsky's subsequent analysis of binding in a different syntactic domain, and together the two analyses support the idea that different syntactic domains can serve as the relevant domain for the conditions on binding.

The validity of Hestvik's claim that reflexive pronouns are bound differently from nonreflexive pronoun can be evaluated in different ways. First, the difference between true binding and independent coreference suggested by Hestvik should be testable with verb-phrase ellipsis. If the verb phrase can be elided in an adjoined clause and the nonreflexive pronoun allows for a strict reading, then Hestvik's argument for a difference in binding (between the nonreflexives and reflexives in 2.15 and 2.16) is upheld since a strict reading assumes independent coreference. The examples created in 2.18 do show in fact that in some of Hestvik's examples, the nonreflexive allows for a strict reading under VP ellipsis:

(2.18) a. *In the tunnel, John$_i$ searched above him$_i$ and Bill did too.*
(Possible reading: *Bill searched above John.*)

b. *John$_i$ piled newspapers in front of him$_i$ and Bill did too.*
(Possible reading: *Bill piled newspapers in front of John.*)

Some of Hestvik's examples pass the VP ellipsis test. The examples that pass, such as those in 2.18, involve an action that is not directional. It is our observation that those examples that do not pass this test are semantically incompatible with the test. For example in 2.15a, represented below in 2.19, one cannot get the strict reading under the VP ellipsis test because looking events originate with the subject and the goal of the line of sight is predicated to be around the subject:

(2.19) *John$_i$ looked around him$_i$ and Bill did too.*

(Impossible reading: *Bill looked at the area around John*)

(Only possible reading: *Bill looked at the area around Bill.*)

The fact that the strict reading is impossible in 2.19 suggests that the semantics of the event itself limits the referential relations of the nonreflexive pronoun. That is, in an event of looking around, meaning staring at different points in space around one's body, the nonreflexive can only refer to the looker of the event.

A second issue for Hestvik is the fact that the sloppy reading is possible with the nonreflexive pronoun in 2.19. Although not addressed in his paper, Hestvik's examples seem to allow for an interpretation in which the nonreflexive pronoun is bound. In 2.18, repeated below in 2.20, the conjoined clause with the elided VP also allows not only for the strict interpretation but also for the sloppy interpretation:

(2.20) a. *In the tunnel, John$_i$ searched above him$_i$ and Bill did too.*

(Possible reading: *Bill searched above Bill.*)

b. *John$_i$ piled newspapers in front of him$_i$ and Bill did too.*

(Possible reading: *Bill piled newspapers in front of Bill.*)

Remember, according to Hestvik, nonreflexive pronouns are sensitive to the PP domain and are not bound by the subject because the minimal PP domain shields them. It appears here that Hestvik would have to concede that these interpretations require the subject to bind the nonreflexive pronoun in order to get the sloppy interpretation. Thus, in some cases the subject needs to bind the nonreflexive pronoun.

A remaining problem for Hestvik, is how he should account for the ungrammaticality of the examples in 2.17. Based on the theory he outlines, the reflexive should be grammatical in the PP because it is bound within the IP. Hestvik decides that the ungrammaticality of 2.17 is due to a fundamental distinction between adjunct PPs and complement PPs. Based on the assumption that adjunct PPs are "optional" (a term he borrows from Grimshaw (1990)), the PP in 2.17 is labeled as an adjunct. He explains that "anaphors (reflexive pronouns) have to raise at LF out of the PPs, but such movement results in ECP violations . . . The trace of the anaphor will fail to be antecedent governed, because the adjunct PP, not being theta-marked, will constitute a barrier for anything inside it (p.464)." If adjunct PPs form barriers over which reflexive pronouns cannot "find" their antecedents, then no reflexive pronoun should be grammatical in an adjunct PP. In summary, Hestvik's theory depends upon the critical division between adjunct PPs and complement PPs. Again, Hestvik would argue that complement PPs are semantically required and adjunct PPs are semantically optional.

Upon close scrutiny, Hestvik's own data raises three questions for further evaluation. First, how do the semantics of a particular event interact with pronominal reference? As in 2.19, certain event semantics seem to force particular reference relations. Second, is the domain for binding really different for reflexive and nonreflexive pronouns? And finally, how important is a clear distinction between PP adjuncts and PP complements for an accurate model of anaphoric distribution?

Büring (2005) tackles one of these issues by further clarifying both the difference and similarity of PP adjuncts and PP complements, and he highlights their differential role in the election of what constitutes Domain D. He characterizes path and place PPs in certain contexts as complement PPs. In these cases such as *look behind NP*, *pull it*

around NP, and *hide it next to NP* (p.71), following Reinhart and Reuland (1993) (discussed in the next section), he theorizes that the PP serves as the co-argument domain for the NP object. When the nonreflexive pronoun is grammatical in cases such as 2.27, he hypothesizes that it receives its theta-role directly from the preposition (the PP gets its theta-role from the verb) and is consequently grammatical:

(2.27) *John$_i$ looked around himself$_i$/him$_i$.* (Büring 2005: 71 #3.27b)

In 2.27, *him* receives its theta role from *around*, which receives its theta role from *look*. Although he invokes the theory of theta role assignment, he essentially argues the same point as Hestvik. Nonreflexive pronouns are sensitive to the PP domain, and the PP domain is the relevant domain concerning binding relations. The fact that the nonreflexive pronoun's semantic value is the same as the subject's is coincidental. That is, the nonreflexive pronoun can be coreferential to the subject, but, according to him, that reference is not gained through binding. As for the reflexive pronoun, Büring simply restates Hesvik's principles while incorporating case theory. He states that the reflexive pronoun "must be bound within the smallest category containing it, its case assigner (the preposition in these examples), and a Subject (its GC) (p.70)"; thus the reflexive is also grammatical in 2.27 because it is bound at the IP level –there is a subject, *John*, to bind it. Unlike Hestvik, however, Büring argues that adjunct PPs function in the same way as complement PPs; they allow either reflexive or nonreflexive as in 2.28 for the reasons summarized above:

(2.28) *John$_i$ saw a cat next to him$_i$/himself$_i$.*

Büring does not exclude the reflexive pronoun from appearing in all PP adjuncts like Hestvik does. Büring's contribution is to incorporate a new, semantic-based explanation for why the PP serves as its own domain in the case the nonreflexive, but he doesn't address details about the role semantics plays in the awkwardness of Hestvik's PP adjunct examples in 2.17.

The examples in 2.20, which suggest, by VP ellipsis test, that the nonreflexive can be bound, seem to be a problem for Büring for the same reasons they are a problem for Hestvik.

A final point to make when evaluating these two syntactic models of anaphoric distribution is the fact that neither researcher discusses the disproportional frequency of one pronoun over the other in the contexts which allow both. A question arises as to whether or not it is within the domain of syntax to account for distributional preferences. An answer to this question will be saved for Chapter 6. The next researchers to be discussed do not address distributional preferences, but they do question how much of anaphora in the PP can be modeled with only a syntactic apparatus. They advocate for a distinction between syntactic and semantic domains in relation to binding.

2.4.2 The PP as a semantic domain

Hestvik and Büring are not the only researchers to propose that the PP forms its own local domain. Reinhart and Reuland (1993) also argue that the PP forms its own domain; however, in their analysis this domain is semantic not syntactic:

> *What P and N predicates have in common is that the predicates they form are not syntactic . . . Hence, we find a different pattern of anaphora here than in the case of V predicates. On this analysis, all predicative heads form semantic predicates, but only a subset also form syntactic predicates* (p.690).

In order to understand the division created by Reinhart and Reuland, one must look at their definition of what constitutes a syntactic predicate versus a semantic predicate. They state:

(2.21) a. The *syntactic predicate* formed of (a head) **P**(redicate) is P, all its syntactic arguments, and an external argument of P (subject).
The *syntactic arguments* of P are the projections assigned theta-role or Case by P.

b. The semantic predicate formed of P is P and all its arguments at the relevant semantic level.

c. A predicate is *reflexive* iff two of its arguments are co indexed.

d. A predicate (formed of P) is *reflexive-marked* iff either P is lexically reflexive or one of P's arguments is a SELF-anaphor.

(Reinhart & Reuland 1993: 678 #40)

Based on these new definitions, Reinhart and Reuland make a revision of Condition A and Condition B:

(2.22) A: A reflexive-marked syntactic predicate is reflexive.

B: A reflexive semantic predicate is reflexive-marked.

(Reinhart & Reuland 1993: 678 #41)

Reinhart and Reuland propose that in locative and directional PPs, the preposition forms its own syntactic and semantic predicate; the NP complement of the preposition is an argument of the preposition rather than of the verb. They list the following examples to support this idea:

(2.23) a. *Max$_i$ saw a ghost next to him$_i$ /himself$_i$.*

b. *Max$_i$ put the book next to him$_i$ /himself$_i$.*

c. *Max$_i$ pulled the cart toward him$_i$ /himself$_i$.*

(Reinhart & Reuland 1993: 686 #59)

In 2.23, the pronouns in the P(reposition) predicate are free from the domain of binding. That is, since the V does not select for the NP argument of the P predicate, the V-predicate is irrelevant for Condition B. Also within the PPs in 2.23, there is no subject, thus there is no syntactic predicate either, so Condition A does not apply. They state that the reflexive pronoun in these cases is perhaps more marked and that it demonstrates logophoric properties:

> In the context where the syntax allows both a pronoun and a SELF anaphor to be co indexed with a given antecedent, the choice between them is motivated by discourse considerations . . . For this reason, the use of an anaphor in such contexts may appear more marked than in the reflexivity environments, where the anaphor is the only grammatical option. There is no reason to assume that this type of discourse consideration is encoded in the syntax (p. 672).

Turning to a slightly different type of construction, Reinhart and Reuland show that some PPs form semantic predicates which are semantically reflexive. This is the case in 2.24:

(2.24) a. **Max rolled the carpet$_i$ over it$_i$.*

b. *Max rolled the carpet$_i$ over itself$_i$.* (Reinhart & Reuland 1993: 689 #67)

Since this P predicate, *over*, is not a syntactic predicate, Condition A is irrelevant, but because it is a semantic predicate Condition B does apply. Condition B allows the reflexive-marked predicate in 2.24b but rules out the unmarked predicate in 2.24a.

In summary, Reinhart and Reuland argue that the only syntactic binding relevant to PPs is when the NP in the PP is selected for by the V as in 2.25. In cases

such as 2.25, the preposition forms a semantic unit with the verb, unlike the examples in 2.23[8]. It should be noted that these prepositions are not semantically locative:

(2.25) a. *Lucie$_i$ explained Max to *her$_i$/herself$_i$.*

b. *Lucie explained Max$_i$ to *him$_i$/himself$_i$.* (Reinhart & Reuland 1993: 687 #58)

According to Reinhart and Reuland, the preposition *to*, in 2.25, is selected for by the verb *explain*, thus allowing the subject to bind the reflexive, and the nonreflexive in this position is ruled out by their Condition B since *him* does not mark the predicate as reflexive. In support of Reinhart and Reuland, Büring argues that PPs in phrasal verbs as well, such as *send to* or *rely on*, are not co argument domains for the NP object because the preposition forms a semantic unit with the verb; thus the verb selects for the NP object of the "semantically-inert" (p.301) preposition. Thus, the reflexive in 2.26 receives its theta role from the verb not the preposition, hence disallowing the nonreflexive pronoun:

(2.26) *John$_i$ relies on himself$_i$/*him$_i$.* (Büring 2005: 71 #3.27a)

Unlike Büring, however, Reinhart and Reuland's full analysis abandons a syntax-driven account of the distribution of anaphora and relies on the semantic structure of the PP and its relation to the VP to account for the data they use. In this way their account also differs from Hestvik 1991. Reinhart and Reuland would say that the exact mechanism for determining whether a reflexive or nonreflexive pronoun

[8] This semantic unit is similar to a verb-particle (phrasal verb) construction except in this case, the word order is fixed.

should go into a locative PP is not a problem for the syntax to handle -the syntax allows either- this problem is best handled in discourse theory.

Safir (2004) revisits examples analyzed in Reinhart and Reuland (1993). He explains that there is a class of PPs which permit either the pronoun (nonreflexive) or pronoun-SELF (reflexive) to appear in English. Safir's complete approach to anaphoric distribution is based on a theory of competition and will be revisited in Chapter 6. To briefly summarize here, in Safir's theory, anaphoric forms in English (and in every language with an anaphoric system) differ in their position on a scale of dependency; that is, to a given extent, each anaphor depends on its referent to fill its semantic value: nonreflexive pronouns are more dependent forms than epithets or proper names and SELF forms (reflexives) in English are more dependent than simple nonreflexive pronouns. The selection of the appropriate form is based on the standard binding configuration of c-command but adds a notion of competition; thus, "if x c-commands y and y is not the most dependent form available, then co-reference between x and y is not expected (p.50)." In regard to anaphora in the PP he claims that many observations made by Reinhart and Reuland can translate into his approach.

Safir, agreeing with Reinhart and Reuland, explains that in locative and directional PPs, the noncomplementary distribution of pronouns could be due to the possibility that, in these cases, the reflexive pronoun is promoted to a discourse dependent anaphor which has the same status on the scale of dependency as do nonreflexive pronouns.[9] His formal statement is repeated in 2.29:

[9] Pollard and Sag call these reflexive forms in unexpected contexts as "exempt anaphors" (i.e. exempt from Condition A).

(2.29) Promotion of an Anaphor to Discourse-Sensitive UD-Form: An anaphor can be promoted to a discourse-sensitive dependent if it does not participate in a complete thematic complex[10].
(Safir 2004: 179 #61)

Safir, like Reinhart and Reuland, abandons a syntactic explanation of the distribution of anaphora in locative and directional PPs. That is, in his theory, either pronoun is grammatical in these cases, and the appearance of one versus the other is based on discourse mechanisms, which should not be modeled in the syntax.

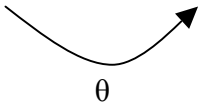
2.5 Summary

We have now seen four related accounts of how to capture the distribution of anaphora in PPs. All researchers in this section agree that "semantically inert" (Büring's characterization) prepositions, which form the second part of a phrasal verb, do not affect the syntax of anaphoric distribution in the PP. In other words, prepositions such as *to* and *on* in *send to* or *rely on* are invisible to binding principles; they function as part of a verb complex, and as such, do not head a separate PP domain to which the syntax is sensitive. Phrasal verbs pattern just like regular verb in regard to the binding principles. Looking ahead to the BNC corpus data, this prediction is in fact born out in the real-world examples.

Hestvik and Büring stay true to a syntactic explanation of the data, however they provide differing grammaticality judgments. In regard to locative and directional PPs, both researchers agree that complement PPs (which receive a theta role from the verb and are syntactically obligatory) form their own syntactic domain; thus, allowing the coreferential nonreflexive pronoun to appear as in 2.16e repeated here in 2.30:

[10] By 'complete thematic complex' Safir assumes that each potential domain for binding must express all possible theta-roles in order to be 'complete'. He explains, "The idea . . . is that picture nominals do not express all of the potential arguments of the picture nominal, and so picture nominals, without pronominal genitives are not complete thematic complexes (p. 179)."

(2.30) John$_i$ put the picture $_{PP}$[behind him$_i$/himself$_i$]. (Hestvik 1991: 463,#11e)



They also agree that the reflexive pronoun in 2.30 is grammatical and that it is licensed because, in this case, it finds a co-indexed binder, *John*, within the minimal IP. Hestvik and Büring disagree on the patterning of anaphora in adjunct PPs (which do not receive a theta role from the verb and are not syntactically obligatory) as in 2.17 and 2.28 repeated here in 2.31:

(2.31) a. *John$_i$ saw a cat next to him$_i$/himself$_i$.* (Büring 2005: 71 #3.27a)

b. *??John$_i$ found a dollar bill in front of himself$_i$.* (Hestvik 1991: 464 #16a)

c. *??John$_i$ left Mary behind himself$_i$.* (Hestvik 1991: 464 #16b)

Büring treats locative adjunct PPs in the same manner as locative complement PPs: the coreferential nonreflexive is licensed because the PP forms its own domain, and the reflexive is licensed because it has a co-indexed binder in the minimal IP. Hestvik, on the other hand, disallows the reflexive in sentences such as those in 2.31. He claims that PP adjuncts form barriers across which the reflexive pronoun cannot find its governing antecedent. Hestvik's model crucially relies on a strict division between PP complements and PP adjuncts. Foreshadowing the corpus data, this strict division will become problematic.

The other researchers, Reinhart and Reuland and Safir, conclude that the syntax should allow either reflexive or nonreflexive in locative PPs of both kinds (complement

and adjunct), and the appearance of one versus the other is due to discourse factors which should not be modeled in the syntax.

Beyond a lack of consensus about how to accurately model the PP data, these four researchers leave many questions unanswered. First, it is unclear what the real-world data look like. Does speaker-generated data follow the same grammaticality judgments given by the researchers? Second, if discourse factors are responsible for the distributional pattern, then what are they? And finally, there is no explicit, detailed investigation from any of the six researchers profiled in this chapter as to why the syntactic apparatus so cleanly models the core data but not the PP data. Why should the syntactic apparatus be responsible for anaphoric distribution in simple clauses while semantic factors are at least partially, if not exclusively, responsible for the distribution of anaphora in the PP? These questions will be address in the subsequent analysis.

In this chapter, it has been shown that even the most syntactically faithful analyses of anaphoric distribution in the PP must rely on semantic principles to create categorical distinctions to model the data. Categories like 'phrasal verb', 'PP complement', and 'PP adjunct' have fuzzy boundaries that complicate their role as predictive forces for anaphoric distribution. Some sort of linguistic principles are guiding anaphoric distribution in PPs, otherwise there would be no preference for one anaphor over the other in this syntactic context. The goal of the rest of this thesis is to embrace the complications outlined above by shifting the focus of the analysis toward more detailed semantic distinctions brought up in the data. There are two specific goals for the ensuing corpus study. First, the corpus analysis allows for quantitative data. With a large corpus, distributional counts and probabilities can be examined. Along any proposed parameter, it can be shown that one pronoun occurs more frequently than

another. This type of 'macro-analysis' (to be explained below) is impossible when one relies only on author-generated data in limited quantities. Secondly, a large corpus allows the data to be coded by different semantic parameters that incorporate more spatial details. Data from the syntactic literature is limited in breadth. As mentioned in this chapter, historically, one set of syntacticians like Büring has looked at more general categories of PPs like PPs of 'location' and 'direction/path'. Thus, not only have they grouped the data into very broad semantic categories, which may not be the most useful categories for analysis, but also they often work with a limited number of examples from within those categories. Another set of syntacticians have lumped all prepositions together into one analysis: those that are more functional, more polysemous, and arguably less locative, like *to*, with those that are less functional, less polysemous, and primarily locative, like *beneath*. The danger in these analyses is the assumption that all prepositions, regardless of function type, interact with anaphora in the same way. Thus, this thesis offers a new approach. In the new approach the corpus data helps drive the analysis rather than the analysis guiding the data. The hope in the new approach is that a closer examination of semantic factors will reveal why the syntactic factors outlined above have received so much attention.

Chapter 3

The Data

3.1 Introduction

Research outlined in the previous chapter references various syntactic and/or semantic factors that can contribute to anaphoric distribution in the prepositional phrase. One such factor outlined was the PP's status in regard to an oppositional scale of complement vs. adjunct. MacDonald (2004), among others, argues that the semantics of the preposition can influence the PP's status on this scale: specifically he shows that PPs that denote directionality, Path PPs, tend to be complement-like, patterning differently from non-directional, locative PPs, which pattern more like adjuncts. However, because English spatial prepositions form a closed class of functional heads, the syntactic-based research on the internal semantics of this class has been limited in regard to its affect on anaphoric distribution. Most syntactic theories (with a few notable exceptions to be discussed in Section 4.5) do not model the internal semantics of the P head – placing all prepositions in the same category and ascribing differences among them to external factors such as their node of attachment in the syntactic tree or their functional characteristics in regard to what complements they can or cannot take.

Although the internal semantics of the preposition are not attributed to the syntactic apparatus, some syntactic researchers apart from those outlined in Chapter 2 have noted the semantic effects of the individual preposition on the distribution of anaphora. For example, Wechsler (1997: 17), working in a lexicalist framework, noted that English directional prepositions pattern differently from locative prepositions with regard to anaphoric distribution. Sentences 3.1a and 3.2a show a directional preposition

occurring with the reflexive pronoun, and sentences 3.1b and 3.2b show a locative, nondirectional preposition occurring with the nonreflexive pronoun:

(3.1) a. *Bubba$_i$ tossed the beer can* to **him$_i$ / himself$_i$.*

b. *Bubba$_i$ tossed the beer can* behind *him$_i$ / *himself$_i$.* (Wechsler 1997: 17 #38)

(3.2) a. *Corporal Crump$_i$ pinned the medal* onto **him$_i$ / himself$_i$.*

b. *Corporal Crump$_i$ pinned the medal* beside *him$_i$ / *himself$_i$. (on the wall).*

(Wechsler 1997: 17 #38)

These minimal pairs indicate that directional prepositions, specifically those which form part of a self-directed action such as *to* and *onto*, pattern with the reflexive pronoun; whereas prepositions that do not lexically imply directionality, such as *behind* and *beside*, pattern with the nonreflexive pronoun.

Idiosyncratic data found in the semantic literature on anaphora suggest a second semantic parameter that may influence anaphoric distribution. When used in contexts that require anaphoric reference, prepositions that denote containment also appear to disproportionately pattern with the reflexive pronoun, as found in Lakoff (1996) in a paper written about metaphors for the self:

(3.3) *You need to step outside yourself.* (Lakoff 1996: 99 # 17)

These cases involve a metaphorical understanding of the body and mind (to be discussed below in section 4.3) and isolate another semantic category of prepositions that may influence anaphoric distribution.

Taking into account these clues buried in the literature on anaphora, the goal of this part of the thesis is to use a substantial body of corpus data to compare how prepositions, representing different semantic classes, pattern in regard to anaphoric distribution. By subjecting a large quantity of data to simplified categorical comparisons, several semantic dimensions will be highlighted. These factors play an important role in determining which anaphor is more likely to appear in an individual prepositional phrase. The spatial semantic factors heretofore mentioned have only been speculated on in short passages by just a few authors; thus a sizable corpus analysis is warranted not only to further investigate the topic from a macro-semantic lens, but also to complement the limited set of conflicting author-generated examples in the literature.

This chapter outlines the overall methodology used in the corpus collection process. The results of the analysis are saved for Chapters 4 and 5 in order to integrate them into the relevant semantic theories.

3.2 Methodology

Although vast amounts of literature have been written on the topic of anaphora, few researchers have drawn their conclusions from corpus data, and even fewer have drawn their conclusions from a statistical comparison of large quantities of data. This section details the process used to compile the 1,667 example English corpus and 386 example Spanish corpus.

In corpus linguistics, the process of corpus building falls under the rubric of a quantitative approach, or "macroscopic" analysis (Biber 1988: 61), in which exceptional data is statistically disregarded as random exception and not factored into the overall

analysis. The focus in this method is on the overall variation found in the distribution of reflexive and nonreflexive pronouns. Generally speaking, the advantage of this type of analysis is the ability to identify and give explanatory power to semantic trends in the data, which may otherwise go unnoticed at the level of the individual example. Nevertheless, semanticists often criticize this process, arguing that the semantic detail of the individual example necessitates introspection on the part of the researcher. Likewise, there is a danger in ignoring infrequent data and throwing them into the black box of "outliers", since they have obviously been uttered by someone and cannot be assumed to be ungrammatical or unimportant.

Thus, in order to complement the macroscopic analysis, the present study includes the analyses of individual examples. This method is called a qualitative approach, a "microscopic" analysis (Biber 1988: 119), in which the data are described and conclusions are drawn from the description. These two methodologies compete in the literature on corpus linguistics. McEnery and Wilson (1996: 62) summarize:

> In qualitative analysis, rare phenomena receive . . . the same attention as more frequent phenomena, and because the aim is complete detailed description rather than quantification, delicate variation in the data is foregrounded: qualitative analysis enables very fine distinctions to be drawn … The fact that qualitative analysis is not primarily classificatory also mean that the ambiguity which is inherent in human language – not only by accident but also through the deliberate intent of language users - can be fully recognized in the analysis… In contrast to qualitative analysis, the quantitative analysis of a sampled corpus does allow for its findings to be generalized to a larger population, and, furthermore, it means that direct comparisons may be made between different corpora … Quantitative analysis … enables one to discover which phenomena are likely to be genuine reflections of the behavior of a language …and which are merely chance occurrences."

It is quite apparent from this description of the two different analytical methods that both routes have different virtues warranting the chosen "multi-method approach" (McEnery and Wilson 1996: 62), which consists of presenting the data both

quantitatively and qualitatively. The statistical analysis shows what semantic factors are significantly dictating the distribution and should thus generalize to the grammars of English and Spanish, or more broadly to the human cognitive system. However, the qualitative analysis ensures that every potential hypothesis relevant to the data is considered.

The use of corpus data does not negate the value of author-generated data and author-produced judgments of grammaticality (see Fillmore 1992 for a humorous discussion of the two approaches). Often the process of constructing artificial data (as opposed to discovering naturally occurring data) helps the researcher tease apart semantic nuances that are impossible to recognize if one uses only limited sets of data from corpora. Thus, each chapter of analysis in this thesis will include a discussion of both author-generated data and insight as well as corpus data. The analytic strategy of the subsequent chapters is to help explain the semantic insight and intuition of previous researchers with semantic trends found in the corpora. Hence, the corpus data can help us figure out why certain sentences sound more 'natural' with one pronoun over the other.

3.2.1 Quantitative Methods

Both the online versions of the British National Corpus (BNC) and the Corpus Real Academia Española (version Corpus de Referencia Español Actual (CREA)) allow for rapid searches of the entire corpus by entering a word or word-strings into the search field of the main search webpage (see Appendix 1 for detailed description of both corpora's size and content). Prepositional phrases headed by the following seventeen prepositions were searched for in the BNC: *next to, beneath, behind, in front of, before, below, above, beyond, on top of, on, toward, around, out of, into, in, inside, and within.* Prepositional

phrases headed by the following nine prepositions were searched for in the CREA: *através de 'through'*, *al lado de 'next to'*, *alrededor de 'around'*, *debajo de 'under'*, *delante de 'in front of'*, *dentro de 'inside'*, *detrás de 'behind'*, *encima de 'on top of'*, *hacia 'toward'*. The impetus for this specific set of seventeen prepositions in English and nine in Spanish has two explanations. First, the set needed to be large enough for internal comparison across semantic type (see Chapter 4); that is, the smaller the data set is, fewer are the conclusions that can be inferred. Second, all the prepositions in this set have at least one sense which is overtly spatial in nature. Unlike the prepositions *to* and *of*, for example, which fall at the functional end of the semantic spectrum (and are often characterized as semantically inert (see Büring 2005)), the selected prepositions participate frequently in the description of spatial relations among event participants. The English prepositions were paired with either the 3rd person reflexive or nonreflexive pronoun (*himself* or *him*) or the first person reflexive or nonreflexive pronoun (*myself* or *me*). In short, the reason for searching in both person categories was not only to bolster the total corpus count, but, secondarily, to look for frequency differences between the first and third person contexts. To be discussed in detail in Chapter 5, the different person categories might have aided in explaining how perspective marking plays a role in anaphoric distribution. The Spanish prepositions were paired only with the third person nonreflexive pronoun *él 'him, it'* and the third person reflexive pronoun *sí 'himself, itself'*. Spanish prepositional pronouns have no form distinction between the first person nonreflexive and reflexive forms, thus making a categorical distinction impossible. The following table lists all 68 search strings in English and 18 in Spanish.

Third Person		***First Person***	
next to him	*next to himself*	*next to me*	*next to myself*
beneath him	*beneath himself*	*beneath me*	*beneath myself*
behind him	*behind himself*	*behind me*	*behind myself*
in front of him	*in front of himself*	*in front of me*	*in front of myself*
before him	*before himself*	*before me*	*before myself*
below him	*below himself*	*below me*	*below myself*
above him	*above himself*	*above me*	*above myself*
beyond him	*beyond himself*	*beyond me*	*beyond myself*
on top of him	*on top of himself*	*on top of me*	*on top of myself*
on him	*on himself*	*on me*	*on myself*
toward him	*toward himself*	*toward me*	*toward myself*
around him	*around himself*	*around me*	*around myself*
out of him	*out of himself*	*out of me*	*out of myself*
into him	*into himself*	*into me*	*into myself*
in him	*in himself*	*in me*	*in myself*
inside him	*inside himself*	*inside me*	*inside myself*
within him	*within himself*	*within me*	*within myself*
através de él	*através de sí*		
al lado de él	*al lado de sí*		
alrededor de él	*alrededor de sí*		
debajo de él	*debajo de sí*		
delante de él	*delante de sí*		
dentro de él	*dentro de si*		
detrás de él	*detrás de sí*		
encima de él	*encima de sí*		
hacia él	*hacia sí*		

Table 3.1 All search strings inputted into BNC and CREA search engines

In order to collect each preposition-pronoun combination included in the English and Spanish corpora, the exact word-string was typed into the search field, and results appeared as a list of all the examples in the corpus that contained the particular preposition-pronoun combination (see Appendix 2 for all English and Spanish search strings and hits). For example, *around him* was put into the search field of the BNC and all 922 sentences containing *around him* are listed in by the search program; or *dentro de él* 'inside him/it' is put into the search field of the CREA and all 375 paragraphs containing *dentro de él* are listed by the corpus' search program.

Even though the initial BNC examples are sentence-length, paragraph-length context can be accessed if necessary. It is necessary to access more context for examples in which co reference is unclear within the sentence alone. For example, in 3.4, it is initially unclear whether the nonreflexive pronoun *him* refers to the subject, Charles, or to another person:

(3.4) *Charles drew the bedclothes around him.* (ACE 3540)

Accessing a larger chunk of textual context reveals that the nonreflexive pronoun is in fact co referent with the subject:

3535 Officers sat on the balustrade with their brandy and cigars after dinner, watching the anti-aircraft batteries around Lille piercing the night sky with brilliantly coloured tracer shells.

3536 "Sir! sir! 3537 Wake up, sir!"

3538 "Go away. 3539 It's the middle of the night." 3540 Charles drew the bedclothes around him. 3541 He could hear the dawn chorus in full concert, and when he opened his eyes unwillingly light was filtering faintly through the curtains. 3542 "What time is it?"

3543 Just on five, sir, and the war's begun.

3544 "What?" 3545 He was on his elbows.

The scene described in the paragraph of text is compatible with a semantic reading that *him* refers to Charles, not to a different referent, especially since no such possible referent is mentioned in the larger chunk of text. [11]

3.2.2 Coding

After building the corpora in the two languages, the data was coded using a system of manual tagging (Meyer 2002:111). "Manual tagging" refers to a process of systematically labeling and categorizing each example in the corpus for specific categories decided upon by the researcher. The same system was used for both the English corpus and the Spanish corpus.

Before tagging the data into different semantic and/or syntactic categories, the data was filtered to include only that data relevant to the research questions: cases where the coreferential pronoun refers back to the subject of the clause in which it is contained. The corpus data was limited in this way so that the constructed corpus would most replicate the syntactic configurations discussed in Chapter 2. In the English data, quite frequently the subject of the clause is not overt in the text but rather identified in a preceding clause; for coding purposes only, the convention PRO was used in this scenario. PRO represents a phonologically null pronoun; thus, in the data, PRO was inserted before overtly subjectless verbs for ease of reference in coding coreferential relationships among pronouns[12]. Following standard syntactic practice, subscripted letters were used to mark two clause participants as either coreferential or

[11] The Spanish Corpus data comes already in a paragraph format.
[12] I do not advocate the idea that PRO is a psychologically real, phonologically null pronoun. I use it only for ease of reference identification in the written corpus material.

noncoreferential. The following example shows indices marking as well as the marking of PRO:

(3.5) *He$_i$ failed to notice the borrowed things that the girls wore, **PRO**$_i$ looking around him$_i$ instead in dumb bafflement: it was a wedding day, a shining moment in his life, and, except for the dressed children, it could be any ordinary day.* (A6N 1071)

In this example, the subject of *looking* is not overtly stated in the text; thus, 'PRO' represents the subject and is marked with the subscripted indices to clarify the referential relationship it shares with the pronoun in the prepositional phrase. The judgments were supported by the semantics of the scene described in the text. Sentences in which the pronoun's reference remained ambiguous, even when surrounding text was consulted, were coded as such and eliminated from the analysis.

Very frequently in Spanish, the subject of the morphologically inflected verb is also not overtly represented. Following the syntactic literature, again for coding purposes only, this type of so-called phonologically null subject is represented with little 'pro', and the coreferential relationship is marked with subscripted letters as shown in the following example[13]:

(3.6) *Y la **pro** agarra a Nati, la **pro**$_i$ pone delante de **él**$_i$ y usa su brazo como espada …(CREA)*

And, he grabs Nati, he$_i$ puts her in front of him$_i$, and he uses his arm like a sword …

It should be noted that "big PRO" subjects also exist in Spanish when a verb that is not inflected for person (i.e. gerundial and infinitival forms) appears in the text and the subject of that verb is inferred from preceding context. These subjects as well were marked by PRO.

[13] As stated for PRO, I do not advocate the notion that *pro* is a psychologically real, phonologically null pronoun. My use of it is for notational convenience only.

Because only examples in which the subject of the clause is coreferential with the pronoun in the PP are of interest, examples such as 4.7 and 4.8 were eliminated from the two corpora:

(4.7) *I$_j$ shut the door behind him$_i$ and locked it.* (A08 3108)

(4.8) *Bueno, yo$_j$ he estado sentado al lado de él$_i$ muchas veces…* (CREA)

Well, I$_j$ have been seated next to him$_i$ many times…

In these examples, the nonreflexive pronoun is not coreferential with the clausemate subject. These examples represent the typical use of the nonreflexive pronoun to refer to a scene participant outside of the clause. In all remaining examples, the pronoun, either reflexive or nonreflexive, in the event-modifying PP, refers back to the subject of the clause, with four categorical exceptions to this norm. These four categories were coded for, but excluded from the macro-semantic analysis because each, for differing reasons, does not represent the relevant construction: a complex clause with a pronoun (coreferential to the subject) in a PP that modifies the event, a PP that forms part of the predicate.

First, "long-distance" reflexives in PPs were eliminated; these are constructions in which the reflexive pronoun does not have a clausemate antecedent as in 3.9 and 3.10:

(3.9) *Of all George$_i$ went through in his Beatle days, there was this kind of commitment within himself$_i$ to say "I really want to see you*".(CBC 7906)

(3.10) *Con el caballo sigiéndolo unos pasos detrás de sí$_i$, pro$_i$ se dirigió hacia la tumba.* (CREA)

With the horse following a few feet behind himself$_i$, he$_i$ turned toward the grave.

These examples were coded "LD" for reference in the subsequent analysis, and excluded from the macro-analysis performed on the within-clause examples of coreference.

Secondly, reflexive pronouns whose antecedent is the object of the verb as in 3.11 and 3.12 were eliminated from the corpus:

(3.11) *We are counting on you to take him$_i$ out of himself$_i$.* (FU6 1414)

(3.12) *pro$_j$ acosó a la Nana$_i$ hasta que la$_i$ puso fuera de sí$_i$…* (CREA)
(He$_j$) harassed Nanai until (he$_j$) made her$_i$ outside of herself$_i$…
He harassed Nana until he made her crazy …

These examples were coded "D.O.", and they too were eliminated from the statistical comparisons. These examples, like those in the LD category, did not fit the syntactic profile in which we are interested: cases where the pronoun's antecedent is the subject of the full clause.

Contained in the third eliminated group are examples identified in the data and labeled "noun phrase" ("NP"). In these examples the prepositional phrase is unmistakably modifying the location of the object of the action; the prepositional phrase does not describe the location of the event denoted in the clause, rather it refers to the specific location of the object of the verb. We are interested only in PPs which can arguably be interpreted to reference the location of the event. Syntactically, the PPs in the NP category would be identified unmistakably as adjuncts to the NP object of the verb; therefore, these PPs would not signal where the event took place. The following sentences are examples from both the English and Spanish corpora which belong in the NP category:

(3.13) *Cziffra$_i$ only occasionally succumbs to that particular devil within him$_i$ but, when he does, it can be either disastrous, as in the middle section to the slow movement of Tchaikovsky's First Concerto , or it can miraculously work, as in Beethoven's " Rage over a Lost Penny ", where an unexpected surge of tempo cannot fail to take his listeners with him.* (BMC 2450)

With many verbs, it is often unclear as to whether or not the PP in a particular example is denoting the location of the object of the event or the event itself; this ambiguity can be seen in 3.14:

(3.14) *He "felt a surge of patriotic emotion within him" but almost at once he "incontinently began to analyse his wave of emotion and to wonder how much of it was due to the romantic beauty of his surroundings".* (EC8 504)

In 3.14, one could make the argument that the PP *within him* is modifying the location of the emotion, the emotion is located within the referent. However, it is also possible to interpret the PP as designating the location of the event of feeling; it is the whole event of feeling that is located within the referent not just the particular emotion. Ambiguous cases such as these were not excluded, but rather included it in the main corpus.

The fourth excluded category is called "Phrasal Verb" (PV). In this category the prepositional phrase is clearly semantically unified with the immediately preceding verb and its meaning is often idiomatic. The preposition in the phrasal verb is unique to the phrasal verb, forms part of a unique semantic structure, and is not exchangeable with another preposition. Büring's deictic adverb test for phrasal verbs was mentioned in Chapter 2. Another regular test used to decide if a verb-preposition sequence is a phrasal verb is if the construction can be replaced by a semantically similar, single verb such as *have confidence in = trust* (Lindstromberg 1998: 248). The final test is syntactic: if the PP in question can be fronted to the beginning of the sentence, then the verb-PP sequence is not labeled a phrasal verb, but if the PP cannot be fronted, such as ***On him**, I am counting* _____ , then the verb-PP sequence is labeled a phrasal verb (Lindstromberg 1998: 249). The examples put into this category passed all three of

these tests. Because Spanish prepositions which participate in phrasal verbs were not used in the corpus, this category is only relevant to the English data. The following sentences are examples from the English search results which have been put into the PV category:

(3.15) *His old friend Barfield ruefully suggested that "at a certain stage in his life he$_i$ deliberately ceased to take any interest in himself$_i$ except as a kind of spiritual animus taking stock of his moral faults ...* (A7C 308)

(3.16) *He$_i$ has total faith in himself$_i$ that what he set out to do, he will achieve.* (HUB 786)

(3.17) *Mike$_i$ took a grip on himself$_i$.* (CMJ 273)

(3.18) *The competitive Morris$_i$ was being unduly hard on himself$_i$, though there was no getting away from the fact that a missed kick by him to touch in the final second against Wasps at Sudbury had spelt defeat.* (AKE 1355)

These examples were excluded from the macro-semantic analysis because once looked at as a group, it is clear phrasal verbs pattern identically to regular verbs in regard to anaphora. That is, the pronoun object of the particle in the phrasal verb is the target of the event denoted by the verb-preposition combo. The grammar treats the verb-preposition combination as one unit, whose individual parts should not be read separately in the grammar.

By removing these four categorical exceptions, a functionally distinct group of data emerges. In this data, the prepositional phrase, whether it be a syntactic argument, as in ditransitive constructions, or a syntactic adjunct, as in most PPs adjoined to the VP, semantically marks the location of the event denoted in the VP.

Chapter 4

A Macro-semantic analysis of spatial factors

4.1 Introduction

As discussed in Chapter 3, various potential semantic hypotheses to account for anaphoric distribution in the PP have been proposed in the literature. None of these hypotheses have been tested on large quantities of real-world data. Playing on the term "macroscopic" (Biber 1988: 61), coming from the domain of corpora-based analysis (discussed in Section 4.2), I introduce here the term 'macro-semantic'. We can define a macro-semantic analysis as one in which a large body of data is statistically subjected to broad semantic categorization. The goal of this approach is to take the semantic predictors (of pronominal distribution) that have been identified in the literature and see how they fare when subject to large quantities of data. The assumption in this approach is that if a particular semantic parameter is significantly statistically correlated with one pronoun type over the other, then than that parameter is directly or indirectly influencing the speaker's choice to use that pronoun. This chapter will integrate and adapt these proposals to a cognitive-based model of the English corpus distribution and follow up with supporting analysis from the Spanish Corpus.

The use of the corpus in the macro-semantic analysis, in and of itself, does not negate the validity or usefulness of individual grammaticality judgments. The macro-semantic analysis is used here as a tool to a) substantiate certain grammaticality judgments given in the literature and b) support or refute, and most importantly, explain the intuitions of the researcher who gives those judgments. Here I am not

advocating to supplant previous analyses on the topic; rather I am arguing that the new approach will inform existing debates on how to model the distribution and lead to new conclusions about the 'best' way to do so. With that said, individual examples provided in past literature will be referenced and relied on in the overall presentation of the phenomenon. Corpus based methodology should not be seen as an enemy of author-generated data, nor vice-versa. There is a place for both approaches in linguistic analysis.

Three spatial parameters have been identified explicitly or implicitly in the literature on anaphora as influential to anaphoric distribution. They are containerhood, peri-personal space (to be defined below), and directionality. For each parameter tested, background research, if any, will be incorporated into an explanation of the statistical results. The aim is to present a coherent picture of why the distribution patterns as it does by identifying and exploring these interactive predictor variables.

We can describe these predictor variables as interactive because here we will demonstrate that each variable has an (unquantifiable) predictive power. As each variable is 'added' to the data, the pronoun choice becomes more and more skewed to the reflexive pronoun. Some data fall into all three categories, and that data most favors the reflexive pronoun. In our discussion of the data trends we will explore why these semantic factors are related to one another and to what extent these relationships reflect universal cognitive processes encoded in language.

The chapter begins with a description of the pronominal distribution by preposition. As Faltz (1977) would predict, the distribution is not just sensitive to the semantics of the main verb but also sensitive to the semantics of the preposition involved. After confirming Faltz's intuition, it is shown that the lack of uniformity

across the seventeen prepositions serves as evidence against the syntactic models discussed in Chapter 2. The chapter continues with frequency counts based on the semantic factors outlined above. The final section of the chapter is an exposition on how the most predictive semantic variables relate to one another. Later in the thesis it will be argued that these semantic variables at work in the PP distribution can be applied to other anaphoric contexts as well.

4.2 Pronominal distribution by preposition

The coded corpus includes anaphoric data from examples with the following seventeen prepositions: *next to*, *beneath*, *behind*, *in front of*, *before*, *below*, *above*, *beyond*, *on top of*, *on*, *toward*, *around*, *out of*, *into*, *in*, *inside*, and *within*. Table 4.1 shows the number of examples of reflexive and nonreflexive pronouns that appear within the prepositional phrase headed by each preposition.

PREPOSITION	NONREFLEXIVE	REFLEXIVE	TOTAL
next to	**5**	**0**	5
beneath	**24**	**0**	24
behind	**593**	**2**	595
in front of	**130**	**2**	132
before	**58**	**4**	62
below	**5**	**2**	7
above	**16**	**11**	27
beyond	**1**	**4**	5
on top of	**4**	**2**	6
on	**73**	**41**	114
toward(s)	**129**	**5**	134
around	**176**	**23**	199
out of	**0**	**13**	13
into	**0**	**37**	37
in	**97**	**81**	178
inside	**36**	**21**	57
within	**18**	**54**	72
TOTAL	1365	302	1667

Table 4.1-Counts of pronoun distribution according to head preposition

Table 4.2 characterizes these same pronoun counts as percentages. Thus, each cell represents the percentage of reflexive versus nonreflexive pronouns to occur within the individual preposition context.

PREPOSITION	NONREFLEXIVE	REFLEXIVE	TOTAL
next to	**100%**	**0%**	100%
beneath	**100%**	**0%**	100%
behind	**99.6%**	**0.4%**	100%
in front of	**98%**	**2%**	100%
before	**94%**	**6%**	100%
below	**71%**	**29%**	100%
above	**59%**	**41%**	100%
beyond	**20%**	**80%**	100%
on top of	**67%**	**33%**	100%
on	**64%**	**36%**	100%
toward(s)	**96%**	**4%**	100%
around	**88%**	**12%**	100%
out of	**0%**	**100%**	100%
into	**0%**	**100%**	100%
in	**54%**	**46%**	100%
inside	**63%**	**37%**	100%
within	**25%**	**75%**	100%
TOTAL	83%	17%	100%

Table 4.2- Percentages of pronoun distribution according to head preposition

The distribution is clearly not equal among the seventeen prepositions; varying widely, some prepositions such as *next to*, *beneath*, *behind*, and *in front of*, almost exclusively pattern with the nonreflexive pronoun and other prepositions such as *into* and *out of* show no examples with the nonreflexive.

The fact that the distribution is so varied across this set of prepositions suggests that spatial and event semantics are playing a part in the appearance of one pronoun versus the other. Thus, the ensuing analysis will investigate proposed spatial factors that may be influencing the broad range of distributions seen above.

To investigate these differences the corpus data was grouped into categories defined by selected spatial parameters. To gage if these spatial factors are influencing the distribution, we will want to see if the selected spatial parameter changes the frequency of the distribution of the pronouns. A change is measured by a deviation from our corpus baseline percentages (1365 (83%) nonreflexives, 302 (17%) reflexives). These baseline totals consist of all the PPs that contain a pronoun which is coreferential with the subject of the clause (intra-clausal coreference). If there is a change from the baseline, we can assume that there is something about a particular semantic context that is 'more friendly to' the reflexive pronoun (or in some cases we may want to characterize this as 'more friendly' to the nonreflexive) .

Since we are looking at frequency counts, our numbers need to be understood within the context of the larger picture of anaphoric distribution. The individual distributions we examine in this section should be interpreted against a backdrop of asymmetrical frequency between the nonreflexive and reflexive pronouns in the English language as a whole. If we look to the BNC for a representation of total frequency difference between all uses (both coreferential and not; in all syntactic contexts) of the 1st and 3rd person nonreflexive and reflexive pronouns, *him*, *me*, *himself*, and *myself*, we find that the nonreflexives are overall much more frequent in usage than are the reflexives. The nonreflexives, *him* and *me*, combine to get 284,105 total hits (87.4%) compared to the reflexives, *himself* and *myself*, which combine to get 40,988 total hits (12.6%). By using a four-cell chi-square analysis in most cases to measure the effect of the coded semantic conditions on the overall frequency pattern (where nonreflexive is more frequent than reflexive) we can make simple conclusions about categorical differences -conclusions of the sort condition 1 is different from condition 2. Thus the

intention of the analysis will not be to model distributional frequency based on all the psycholinguistic, syntactic, semantic, and phonological predictor variables (of which overall frequency is, according to some theories, one) but rather to simply demonstrate and then try to explain statistically significant semantic-based tendencies that show up in the data.

However, before examining any frequencies based on event semantics, we must return to the syntactic proposals described in Chapter 2. By looking at individual examples from the new corpus, it becomes clear that the syntactic factors outlined in Chapter 2 do not hold true as predictor variables once subjected to large quantities of corpus data.

The most coherent argument in the syntactic literature, argued by both Hestvik (1991) and Büring (2005) (detailed in Section 2.4) is that the PPs which function as syntactic arguments should require the reflexive pronoun; whereas the PPs which function as syntactic adjuncts should require (Hestvik) or prefer (Büring) the nonreflexive pronoun. For examples in which the PP encodes the spatial semantics of the verb-denoting event, argument/adjuncthood is tested for via the semantic parameter of optionality (Grimshaw 1990); that is when the PP forms part of the event-internal semantics (when it is "theta-marked by the predicate as a function of the predicate's argument structure (Grimshaw 1990: 108)"), it is labeled an argument; however, when the PP serves more as a scene-setting descriptor (not theta-marked by the verb), easily replaced with a different PP (without changing the verbal semantics) or left out of the sentence entirely, it is labeled an adjunct. Based on these standard divisions, all of the following sentences 4.1-4.6 would be categorized as having PP adjuncts; however, both pronouns are easily found within this syntactic context:

(4.1) *As he looked at them walk away, something hard in him revolted and* ***he$_i$ cried inside himself$_i$.*** (A7J 121)

(4.2) *Jack could see hard frost gleaming on the windows and* ***he$_i$ could feel the bleak coldness of the beds inside himself$_i$.*** (BPD 2898)

(4.3) ***He$_i$ was smiling, but holding a gun out in front of him$_i$.*** (HTY 2134)

(4.4) *He$_i$ rushed from the cell,* ***PRO$_i$ shutting the door behind him$_i$.*** (FAB 3273)

(4.5) ***He$_i$ put the opened bottle down next to him$_i$*** *and smelled the top.* (CA3 641)

(4.6) *Another --; and this is one of the most perfect sights in all cycling --; is the silken action of one of those men* ***who$_i$ can sit perfectly still, hands lightly on top of the handlebars, and spin his legs beneath him$_i$*** *as though their motions came from a quite different source of power.* (B35 401)

The difference among 4.1-4.6 is not syntactic but rather semantic: examples 4.1-4.2 contain prepositions which denote containment, whereas 4.3-4.6 do not. Similarly, the examples in 4.7-4.10 would be classified as containing PP arguments, yet both pronouns are readily found in this syntactic context contrary to both Hestvik and Büring's analysis:

(4.7) ***He$_i$ could have stuck pins into himself$_i$*** *and it would have taken ten seconds for his body to complain.* (FSP 2109)

(4.8) *He$_i$ fell silent,* ***PRO$_i$ looking into$_i$ himself.*** (H84 2873)

(4.9) ***He$_i$ gazed around him$_i$*** *at the late afternoon sky.* (HWA 97)

(4.10) ***He$_i$ gathered his own particular circle of close advisers around him$_i$*** *and was prone to discount views that did not reflect those prevailing in his own circle.* (EDP 1205)

Examples 4.9 and 4.10 are in violation of the supposed syntactic constraint against nonreflexives in argument PPs. Again, in this set of examples the difference between the PP contexts relates to the spatial semantics denoted in the PP (not a syntactic division): examples 4.7 and 4.8 contain prepositions which denote containment, whereas 4.9 and 4.10 do not.

At this point, further discussion of an argument/adjunct distinction will be put on hold. The semantic basis for this syntactic division will be explored in section 4.6. There is good reason to believe that a semantic version of the syntactic notion does contribute to the distribution of anaphora, but by its pure syntactic definition, this categorical boundary is not a good predictor.

If argument- or adjuncthood (as represented in the syntactic literature) does not serve as a good predictor variable, then what categories do make a difference in the distribution? In order to begin to answer this question, we must look back at Tables 4.1 and 4.2 and notice that the prepositions toward the right side of the tables pattern with a much higher rate of reflexive pronouns than those to the left side of the tables. The commonality shared by these prepositions is their inclusion in the category of container prepositions. The next section will detail a hypothesis as to why this semantic factor yields so many examples with the reflexive pronoun.

4.3 Containerhood

Faltz (1977) inadvertently implicates containerhood as a potential semantic influence on pronoun distribution with the following author-generated examples:

(4.11) *Krag$_i$ the robot placed a sandwich in front of him$_i$/?*himself$_i$.* (Faltz 1977: 106 #3.20)

(4.12) *Krag$_i$ the robot unscrewed a panel in his abdomen and placed a sandwich inside himself$_i$/?him$_i$.* (Faltz 1977: 106 #3.21)

He argues that the reflexive pronoun seems more grammatical after the container preposition *inside* than it does after the noncontainer preposition *in front of*.

This intuition is confirmed in the BNC data: prepositions which denote containment (*in, inside, within, into, out of*) much more readily allow for the reflexive pronoun and in certain instances seem to 'prefer' it ($\chi 2(1, N = 1667) = 479.91, p < .001$).

Pronoun Distribution by Containerhood

	ContainerPrep	**NonContainerPrep**	**Total**
Nonreflexive	151	1214	1365
Reflexive	206	96	302
Total	357	1310	1667

Pronoun Distribution by Containerhood

	ContainerPrep	**NonContainerPrep**	**Total**
Nonreflexive	42%	93%	83%
Reflexive	58%	7%	17%
Total	100%	100%	100%

Table 4.3- Pronoun distribution according to semantic environment

As shown in Table 4.3, examples involving container prepositions pattern with the reflexive pronoun 58% of the time as opposed to examples with the remaining noncontainer prepositions, which pattern with the reflexive pronoun only 7% of the time. By default, all the cases in the corpus with a container preposition reference the human body as a metaphorical (and in very few cases literal) container as illustrated by the following examples:

Container Prepositions

(4.13) a. ***He$_i$ could have stuck pins into himself$_i$*** *and it would have taken ten seconds for his body to complain.* (FSP 2109)

b. *He$_i$ fell silent,* ***PRO$_i$ looking into$_i$ himself.*** (H84 2873)

c. *He$_i$ remembered sitting at his desk in the bothy before Bicker arrived,* ***PRO$_i$ trying to tear the story out of himself$_i$*** *and filling pages with blizzards and killing.* (GWF 1208)

d. *In the last chapter, in prison, he says that without the crime* ***he$_i$ would not have found within himself$_i$*** *such questions, desires, feelings, needs, strivings, and development.* (A18 201)

e. *This extraordinary man, explorer, farmer, writer and mystic, taught* ***Charles$_i$ to seek the depth and strength within himself$_i$****, to use the experience of Mountbatten's death, to use the suffering and learn from it: to grow in the spiritual sense; and above all, to ensure that his great-uncle had not died in vain.* (A7H 295)

f. ***The researcher$_i$ who is really at home as one of the natives must seek inside himself$_i$*** *to create his `remote areas'.* (E. Ardener 1987). (A0K 81)

g. *And it illumines too the politics of personal relations: the vital fabric of social life that exists in the silence between people --; exactly that space which is filled by music: `As the person talked to me in a conventional conversation, I knew, I heard that,* ***inside himself$_i$, the person$_i$ perhaps wept****.* (A35 167)

h. *As he looked at them walk away, something hard in him revolted and* ***he$_i$ cried inside himself$_i$****.* (A7J 121)

i. *Jack could see hard frost gleaming on the windows and* ***he$_i$ could feel the bleak coldness of the beds inside himself$_i$****.* (BPD 2898)

j. *And if* ***Jesus$_i$ incorporated in himself$_i$*** *the double role of royal and priestly Messiah, he would indeed have been a figure worthy of such adherence.* (EDY 1588)

k. ***He$_i$ had learned to live in himself$_i$*** *and be self-sufficient, Instead of feeling Kee's presence as an intrusion, upsetting his concentration at a critical stage, he was inspired by her.* (CBN 628)

In each example the protagonist's body serves as a landmark for some type of event (usually metaphorical), which occurs inside of it.[14] Sentences 4.13g and 4.13h refer to emotions, which occur inside the body. Sentences 4.13b, 4.13d, 4.13e, and 4.13f evoke the notion of self-introspection; the metaphorical idea that a person can look for certain truths stored within his interior.

In all the corpus data with container prepositions, the protagonist is understood as a container, which houses his subjective self, his consciousness. It is argued that this unique, metaphorical understanding of the body is the reason for the higher yield of reflexive pronouns in this particular spatial context. Once we understand the details of the metaphor responsible for this interpretation, we can see why the reflexive as opposed to the nonreflexive pronoun would more readily evoke the metaphor.

[14] Example 4.13a is the only example listed in which the action described is a literal action of sticking pins into the body. This example illustrates an understanding of the body as a container with the skin serving as the exterior surface of the container.

The metaphorical conception that humans house internal divisible parts of the self is not unique to present history, nor English for that matter. For example, the metaphor shows up in Freud's 1923 theory of the *Id*, *Ego*, and *Superego* (Freud, A. 1986: 450). In the cognitive literature this metaphor is referenced as the "Divided Person Metaphor" (Lakoff 1996: 100) or the "Divided Self" conceptualization (Talmy 2000a: 431) outlined in 4.14:

(4.14) **The Divided Person Metaphor**
Mappings:
(a) A person is an ensemble (containing one person, the Subject, and at least one other entity, a Self).
(b) The experiencing consciousness is the Subject.
(c) The bodily and functional aspects of a person constitute a Self.
(d) The relationship between Subject and Self is spatial: the Subject is normally either inside, in possession of, or above the Self.

(Lakoff 1996: 100)

The details of the Divided Person Metaphor are as follows:

> The Subject is supposed to be in control of the Self. The Subject can reason, but cannot function directly in the world, as the Self can. The Subject is always the locus of consciousness, subjective experience, perception, reason, and judgment. The Self consists of other aspects of a whole person –the body, emotions, a past history, social roles, and much more (p.102).

Apart from the examples with container prepositions, other, even more explicit examples of the metaphor are readily found in the BNC data, shown in 4.15. In 4.15a and b, the Subject, the referent's judgment and consciousness, is predicated to be above the Self, the referent's 'real-world' status, yielding an idiomatic meaning that the referent is conceited. In 4.15c, the referent's Subject, his consciousness, is predicated to be either inside of outside of his Self, his brain/body, yielding a meaning in which the referent is able to both understand his own actions internally and understand an external perspective of his actions.

(4.15) a. *He was getting above himself.* (FRF 2884)

b. *With a face like that he would naturally be constantly above himself, and he wasn't English, that was for sure.* (HGK 63)

c. *He was at an age when he$_i$ could be both inside himself$_i$ and outside himself$_i$ at the same time: he knew he was only playing a game, but the game still remained real to him.* (HH9 2452)

As Lakoff suggests, in each case in 4.15, the reflexive pronoun cannot be construed as containing the exact same semantic value as its antecedent. (This inequality provides evidence for the argument that coindexation should be used as a formal tool for imprecise, basic identity relations only, not as an indicator that coindexed elements are exactly the same.) In this type of construction, a metaphorical spatial relationship is set up between the two parts of the self. The reflexive pronoun is required to maintain the interpretation that both aspects of the self match the same human referent. These examples are similar to the 'proxy' use of the reflexive pronoun in which two separate referents share the same identity as in 4.15:

(4.16) *Ringo$_i$* (the man) *fell on himself$_i$* (the statue).

Just as the reflexive is needed in 4.16 to assert identity between a human referent and a proxy, the reflexive is needed in 4.15 in order to assert identity between the Subject and Self of the human referent. For this reason the reflexive pronoun is obligatory in these cases. If *him* were substituted for *himself* in 4.15a, for example, the coreferential construal is eliminated. *Him* must refer to a second party:

(4.17) *He$_i$ was getting above him$_{*i}$.*

Looking back to Chapter 2, these examples of a divided self, such as 4.15a, parallel Reinhart and Reuland's semantically reflexive P-predicates in 2.24 repeated here in 4.18:

(4.18) a. **Max rolled the carpet$_i$ over it$_i$.*

b. *Max rolled the carpet$_i$ over itself$_i$.* (Reinhart & Reuland 1993: 689 #67)

Although the syntax of the examples in 4.15 is different from the examples in 4.18 (the antecedent in 4.18 is the direct object of the verb, whereas the antecedent in 4.15 is the subject), the metaphorical (4.15) and literal (4.18b) spatial relationships are the same. In both cases one aspect of the referent is construed as separated from the other aspect. The carpet as a whole is construed as multiple entities, with one part on top of the other, and the person as a whole is construed as multiple entities with one part of the self above the other. This spatial configuration can be termed 'Dual Position Occupancy': one entity (or part of one entity) is understood as occupying two different locations simultaneously. In the English corpus, the reflexive pronoun is used 98% of the time in this construction as indicated in Table 4.4:

Pronoun Distribution by Dual Position Occupancy

	Other	**Dual Position**	**Total**
Nonreflexive	1364	1	1365
Reflexive	251	51	302
Total	1615	52	1667

Pronoun Distribution by Dual Position Occupancy

	Other	**Dual Position**	**Total**
Nonreflexive	84%	2%	83%
Reflexive	16%	98%	17%
Total	100%	100%	100%

Table 4.4- Reflexive distribution in the Dual Occupancy construction

There is an obvious statistical difference between the occurrence of the reflexive pronoun in the overall corpus distribution (in examples that are not of this type) and the

occurrence of the reflexive in the dual occupancy construction ($p < 0.0001$, Fisher's exact test)[15].

The corpus examples of dual position occupancy along with the high occurrence of the reflexive pronoun in PPs of containment suggest that the reflexive pronoun itself is associated with the activation of a metaphorical understanding of the divided person. Why should this be so?

One hypothesis is that this association comes from the lexical semantics of the reflexive pronoun. The reflexive pronoun in English contains the word *–self* within it and is derived from the world *self* (König and Siemund 2000). Quite frequently, cross-linguistically, reflexive pronouns derive from words referencing the mind or body. In his cross-linguistic analysis of reflexivity, Faltz (1977: 31) explains:

> In some languages with head reflexives, the reflexive nominal stem is identical to some ordinary noun stem, suggesting that the ordinary noun has been impressed into special service as a reflexive head. Commonly found used in this way are noun roots meaning "body", "soul", and "head" ...

Although the English reflexive forms have a diachronic link to *self*, the reflexive pronouns nowadays are often referenced in the literature as 'semantically vacuous' markers of reflexive actions (Hellan 1988) or 'empty' lexical items that receive their semantic input solely from their syntactic binder (Hestvik 1991). However, two different types of evidence below suggest that the reflexive pronouns in English have a synchronic semantic connection to *self*.

Psycholinguistic research over the past several decades indicates that the mental lexicon is structured via various forms of lexical fields (Aitchison 1994: 90, Langacker

[15] This table excludes all container examples. It could be argued that examples like *John searched for the truth within himself* are construed as a 'Dual Position Occupancy' in which one aspect of John is located inside his own body and the other aspect isn't; however, this understanding is less clearly an example of dual position occupancy, and so the container data is excluded from this category. In the container examples the reflexive is more likely to occur; but it is not obligatory as it is in the 'Dual Position Occupancy' construction.

2002: 56), in which both semantically and phonologically related word representations are linked via neurological connections (Feldman 2006). It may be the case that, in a neural connectionist framework, reflexive pronouns of the form *–self/-selves* share a close and strong connection to the word *self.* Evidence for this synchronic connection is shown below; the word *self* often appears in the same syntactic and semantic context as the reflexive pronoun. Texts found in the BNC demonstrate that speakers blur the line between the reflexive pronoun and the possessed self:

(4.19) *Hm, said her friends, and hmmmm said Jay herself to* ***her self*** *when she was alone.* (A0L 1396)

(4.20) *She felt every one of her thirty some years sitting on her face, and stared at this disoriented stranger,* ***her self****.*(A0L 3276)

(4.21) *I was happy in London, free, mistress of* ***my self*** *and my pocket. Here it was impossible not to be happy.* (A0U 1562)

In these examples, 4.19-4.21, the possessive pronoun is orthographically separate from *self,* indicating that the self should be understood as a separate entity from its possessor; however, replacing the D-N structure with the reflexive pronoun does not significantly alter the semantics of these examples, because, as argued, the reflexive pronoun includes the semantics of a possessed self as it shares a close semantic connection with the individual word *self.*

Evidence from African American English (AAE) also hints that the *–self* part of the reflexive pronoun is both morphologically and semantically separable from the preceding pronoun. Reflexives in AAE can be syntactically separated and internally modified by an adjective as well as quantified with *all.* Based on the following examples,

Pycha et al. (2005) argue that reflexives in AAE behave more like [[pronoun][NP]] structures[16]:

(4.22) a. *they buy their own stuff an' provide for* ***their own self****.* (Pycha 2005:6 # 25a)

b. *I guess ih's more real 'cause you, you, you 'on', you showin'* ***your real self****.* (Pycha 2005:6 # 25b)

c. *You know, but you should definitely be* ***your reg'lar self****.* (Pycha 2005:6 # 25c)

d. *your frien' make comments like, um, "Shut* ***your little boney self*** *up." or sum'in like that.* (Pycha 2005:6 # 25d)

(4.23) a. *Like they trying to put* ***all theyself*** *on they boyfriend.* (Pycha 2005:6 # 26a)

b. *you sat there an' worked* ***all yourself*** *up thinkin' tha' tha' was the person an' it wasn't.* (Pycha 2005:6 # 26b)

It's clear that in 4.22b,c and 4.23b, *-self* refers to the inner character of the referent, *-self* is not semantically vacuous, and yet, the distribution of these modified reflexives is no different from the typical use of a regular reflexive pronoun. It is also important to note that these AAE constructions are not foreign to the standard American dialect. These broken-up reflexives illuminate the internal analyzability of the reflexive pronoun in English, and provide further evidence for the hypothesis that the reflexive pronoun shares a synchronic connection to *self*.

Here we've argued that the high percentages of reflexives in container prepositions comes from the pronoun's cognitive link to the concept of self. The Divided Person Metaphor establishes a spatial relation between the two aspects of the person: the internal entity and the external entity. The internal self is housed inside the body, within the external self. The data presented with the reflexive pronoun in 4.13b-4.13k reference events which metaphorically occur inside the body. The use of the

[16] This type of reanalysis can be closely tied to the use of *hisself* versus *himself* in AAE.

reflexive pronoun reinforces the notion that the described event is spatially predicated to occur at a body-internal location. This semantic connection between the reflexive pronoun and the concept of internal self helps to explain the higher use of the reflexive pronoun in prepositional phrases that denote containment.

We can now return to Krag, the robot. Faltz's judgment is that the short distance, nonreflexive pronoun sounds more grammatical in 4.11 and the reflexive pronoun sounds more grammatical in 4.12 (repeated here as 4.24 and 4.25 for convenience):

(4.24) *Krag$_i$ the robot placed a sandwich in front of him$_i$/?*himself$_i$.* (Faltz 1977: 106 #3.20)

(4.25) *Krag$_i$ the robot unscrewed a panel in his abdomen and placed a sandwich inside himself$_i$/?him$_i$.* (Faltz 1977: 106 #3.21)

Both of these PPs are syntactic arguments and of the same verb, re-confirming the fact that syntactic argument/adjuncthood is not a good predictor of pronoun distribution. We can now speculate that 4.25 sounds more natural to Faltz with the reflexive pronoun because the Divided Person, or in this case, Divided Robot, understanding is active. Krag is understood as a human like entity whose body demarcates an interior and exterior boundary. Since the reflexive pronoun is linked to the self, housed in the interior of the body, its use with body-internal predication is natural in 4.25. On the other hand, in 4.24, there is no mention of Krag's interior, no evocation of the Divided Person metaphor; the reflexive sounds unnatural in this context because the Divided Person metaphor is not evoked.

Additionally, this quasi-minimal pair highlights the spatial underpinning of the divided person construal. In this case, because Krag is a robot, the interior of the protagonist serves as a literal spatial predication. In the cases from the BNC data, the

body-internal predication is metaphorical. Nevertheless, even the metaphorical predication of the corpus data sets up a spatial relationship between the protagonist (external and internal) and the entity on which he is acting (internal). For example, in 4.13c the protagonist metaphorically reaches within to get the story. Although not literal, the story is predicated to be inside the protagonist's body. The common use of the reflexive for body-internal events reinforces the link between its use and the physical space which its referent occupies. This spatial association is the subject of the next section of data analysis.

4.4 Peri-personal Space

In his detailed semantic analysis of reflexivity, Kuno (1987) addresses cases of anaphora in the PP, which allow for either the reflexive or nonreflexive pronoun as in 4.26:

(4.26) a. *John$_i$ pulled the blanket over himself$_i$.* (Kuno 1987: 66 #9.18a)

b. *John$_i$ pulled the blanket over him$_i$.* (Kuno 1987: 66 #9.18b)

Kuno argues that the use of the reflexive in 4.26a implies that John is more affected by the event than he is when the nonreflexive is used in the same position. Kuno's intuitions may, in part, be explained by a spatial parameter.

As argued in the last section, the reflexive pronoun is associated with a version of the self that is located in the body, triggering an association of the reflexive with actions that are metaphorically performed within the container-construed body. And, we saw frequency evidence to show that the reflexive appears more frequently with container prepositions than within the set of noncontainer prepositions. Conversely, this means, as is shown in Table 4.3, that the nonreflexive is associated with most body-external events. How, then, can we explain the seemingly contradictory intuition that

the reflexive pronoun feels more appropriate in contexts in which the action is performed outside of, but still close to the referent's body? Kuno's intuition leads us to search for another relevant spatial category that indicates the area on or closely around the referent. If the reflexive pronoun feels more natural in the description of this type of close spatial relationship, it seems reasonable to hypothesize this is the case because the reflexive pronoun is associated with body-internal actions. We could even go so far as to say that because the reflexive pronoun is associated with body-internal action, its use in body-external events is unconsciously 'purposeful'. When a body-external event is described, the verb and preposition alone may not provide much semantic specificity as to whether the event is close or far from the referent's body. Within our category of body-external events we may need some marker to reinforce that the action occurs close to the body. The reflexive can serve as this type of marker. In this section, a connection will be made between the use of the reflexive pronoun and the expression of events which are performed on and around the body. First we will look at neural evidence for the cognitive reality of this spatial category, and then we will look at how the data aligns according to it.

The term *peri-personal* space refers to the motor space around the body in which the body's effectors (hands, arms, feet and head) freely move during action events (Gallese and Lakoff 2005). Thus, peri-personal space can be understood as a region of space around the body whose outer limits are defined by the distance one can extend the body's limbs to perform an action. Actions occur at locations, and Gallese and Lakoff (2005) demonstrate that the semantic relation between an action and its location is part of a conceptual structure which is characterized neurally by specific firing relations that occur both when a monkey (and most likely a human) performs an action in peri-

personal space and then sees or hears an object in that same location (for more explanation and a summary of the "mirror neuron" phenomenon see Feldman 2006: Ch. 13). Hence, the region around the body defined as peri-personal space forms a neurologically recognized and encoded spatial category.[17]

Example 4.27 illustrates an action performed within the referent's peri-personal space:

(4.27) *He$_i$ spends all his time hiding, cloaking and padding things around himself$_i$.* (CH1 7617)

Whereas, example 4.28 shows an event that is most likely not construed as occurring within the referent's peri-personal space:

(4.28) ***He$_i$ gathered his own particular circle of close advisers around him$_i$*** *and was prone to discount views that did not reflect those prevailing in his own circle.* (EDP 1205)

Apart from its status as a nonlinguistic spatial category, there is ample linguistic evidence to treat peri-personal space as a spatial category to which the grammar is sensitive. The reflexive clitic in Spanish serves this role. The use of the Spanish reflexive clitic confirms that at least one language associates the semantic notion of reflexivity with actions performed inside the peri-personal space. Maldonado (1999) discusses several examples of the reflexive clitic *se* in which its use is related to the so-called dative of possession. These cases are when the reflexive clitic is used for verbal actions which are performed on the body or on inalienable objects as in 4.29:

(4.29) ***Se*** *lavó* *los dientes.*
Refl. wash-past-3rdP the teeth
'He brushed his teeth.'

[17] The spatial parameter of peri-personal space may be the literal basis for the abstract notion of *personal sphere* (Dabrowska 1997: 17), which comprises the "persons, objects, locations, and facts sufficiently closely associated with an individual that any changes in them are likely to affect the individual as well." Dabrowska applies this notion to understand the use of the dative in Polish.

In 4.29, the teeth are considered an inalienable part of the body and so the verb must appear in its reflexive form, with the reflexive clitic. However, if the object of the verb is alienable, and not part of the body, such as the car in 4.30, the nonreflexive verb is the only grammatical choice:

(4.30) *Lavó el carro.*
wash-past-3rdP the car
'He washed the car.'

The use of the reflexive in 4.29 is related to another category of reflexive use in which the reflexive clitic indicates that the object of the action is located on the Initiator (Maldonado 1999: 183) of the action. These examples are particularly interesting because the parameter of alienability is neutralized. The following minimal pairs demonstrate this difference:

(4.31) a. *Doña Maru guardó el cambio.* (Maldonado 1999:183 #135a)
Doña Maru put-away-past-3rdP the change
'Doña Maru put away the change.'

b. *Doña Maru* ***se*** *guardó el cambio.* (Maldonado 1999:183 #135b)
Doña Maru refl, put-away-past-3rdP the change
'Doña Maru put away the change.'

The example in 4.31a without the reflexive clitic can mean that Doña Maru put the money in her drawer or in her closet; whereas in 4.31b. the only reading possible is that of Dona Maru putting the money directly into her pocket. The same contrast is represented in the following pair of sentences in 4.32:

(4.32) a. *Mari Carmen guardó el boleto en la bolsa del pantalón*
Mari Carmen put-away-past-3rdP the ticket in the pocket of the pants
'Mari Carmen put the ticket away in the pants' pocket.'
(Maldonado 1999:183 #136a)

b. *Claudia* ***se*** *guardó el boleto en la bolsa del pantalón*
Claudia refl. put-away-past-3rdP the ticket in the pocket of the pants
'Mari Carmen put the ticket away in her pants' pocket.'
(Maldonado 1999:183 #136b)

The interesting contrast is that 4.32a presupposes that *Mari Carmen* isn't wearing the pants; whereas, because of the reflexive *se*, the only possible reading in 4.32b is that *Claudia* is wearing the pants. Maldonado explains that the notion of proximity presupposes the presence of containment and control on the part of the subject.

A further example illustrates this idea. In 4.33, the object is not only within the subject's control, but the subject's body demarcates and restricts the placement of the object:

(4.33) *En* Los siete samurais, *Toshiro Mifume* ***se (*Ø)*** *saca una bala del*
In *The seven samurais* Toshiro Mifume refl. pull-out-pres.-3rd from-the

brazo con los dientes.
arm with the teeth

'In The Seven Samurais, Toshiro Mifume pulls a bullet out of his arm with his teeth.'
(Maldonado 1999:184 #137)

In 4.33, the absence of *se* would make the sentence ungrammatical.

It is clear from these examples that Spanish speakers have extended their understanding of reflexivity to encompass situations in which there is a very close spatial connection between event participants and the objects they manipulate; the reflexive clitic is used when the actor's body is identified as the landmark for the action and the trajector remains within the peri-personal space of the human landmark, and it is the grammar that conveys these spatial semantics with the addition of the reflexive clitic. In Spanish, reflexivity is most often manifested with the reflexive clitic (the details of the reflexive pronoun in Spanish will be outlined in Section 4.9), whereas in English the semantic property of reflexivity is encoded in the reflexive pronoun.

If the same semantic connection between the reflexive marker and peri-personal actions hold for English, then one would expect to find a higher rate of reflexive pronoun in contexts which encode an action performed within the referent's peri-personal space than in contexts which encode an action performed outside of the referent's peri-personal space. (Remember, in all the English data, the referent's body is always the landmark for the action.) This hypothesis is confirmed in the BNC data shown in Table 4.5:

Pronoun Distribution by Peri-personal Space

	Out Per Space	**In Per Space**	**Total**
Nonreflexive	482	515	997
Reflexive	28	261	289
Total	510	776	1286

Pronoun Distribution by Peri-personal Space

	Out Per Space	**In Per Space**	**Total**
Nonreflexive	95%	66%	78%
Reflexive	5%	34%	22%
Total	100%	100%	100%

Table 4.5- Pronoun distribution according to whether or not action is performed inside or outside the referent's peri-personal space

A higher use of the reflexive pronoun occurs when an action is performed inside the referent's peri-personal space (χ^2 (1, N = 1286) = 139.91, $p < .001$). Table 4.5 shows that 34% of the actions performed in the peri-personal space of the referent patterned with the reflexive pronoun; whereas the reflexive pronoun was only used 5% of the time when the action was determined to occur outside of the referent's peri-personal space[18].

[18] Many examples were concluded to be indeterminate as to whether they fit one of the two categories and were thus eliminated for the overall count. For example in (i), it is unclear if John is looking for the keys near his body or farther away on the floor, for example:

We should note here that the nonreflexive is still used more often even when the action is performed in the referent's peri-personal space (hence giving an explanation as to why Faltz judges the nonreflexive to be grammatical in 4.11). The important distinction we want to highlight is the relative use of the reflexive pronoun: the fact that it is used *more often* in an in-peri-personal space construal than in an out-of-peri-personal space construal.

One confounding factor that should be controlled for is the fact that many of the examples categorized to belong to the within-peri-personal space category are also examples that involve the container prepositions. That is, the in-peri-personal space predictor partly subsumes the semantic predictor of containerhood (outlined above in section 4.3). If an event occurs within the referent's body, it automatically occurs within the outer limit defined by the peri-personal space. In order to control for this conflation, in order to analyze peri-personal space as its own independent predictor for the use of the reflexive pronoun, it is necessary to exclude the semantic parameter of containerhood. By doing so, there should be no assumption made that these two semantic factors are unrelated cognitively. Separating the two factors simply allows them to be teased apart for statistical analysis.

Table 4.6 shows that peri-personal space still remains as a significant, although weak, correlate to a higher percentage of reflexive use even after the elimination of the container examples from the sample ($\chi^2(1, N = 933) = 19.56, p < .001$):

(i) *John looked in front of him for the keys*.

Of the examples judged to be indeterminate as to whether or not they occurred in the referent's peri-personal space, 367 (97%) patterned with the nonreflexive pronoun and 13 (3%) patterned with the reflexive pronoun.

Pronoun Distribution by Peri-personal Space (NonContainer)

	In Per Space	Out Per Space	Total
Nonreflexive	364	482	846
Reflexive	59	28	87
Total	423	510	933

Pronoun Distribution by Peri-personal Space (NonContainer)

	In Per Space	Out Per Space	Total
Nonreflexive	86%	95%	91%
Reflexive	14%	5%	9%
Total	100%	100%	100%

Table 4.6- Pronoun distribution according to whether or not action is performed inside or outside the referent's peri-personal space for *non-container* prepositions

Within the group of non-container prepositions (*next to*, *beneath*, *behind*, *in front of*, *before*, *below*, *above*, *beyond*, *on top of*, *on*, *toward*, *around*), actions that are performed within the peri-personal space of the referent occurred 14% of the time with the reflexive pronoun as opposed to only 5% of the time when the action was determined to be outside of the referent's peri-personal space[19].

Returning to Kuno's examples in 4.26 above, we can now suggest that peri-personal space provides a spatial explanation for Kuno's semantic notion of affectedness. Working with author-generated examples, Kuno (1987: 66) argues that there is a semantic difference between the reflexive and nonreflexive in the same position. The (a) examples in 4.34-4.40, with the reflexive pronoun, are described by Kuno as conveying more involvement on John's part, along with an extra semantic nuance that John is more directly affected or touched by the action:

[19] Of the noncontainer examples, 376 were indeterminate for the parameter of peri-personal space, 367 (98%) of those pattern with the nonreflexive and 9 (2%) with the reflexive.

(4.34) a. *John$_i$ pulled the blanket over himself$_i$.* (Kuno 1987: 66 #9.18a)
b. *John$_i$ pulled the blanket over him$_i$.* (Kuno 1987: 66 #9.18b)

(4.35) a. *John$_i$ put the blanket next to himself$_i$.* (Kuno 1987: 66 #9.19a)
b. *John$_i$ put the blanket next to him$_i$.* (Kuno 1987: 66 #9.19b)

(4.36) a. *John$_i$ put the blanket under himself$_i$.* (Kuno 1987: 66 #9.20a)
b. *John$_i$ put the blanket under him$_i$.* (Kuno 1987: 66 #9.20b)

(4.37) a. *John$_i$ pulled the rope toward himself$_i$ and won the game.* (Kuno 1987: 66: #9.21a)
b. *John$_i$ pulled the rope toward him$_i$ and won the game.* (Kuno 1987: 66 #9.21b)

(4.38) a. *John$_i$ strung the rope around himself$_i$.* (Kuno 1987: 67 #9.23a)
b. *John$_i$ strung the rope around him$_i$.* (Kuno 1987: 67 #9.23b)

(4.39) a. *John$_i$ spilled gasoline all over himself$_i$.* (Kuno 1987: 67 #9.24a)
b. *John$_i$ spilled gasoline all over him$_i$.* (Kuno 1987: 67 #9.24b)

(4.40) a. *John$_i$ poured the gasoline all over himself$_i$.* (Kuno 1987: 67 #9.25a)
b. *John$_i$ poured the gasoline all over him$_i$.* (Kuno 1987: 67 #9.25b)

In 4.34a for example, Kuno would argue that John is more affected by the event of pulling the blanket than he is in 4.34b; the blanket would be interpreted as covering more of John's body in 4.34a than it would in 4.34b. As suggested, another way to characterize Kuno's semantic intuitions is to reinterpret them through the lens of peri-personal action construal. Kuno's examples with the reflexive pronoun may be focusing the close spatial relationship between trajector and landmark to a level that is not achieved with the use of the nonreflexive pronoun. All the events in 4.34–4.40 can be interpreted as occurring within the John's peri-personal space, but the if the reflexive pronoun is used, the event must be understood as taking place within the referent's peri-personal space. This entailment can be teased out of the data in two ways. First, it is shown in 4.41 that the nonreflexive pronoun is natural in a scenario in which the action and object are distanced from the referent's body:

(4.41) *John received his million dollars by winning the timed contest. John was locked in the cage. Although the rope had to remain outside of the cage, somehow John$_i$ strung the rope around him$_i$, so that it formed the required circle shape outside of the cage.*

The nonreflexive does not entail that the action is performed in the referent's peri-personal space. Because the reflexive does entail a within peri-personal space construal, it sounds unnatural in the same scenario in 4.42:

(4.42) *John received his million dollars by winning the timed contest. John was locked in the cage. Although the rope had to remain outside of the cage, John$_i$ strung the rope around ???himself$_i$, so that it formed the required circle shape outside of the cage.*

In 4.42, the use of the reflexive triggers an event simulation (in the sense of Bergen 2005) in which John is tying the rope around his body; this action is inconsistent with the subsequent information that the rope lays outside of the cage. There is no way for John to tie the rope around his body if the rope sits beyond a physical barrier.

The reflexive's entailment of an in-peri-personal space construal is secondarily demonstrated by the use of the long-distance reflexive. The long-distance reflexive sounds natural in 4.43:

(4.43) *John reached out, grabbed the can, and the gasoline spilled all over himself.*

The reflexive is natural in 4.43 because it is implied that the gasoline covers John's body. If the gasoline spills onto a location near John, but not onto his body, as in 4.44, the reflexive becomes ungrammatical:

(4.44) **John reached out, grabbed the can, and the gasoline spilled next to himself.*

These two entailment tests suggest that the use of the reflexive pronoun in these contexts requires an in-peri-personal space construal. As both the Kuno data and the corpus data show, the nonreflexive can be used for actions that are performed on the referent's body, but its use does not entail or force close spatial relationship between protagonist/landmark and object/trajector.

We can now explain the higher use of the reflexive pronoun for actions inside (as opposed to outside) the referent's peri-personal space (shown in Table 4.6). In the relevant examples, the reflexive more focally implicates the body as the relevant event location, while the nonreflexive does not. At a broader level of analysis, the revelation that peri-personal space is associated with a higher occurrence of the reflexive pronoun again confirms that certain types of spatial relation are correlated with the reflexive pronoun. Thus far we have seen these relationships to be ones of containment and personal space. The next section investigates a final schematic relation that has been implicated as a semantic context friendly to the reflexive pronoun.

4.5 Directionality

As mentioned in the introduction to this chapter, Wechsler (1997), like MacDonald (2004), notes that English directional prepositions pattern differently from locative prepositions with regard to anaphoric pronoun distribution. Sentence 4.45a and 4.46a show nondirectional prepositions occurring with the nonreflexive pronoun, and 4.45b and 4.46b show directional prepositions occurring with the reflexive pronoun:

(4.45) a. *Bubba$_i$ tossed the beer can behind him$_i$ / *himself$_i$.* (Wechsler 1997:15 #38a)

b. *Bubba$_i$ tossed the beer can to *him$_i$ / himself$_i$.* (Wechsler 1997:15 #38b)

(4.46) a. *Corporal Crump$_i$ pinned the medal beside him$_i$ / *himself$_i$ (on the wall).*
(Wechsler 1997:15 #39a)

b. *Corporal Crump$_i$ pinned the medal onto *him$_i$ / himself$_i$.* (Wechsler 1997:15 #39b)

The same contrast is evident in the BNC data shown in 4.47.

(4.47) a. *He$_i$ could have stuck pins into* ***himself$_i$*** *and it would have taken ten seconds for his body to complain.* (FSP 2109)

b. *He$_i$ put the opened bottle down next to* ***him$_i$*** *and smelled the top.* (CA3 641)

As a semantic predictor of pronoun distribution, directionality is a far more complex schematic category than containerhood and peri-personal space, and the analysis of directionality needs to be situated at various levels of event semantics. The first complicating factor is the fact that verbs, as individual lexical items, come with an inherent and canonical direction of their own. Secondly, some prepositions are lexically directional and others are recruited into a directional function based on the semantics of the event. In the analysis of directionality as a macro-semantic parameter, care is required in the assignment of directionality to any one particular syntactic and or semantic level. The notion of directionality as a spatial factor needs precise definition, and directionality as a vague concept doesn't accurately model the relevant BNC data.

The idea that directional PPs generate a different syntax than nondirectional PPs has made its way into several formal theories of syntax. These theories utilize functional projections and aim to incorporate basic spatial semantic categories into the syntactic model. This research focuses on two issues: what types and classes of words should be included in the category P (Svenonius 2004a and references cited there) as well as how to formally model the internal semantic structure of these PP heads

(Svenonius 2004b and references cited there). Within this model head Ps are allowed functional projections such as PATH or PLACE, which are meant to capture the categorical semantics of spatial adpositions and are used to model distributional regularities found among spatial systems. Functional projections have also served as a theoretical tool to explain certain interactions between spatial terms and other areas of the grammar. For Norwegian, Tungseth (2003), argues that directional prepositional phrases differ structurally from nondirectional, locative PPs. She uses the *do-so* substitution test to suggest that goal PPs are complements of the verb, whereas locative PPs attach as adjuncts outside the VP, mimicking the similar claim made by MacDonald (2004) for Spanish and Büring (2005) for English.

Because Norwegian nonreflexive pronouns have an anti-subject orientation, a comparison of reflexive versus nonreflexive in the Norwegian syntactic equivalents of the BNC data in 4.1-4.10 is not possible; nevertheless, in Norwegian, the use of a reflexive versus a nonreflexive possessive pronoun interacts with the semantics of motion conveyed in the verb phrase. Tungseth (2003) shows that the choice of pronoun dictates whether the motion in the sentence will be interpreted as directional or nondirectional:

(4.48) *Jeg kastet Per*$_i$ *i svømmebassenget* ***sitt***$_i$ *(directional).* (Tungseth 2003:480 #12a)
I threw Per in swimming pool-DEF REFL
'I threw Per into his swimming pool.'

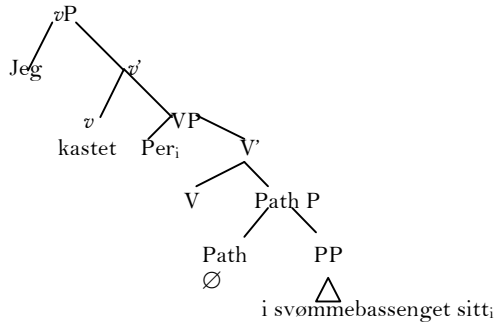
(4.49) *Jeg kastet Per*$_i$ *i svømmebassenget* ***hans***$_i$ *(locative or directional).*
I threw Per in swimming pool-DEF PRON (Tungseth 2003:480 #12b)
'I threw Per in his swimming pool.'

The sentence in 4.48 with the reflexive has only a directed motion reading, in which Per's trajectory starts outside the swimming pool and ends inside the swimming pool, while 4.49 with the nonreflexive is ambiguous between the directed motion reading and

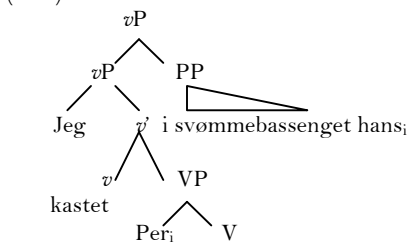
a located motion reading, in which the throwing event occurs within the confines of the pool.

Evoking the functional projection of PATH in the syntax, Tungseth argues that the directed motion reading has a Path P complement with an empty head located inside VP (shown in 4.50) and the located motion reading attaches outside *v*P (shown in 4.51)[20]:

(4.50) Directed Motion



(4.51) Located Motion



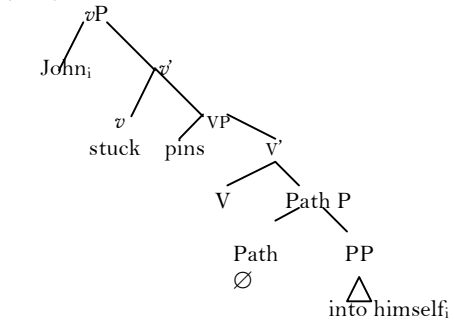
Tungseth's model for Norwegian could also be implemented in English. The resulting structural account mimics the distribution predicted by Conditions A and B of Binding Theory. The generalization in English that a semantic notion of directionality accounts for the PP's complement-like status has thus far, in standard accounts of Binding Theory, not been modeled in the syntax; however, by introducing the semantic notion of 'path' into the syntax, Tungseth implies that a structural account of the distribution in 4.47a and 4.47b is possible.

The tree in 4.52 shows how the directional semantics instigated by *into* motivate

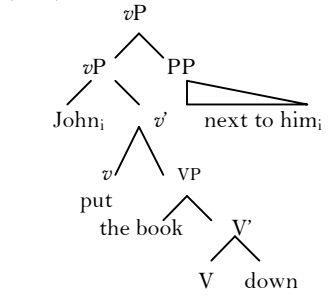
[20] A potential problem arises for Tungseth's structural account when confronted with the directional reading with the nonreflexive pronoun in 4.50. Based on her analysis, the directional reading results in the tree in 4.50; thus, if the nonreflexive is used with a directional interpretation, the structure in 4.50 should be active. As a result, the nonreflexive pronoun would be bound by its antecedent in violation of Condition B. In order to save Tungseth's analysis, one can implement Hestvik's (1991) arguments that the PP serves as a local domain, a CFC (Complete Functional Complex). In 4.50, the minimal domain which includes the pronoun and its governor is the PP. If the pronoun is nonreflexive, then, based on this new characterization of the CFC, it is free, even if it is co-indexed with the subject of the sentence, so it does not violate Condition B.

the PP to be represented as a Path P complement. As part of the VP, the reflexive pronoun in the directional PP is bound by the subject and Condition A is met. The tree in 4.53 shows how the nondirectional semantics prompted by *next to* do not trigger the PP to be represented as a Path P complement, and it consequently attaches outside *v*P as an adjunct. Outside *v*P, the nonreflexive pronoun is free, and therefore, Condition B is met.

(4.52) Directed Motion



(4.53) Located Motion



An analysis such as this relies on strict categorical divisions between PP adjuncts and PP complements. We have already shown this division to be problematic. Many BNC examples contradict the strict division to which these researchers allude. The reflexive pronoun certainly does occur in contexts in which the referent directs an action toward his body as shown in 4.54:

(4.54) a. *The boat was wedged across a rock where the current was not as bad as that elsewhere, and they threw a rope across to McCandless$_i$ who$_i$ tied it around himself$_i$ and the near-unconscious woman.* (BM9 874)

b. *He$_i$ wraps around himself$_i$ the dusty patchwork cloak of his invisibility and the heralds gallop hither and yon in vain.* (ECU 2345)

Yet, more examples involving body-centered directional actions occur with the nonreflexive pronoun as in 4.55. (In a syntactic theory like Tungseth's these directional PPs would have to be treated as complements, therefore violating Condition B.)

(4.55) a. ***He$_i$ pushed the pad aside, PRO$_i$ drew the typewriter toward him$_i$*** *and began to type again, squinting down at his friend's tiny handwriting.* (A08 449)

b. ***He$_i$ pulled the notebook toward him$_i$*** *and wrote: Check early drypoint, Heroin for a Penny , refs.* (A08 486)

c. ***He$_i$ stopped and turned her toward him$_i$*** *and tried to kiss her on the mouth.*(A0N 1858)

"Directionality", the semantic descriptor used by Weschler, MacDonald, and Tungseth, incorporates a large number of corpus examples and includes a broad ranges of prepositions: those that are always directional (*around*, *toward*, *into*, etc.) and those that are sometimes directional (*in*, *on*, etc.). Examples were coded as 'directional' not by the inherent 'directionalness' of the preposition, but according to the directionalness of the PP in the context of the entire predicate. At the macro-semantic level, directional PPs do pattern with a slightly higher rate of reflexive pronouns than the rate of reflexives in nondirectional contexts ($\chi^2(1, N = 1667) = 7.29, p < .01$): among directional PPs the reflexive patterns 22% of the time in comparison to 16% of the time among nondirectional PPs, but, contrary to intuition, within the directional data, the nonreflexive pronoun is actually more common.

Pronoun Distribution by Directionality

	Directional	Nondirectional	Total
Nonreflexive	394	971	1365
Reflexive	111	191	302
Total	505	1162	1667

Pronoun Distribution by Directionality

	Directional	Nondirectional	Total
Nonreflexive	78%	84%	83%
Reflexive	22%	16%	17%
Total	100%	100%	100%

Table 4.7- Pronominal distribution according to directionality

The sharp dichotomy insinuated by Wechsler (and MacDonald 2004) is not only unconfirmed but the data actually demonstrates a pattern opposite to what would be their prediction since, within the BNC sample, 78% of directional PPs pattern with the coreferential nonreflexive pronoun. (Wechsler's self-produced data would suggest that directional PPs should pattern at a much higher rate with the reflexive than with the nonreflexive.)

Wechsler, MacDonald, and Tungseth's claims of an association between directional PPs and argument PPs, although simplistic, are not without merit. Clearly, a vague descriptor of directionality should not be part of the definition of a traditional syntactic argument, a constituent that is selected for by the verb. We have seen many cases thus far from the BNC examples that contradict this notion: many directional PPs pattern with the nonreflexive pronoun. These would be examples of short-distance pronouns that violate the binding conditions upon which these researchers rely. These formal theories, which insert only limited semantics into a syntactic apparatus, fail to explain important idiosyncrasies in the data. However, the intuition put forth by Weschler, MacDonald, and Tungseth deserves further exploration. Self-directed actions expressed with directional PPs do seem exhibit a tight relation between the verb

and the PP. The spatial details of a subset of the directional examples show that this schematic variable is important for pronoun distribution.

When the semantic parameter of directionality is combined with containerhood, the occurrence of the reflexive pronoun is extremely high and significantly differs from the reflexive distribution in container contexts that are not directional ($\chi^2(1, N = 357) = 32.09, p < .001$).

Pronoun Distribution by Directionality of Container Preps

	Directional	**Nondirectional**	**Total**
Nonreflexive	2	149	151
Reflexive	45	161	206
Total	47	310	357

Pronoun Distribution by Directionality of Container Preps

	Directional	**Nondirectional**	**Total**
Nonreflexive	4%	48%	42%
Reflexive	96%	52%	58%
Total	100%	100%	100%

Table 4.8- Pronoun distribution in directional and nondirectional actions with container PPs

In PPs headed by the prepositions which combine both directionality and containment (mostly *into*, *out of*, and sometime *in*), the reflexive pronoun patterns 96% of the time, compared to 52% of the time in all the other container contexts. Examples of this type also necessarily conflate the semantic parameter of peri-personal space since all examples with container prepositions are by default included in the category of actions performed inside the referent's peri-personal space.

Directionality has less of an effect when combined only with the semantic factor of peri-personal space, but the effect is still significant ($\chi^2(1, N = 437) = 21.748, p < .001$). With the container examples excluded, directional actions in the referent's peri-

personal space pattern with the reflexive 21% of the time as opposed to 6% of the time when the action is outside of the referent's peri-personal space.

Pronoun Distribution by Directionality in Peri-personal Space

	Directional	**Nondirectional**	**Total**
Nonreflexive	176	203	379
Reflexive	46	12	58
Total	222	215	437

Pronoun Distribution by Directionality in Peri-personal Space

	Directional	**Nondirectional**	**Total**
Nonreflexive	79%	94%	87%
Reflexive	21%	6%	13%
Total	100%	100%	100%

Table 4.9- Directionality as a Factor within the peri-personal space of noncontainer examples

We have now seen that directionality on its own is not a very strong predictor of a high distribution of reflexive pronouns (see Table 4.7), but its predictive value is increased once it is combined with the other two schematic variables. As a spatial category, it needs to be scrutinized for hidden complexity.

Because directionality as a predictive factor is complicated, it warrants even further discussion in the next section. This section is an exploration of the relationship between the three spatial variables, and the analysis put forth confirms the proposal that these variables are not independent of one another. There is something special about the constructions that incorporate all three semantic predictor variables. It will be shown that the distribution of anaphora in these constructions is, in many ways, similar to the core anaphoric distribution of simple English clauses.

4.6 The symbiotic relationship of directionality, containerhood, and peri-personal space

At several points in this chapter it has been suggested that the semantic factors outlined here are not independent predictor variables, but are highly related to each other. In some sense, separating these factors for statistical coding is an artificial process since most likely the speaker is conflating these factors when he 'chooses' to use the reflexive pronoun. Of primary interest, some examples that conflate all three semantic factors seem to require the reflexive pronoun in order for co reference to be assumed as in 4.56:

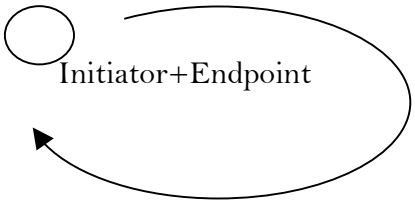
(4.56) *John$_i$ stabbed the knife into himself/???him$_i$.*

Without proper context, the nonreflexive pronoun in 4.56 would most likely be interpreted as referring to someone other than John. In this case, John performs a directional action toward his container-construed body, all carried out within his peri-personal space. Although syntactically different, 4.56 is semantically similar to a reflexive action like in 4.57:

(4.57) *John stabbed himself with a knife.*

In both cases (4.56 and 4.57), John fills both agent and patient/experiencer roles: the very definition of a prototypical reflexive action (Faltz 1977). From a more schematic perspective, Kemmer (1993:52) represents the prototypical direct reflexive situation as an event which involves only one participant, who serves as the Initiator and the Endpoint of the event. She uses the following diagram to illustrate this idea:

(4.58)



Unlike a transitive event schema in which two separate entities fill the Initiator and Endpoint roles, a reflexive event schema requires the same entity to fill both roles.

Van Hoek (1997) further elaborates Kemmer's characterization of reflexivity. In applying event semantics to anaphoric distribution, van Hoek directly addresses examples in which one of the two pronouns is found in a PP: "For some speakers the object of a preposition will be coded as a reflexive (rather than a nonreflexive pronoun) only if it is construed as elaborating a participant in the energetic interaction profiled by the verb; if it describes part of the setting, so that there is no energetic connection linking the antecedent and the anaphor, a pronoun will be used instead (p.181)." Although the BNC corpus suggest that Van Hoek's observation should be qualified as a tendency rather than a rule, her insight holds for examples such as 4.56 in which the "energetic interaction" of stabbing a knife flows from John, the agent, to John, the patient. Van Hoek's characterization of when a reflexive is used in the PP establishes a parallel between the directional PP data and Kemmer's portrayal of a reflexive event prototype. Both involve the circular energy schematic in 4.58. This story, however, is not so simple. If, as Van Hoek suggests, the reflexive pronoun signals this self-directed energetic relation, then why do so many directional actions, nevertheless, pattern with the nonreflexive pronoun? To answer this puzzle, it is useful to investigate how the

semantics of directionality encoded in the PP interact with the semantics of directionality encoded in the verb.

4.7 Directionality of the PP and directionality of the verb

The analysis in section 4.6 above serves as a hypothesis as to why the reflexive is obligatory in example 4.59, but why is it less common in 4.60 and 4.61, and, for that matter, why is the nonreflexive in 4.60 and 4.61 not only grammatical, but also more frequent?

(4.59) *John$_i$ hurled insults at himself$_i$/*him$_i$.*

(4.60) *John$_i$ drew her toward him$_i$/himself$_i$.*

(4.61) *John$_i$ threw the can away from him$_i$/himself$_i$.*

The ensuing analysis explains why the inherent directionality of both verb and preposition matters in the grammar of pronominal distribution.

Within the set of English verbs that denote other-directed actions, the reflexive is required in the simple clauses when the prototypically other-directed action is self-directed. These predicates, which I will call 'extrinsically reflexive' (following Stojanovic 2002; 313), need the reflexive pronoun to indicate that the event is in fact reflexive in nature as in 4.62-4.64:

(4.62) *John$_i$ attacked himself$_i$.*

(4.63) *John$_i$ bit himself$_i$.*

(4.64) *John$_i$ loves himself$_i$.*

Forming a different English category of verbs, there are a small number of canonically self-directed predicates, which I will call 'intrinsically reflexive' again (following Stojanovic 2002: 313) (Reinhart and Reuland term these predicates 'lexically reflexive'). For example, the predicates *bathe* and *shave* are most often performed on one's own body, and often show up as intransitives in the syntax. However, these verbs are different from other intransitives, e.g. *cry*, *sleep*, etc. since they can co-occur with an object, and are construed as semantically transitive –there is a patient involved. These verbs are also different from other semantically transitive verbs because contrary to typically other-directed, outwardly-directed actions, these actions are canonically inwardly-directed and are most often not marked with the reflexive in English:

(4.65) *John bathed.*

(4.66) *John shaved.*

We can see how these verbs, optionally intransitive, are semantically still understood to have a patient (the subject), but since they are canonically self-directed, that patient doesn't have to be syntactically expressed. (Notice how this is different from a canonically outwardly-directed action like kick, which cannot ever show up in intransitive syntax with a reflexive interpretation: **John kicked (himself)*, and for that matter can only show up as an intransitive at all with a very specific meaning like 'move leg in an outward motion'.)

Now, returning to the complex predicate in 4.60, we see an inwardly-directed verb with an inwardly-directed PP. Examples from the BNC corpus indicate that the nonreflexive pronoun is perfectly grammatical and more frequent than the reflexive

when the direction of the verb coincides with the direction encoded in the PP (both are inwardly-directed in this case) as in the representative example constructed in 4.67:

(4.67) *John$_i$ pulled the book toward him$_i$.*

The action of pulling is typically self-directed and would most likely be performed in the direction toward the puller's body. Schematically, this complex predicate functions similarly to an intrinsically reflexive verb like *shave* since the act of pulling, like shaving, is most likely directed at John's own body. Prepositions denoting overt spatial relationships of the type in 4.67 require an object in the syntax; they have to be syntactically transitive. To be used in the description of an event that is self-directed, they must be followed by one of the two pronouns otherwise they are ungrammatical: **John$_i$ pulled the book toward*. In this way they are unlike many verbs such as those in 4.65 and 4.66, which can appear as syntactically intransitive. A question remains though. Why, then, in these cases, is the nonreflexive the preferred pronoun, and why is the reflexive not required but rather optional?

At this point in the explication a digression is needed. The reflexive pronoun began life as a marker of contrastive focus (Koenig and Siemund 2000). In modern-day English it retains this purpose as in 4.68:

(4.68) *John, himself, wants to finish the job.*

Here, the reflexive focuses John as the agent of wanting to finish the job; the reflexive contrasts John from a set of other possible agents. It is argued here that the same semantic feature of contrast marking is signaled with the reflexive in simple clause

structure shown in 4.62-4.64. In these examples the contrastive element is the fact that John, himself, fills the patient role which would canonically be filled by a different participant. These are examples in which an other-directed verb, contrary to expectation, contrary to typicality, is self-directed. Importantly, the reflexive is needed since it is the only marker of reflexivity and the only marker of inward directionality.

Now, returning to the PP data, the reflexive is not needed in 4.67, as it is not needed with an intrinsically reflexive verb, because the action is already assumed to be self-directed. However, just like with an intrinsically reflexive verb, the reflexive can be used with no effect on grammaticality.

These inward-directed actions can be compared to an outward-directed action as in 4.69:

(4.69) *John$_i$ pushed the book away from him$_i$.*

The action of pushing is typically other-directed and would most likely be performed in the direction away from John's body. Again, in this case there is no contrast to mark. The outward direction of the verb coincides with the outward direction of the PP, and together this complex predicate functions just like a simple outwardly directed verb. So here too, the reflexive is not needed since the directionality of the action coincides with the directionality of the preposition.

Under specific (and infrequent) contexts, these directional verbs might be switched to co-occur with a PP whose internal semantics indicate an oppositional direction. In these cases, the reflexive pronoun is, almost obligatorily, needed to indicate coreference as in 4.70 and 4.71:

(4.70) *John$_i$ pulled the book away from himself$_i$/??him$_i$.*

(4.71) *John$_i$ pushed the book toward himself$_i$/??him$_i$.*

The reflexive here serves two functions. It ensures a coreferential relationship between Initiator and Endpoint of the action, and the contrastive semantic element signaled by the reflexive highlights in 4.70 the fact that the typically self-directed action is actually directed away from the referent's body and, in 4.71, the fact that the typically other-directed action is actually directed toward the referent's body.

The analysis has confirmed again that directionality is a complex schematic variable. Inward and outward directionality need to be evaluated at the level of the predicate as a whole. We have seen that each component of the complex predicate, verb and preposition phrase (thus far), carry canonical directionality individually, but once combined in a predicate, the sum needs to be evaluated for canonical directionality at the level of the construction.

Moving even further beyond the combinatorial semantics of the verb and preposition, some examples demonstrate that the verbal object too may be important in the choice of reflexive or nonreflexive pronoun. The following minimal pair shows the importance of the predicate as a whole in determining the choice of pronoun:

(4.72) *John$_i$ stabbed the knife into himself$_i$/???him$_i$.*

(4.73) *John$_i$ stabbed the needle into himself$_i$/?him$_i$.*

It seems that the action of stabbing a needle allows the coreferential nonreflexive pronoun more easily than does the action of stabbing a knife. This difference is presumably accounted for by the expected directionality of the two different predicates.

We will assume knife stabbing is canonically and more commonly an other-directed action. Needle stabbing, in contrast to knife stabbing, occurs more frequently as a self-directed action. Again, the reflexive is needed in 4.72 to reinforce the notion that the action is reflexive in nature. It is less necessary in 4.73 because the predicate itself is more likely to already be self-directed and reflexive. The contrast between these two examples demonstrates that a lexically non-reflexive (in Reinhart and Reuland terms) verb such as *stab* is not evaluated individually, rather it is evaluated for reflexivity at the level of the predicate as a whole which includes verb, object, and preposition.

Further evidence for the inherent directionality of predicates comes from pronoun distribution in Norwegian. Kiparsky (2002: 180) discusses how the complex reflexive in Norwegian is often licensed in reflexivized actions that are typically other-directed, (he calls these "*attack*-verbs"), versus actions that are typically self-directed, (he calls these "*defend*-verbs"). Lødrup (1999: 366) inadvertently illustrates this contrast with the following examples sets:

(4.74) a. *Hun beundrer seg selv / *seg* (Lødrup 1999: 366 #2)
she admires herself

b. *Hun snakker om seg selv / *seg* (Lødrup 1999: 366 #3)
she talks about herself

(4.75) a. *Hun tørker seg* (Lødrup 1999: 366 #6)
she dries herself

b. *Hun pynter seg* (Lødrup 1999: 366 #7)
she dresses-up herself

The actions of admiring and talking about someone in 4.74 are typically other-directed action. In these cases the complex reflexive is used to ensure that the Initiator of the action will be co-construed with the Endpoint of the action; the complex reflexive in Norwegian serves the same function as the reflexive pronoun in English. Contrastively,

the actions in 4.75 are most often self-directed. The Initiator and Endpoint of these actions of drying and dressing-up most often coincide; that is, people most frequently dry and dress-up their own bodies. In these cases the contrastive semantic element of the complex reflexive *seg selv* is not needed. These cases with the simple reflexive pronoun *seg* very much resemble the English example in 4.65 and 4.66 in which no pronoun is necessary in order to construe the Initiator and Endpoint of the action as the same person.

Just as in English, the inherent directionality of Norwegian verbs and PPs can interact to require the complex reflexive. When the inherent directionality of the verb corresponds with the inherent directionality of the PP, the simple reflexive *seg* is allowed as the object of the PP (replicating the pattern of the nonreflexive in English) as in 4.76:

(4.76) a. *Hun dro dynen over seg* (Lødrup 1999: 369 #23)
she pulled the-comforter over herself

b. *Hun kikket rundt seg* (Lødrup 1999: 369 #24)
she looked around herself

c. *Hun kastet det fra seg* (Lødrup 1999: 369 #25)
she threw it away-from herself

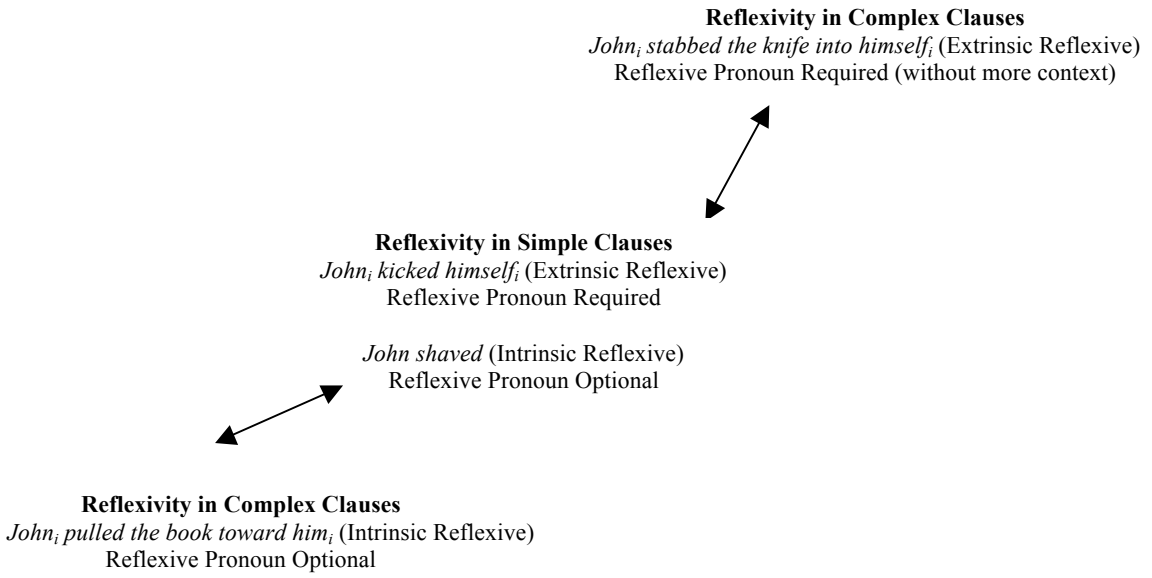
However, when the inherent directionality of the verb opposes the inherent directionality of the PP, the complex reflexive *seg selv* is required:

(4.77) *Hun kastet den mot seg selv/*seg.* (Helge Lødrup 2006, *personal communication*)
She threw it toward herself

The complex reflexive in 4.77 parallels the contrastive semantics of the English reflexive pronoun and ensures a co-construal between Initiator and Endpoint of an action which is typically other-directed.

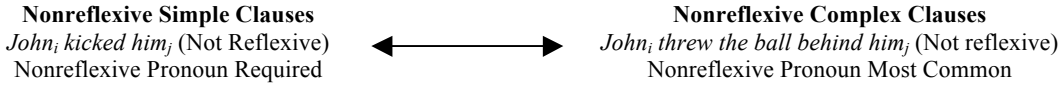
4.8 Interim summary: Reflexive predicates and nonreflexive predicates

We have argued that self-directed actions, performed on the referent's body or within the referent's body, very much resemble the semantics of the prototypical reflexive use in simple clauses. The two clause types share abstract spatial schematics. They are both cases that involve a circular energy transfer in which there is an alignment between the Initiator and Endpoint of the event. Reflexive actions of this type fall into two distinct categories in English: those that are intrinsically reflexive (reflexive marking is optional) and those that are extrinsically reflexive (reflexive marking is required). In a construction-based semantic network, such as that proposed in Cognitive Grammar (Langacker 2002; Van Hoek 1995,1997), structurally different expressions of reflexive actions can be related to one another in a network. In a network such as this, simple and complex predicates (predicates of different form) share basic spatial schematics, and are hypothesized to be part of the same network vis-à-vis shared function. The diagram in 4.78 illustrates this proposed network for the semantic domain of reflexivity:



4.78 Network of Reflexive Clauses

The network diagram reiterates the idea that intrinsically reflexive predicates are optionally marked with the reflexive pronoun and extrinsically reflexive predicates require the reflexive pronoun. If the reflexive pronoun, in an extrinsically marked reflexive predicate, were replaced with the nonreflexive pronoun, the predicate would cease to be reflexive. The corpus data is, of course, also comprised of these nonreflexive predicates. These are cases where the pronoun and subject are coreferent, but the Initiator and Endpoint are not the same entity. These denote outwardly directed actions that have no circular energy transfer. The event, from start to finish, is construed as an outwardly directed action. In this sense, these complex predicates are similar to simple outwardly directed clauses as shown in the relational diagram below:



4.79 Nonreflexive Clauses

If an event is not reflexive in nature, we expect to find the nonreflexive pronoun in the prepositional phrase, since the reflexive pronoun is a marker of reflexivity. When we do find the reflexive pronoun in these cases, we need to look to other reasons why it may be there (the subject of the next chapter).

We can now appropriately characterize Wechsler's examples in 4.45 and 4.46 from the beginning of Section 4.6. They are repeated here. The predicates (verb + direct object + PP) in 4.80a and b are extrinsically reflexive; in order to get a coreferential reading, they must be marked with the reflexive:

(4.80) a. *Bubba$_i$ tossed the beer can to *him$_i$ / himself$_i$.* (Wechsler 1997: 15:#38b)

b. *Corporal Crump$_i$ pinned the medal onto *him$_i$ / himself$_i$.* (Wechsler 1997: 15:#39b)

The predicates (verb + direct object + PP) in 4.81a and b, on the other hand, are not reflexive predicates at all. Both the verb and prepositional phrase are outwardly directed; the predicate as a whole is outwardly directed, so there is no schematic reflexivity involved. In these cases, the nonreflexive pronoun seems more grammatical to Weschler because these actions are not reflexive.

(4.81) a. Bubba$_i$ tossed the beer can behind him$_i$ / *himself$_i$. (Wechsler 1997: 15 #38a)

b. Corporal Crump$_i$ pinned the medal beside him$_i$ / *himself$_i$ (on the wall). (Wechsler 1997: 15 #39a)

Within this section of analysis, two final questions remain. Why are intrinsically reflexive predicates sometimes ("doubly") reflexive; that is, why would an intrinsically reflexive predicate appear with the reflexive pronoun? Secondly, why do some nonreflexive complex predicates occur with the reflexive pronoun? The answer to these questions may lie in discourse factors. This supplemental analysis is the subject of Chapter 5. But before moving on to this different type of analysis, the preceding hypothesis based on spatial factors will be applied to Spanish data.

4.9 Support from Spanish pronouns in the PP

We have thus far argued for a set of three spatial semantic predictor variables that influence the appearance of reflexive pronouns in English. Norwegian data, showing the use of the complex reflexive to mark extrinsically reflexive events, supports the idea that schematic knowledge needs to be built into the grammar of pronoun use. To support the universal nature of these claims, a Spanish corpus, closely replicating the English version, has been built. The Spanish data support the same set of influential semantic variables found in the English data; that is, examples involving self-directed actions, containment, or actions within the peri-personal space yield higher rates of reflexive pronouns than do nondirectional actions, actions involving no containment, or actions performed outside the peri-personal space of the referent.

4.9.1 The Spanish Data

The Spanish corpus includes constructions which most closely parallel the English data. These are sentences in which the 3rd person reflexive or nonreflexive occurs in a

prepositional phrase, which is judged to be part of the verbal predicate. The following chart shows the Spanish nonreflexive pronominal forms that occur after prepositions:

(4.82) **Nonreflexive Prepositional Pronouns**

Singular	**Plural**
me **mí**	us **nosotros**
you **ti**	you **vosotros** (Peninsular) **ustedes** (Latin American)
him, it ***él***	them **ellos** (Masculine)
her, it **ella**	them **ellas** (Feminine)
you **usted**	you **ustedes** (Formal (Peninsular)) (unmarked (Latin America))
it **ello**	

The above set of nonreflexive prepositional pronouns in 4.82 can be compared to the set of reflexive prepositional pronouns in 4.83:

(4.83) **Reflexive Prepositional Pronouns**

Singular	**Plural**
myself **mí**	ourselves **nosotros**
yourself **ti**	yourselves **vosotros** (Peninsular) **Ustedes** (Latin American)
himself, herself, itself, yourself/selves, themselves, oneself - ***sí***	

There is no contrast between the reflexive prepositional pronoun and the nonreflexive prepositional pronoun for the 1st singular, 1st plural, 2nd singular, and 2nd plural forms. Thus, the useful contrast is only possible in 3rd person contexts[21]. Hence, the collected data is limited to contexts in which the PP contains the reflexive *sí* (with a singular, masculine referent) or the nonreflexive, singular, masculine *él*. It should be noted that all of the reflexive prepositional pronouns can be followed by the adjective *mismo* (or *misma/ mismos/ mismas* 'same'- depending on the gender and number of the antecedent) for added emphasis. *Mismo/a(s)* can modify nonreflexive pronouns as well. In the corpus, the presence or absence of mismo after the reflexive pronoun was not controlled for; that is, an occurrence of *sí mismo* was coded in the exact same way as an occurrence

[21] I have listed *sí* as the reflexive prepositional pronoun form for the 2nd person formal singular and plural forms (which would also be the plural informal form in Latin American Spanish), but I am unfamiliar with this use and these examples would be very rare, i.e. *Usted/Ustedes debe(n) hacerlo por sí*. 'You all should do it for yourselves.'

of *sí* alone[22]. There is not a crucial semantic distinction between *sí* and *sí mismo* in the data. One can speculate that the modifier's appearance has to do with dialectal preferences. There were no occurrences of *él mismo* in the data collected for this corpus.

The Spanish data was coded along the same semantic parameters outlined in Chapter 3, save any coding that was specific to first person examples (since the first person data was excluded because of the morphological homophony).

To begin, several differences between Spanish and English reflexive marking need to be highlighted. First, Spanish speakers in general do not use the PP-pronoun structure very often. There are two reasons for this dispreference. One, Spanish, as opposed to English, is a 'verb-framed' language (Slobin 1996). This characterization means that directed motion is often coded in the verb itself, not in a 'satellite' construction, not in an adjoining prepositional phrase. Two, when spatial relations are coded in the satellite construction, they are more likely to appear in nominal form than in a prepositional phrase with a pronoun. For example, the nominal version in 4.84

[22] Pederson (2003) looks at the reflexive vs intensifier distinction in Spanish constructions with prepositional case: i.e. *hacia sí* (*mismo/a/s*) from a historical perspective. He gives the following example to demonstrate the phenomenon:

(ii) *Juan tiró de la cuerda hacia sí (mismo)* (Pederson 2003:1 #1)
[Juan pulled the cord toward himself]

He hypothesizes that *sí (mismo/a/s)*, what he calls "simple reflexive + optional intensifier" is slowly developing toward having a status as a complex reflexive: *sí mismo/a/s*. It should be noted that this complex reflexive, if fully adopted, would replace the simple reflexive, not work in contrast with it as it does in the Norwegian cases. He says this process has occurred in some Danish constructions; he provides the example: *han sad og tænkte over det for sig selv* (*han sad og tænkte over det for sig) [He was thinking about it on his own] (p.1).

The data he looks at indicate a significant increase in the use of *mismo/a/s* in moderrn Spanish compared to Old Spanish. He hypothesizes that there is a gradual interpretive change, by the user, of the intensifier construction toward an interpretation based on a complex reflexive category, as the one observable in Danish and other languages.

(3433 cases of this type in the CREA) is a more common expression of 'next to him' than the PP-pronoun structure in 4.85 (36 cases of this type in the CREA).[23]

(4.84) *Juan$_i$ puso el libro a su$_i$ lado.*
John$_i$ put the book to his$_i$ side.

(4.85) *Juan puso el libro al lado de él.*
John$_i$ put the book next to him$_i$.

When it comes to the expression of reflexivity in simple clauses, Spanish and English are typologically opposite. As discussed in section 4.11, intrinsically reflexive actions in English are optionally marked for reflexivity with the reflexive pronoun. In complete opposition, intrinsically reflexive actions (and of course extrinsically reflexive actions) in Spanish require marking with the reflexive clitic as in 4.86 and 4.87:

(4.86) *Juan se bañó.* **Juan bañó.*
John bathed.

(4.87) *Juan se afeitó.* **Juan afeitó.*
John shaved.

Another typological difference, the one of central concern for the corpus analysis, is the fact that the reflexive pronoun (as opposed to the nonreflexive) in Spanish appears much more frequently in the PP than does the reflexive pronoun in English. (See the differential totals between Table 4.1 for English: 302 total reflexives vs. 1365 total nonreflexives, and table 4.10 for Spanish: 320 reflexives versus only 66 nonreflexive pronouns.) Presumably, the difference between the languages has a historical explanation. In the history of English the reflexive pronoun in various syntactic contexts has come to replace the once sole existing nonreflexive pronoun (Koenig and

[23] The personal pronoun *su* 'his/her/its' does not have a distinct reflexive version as we saw in the Norwegian data.

Siemund 2000); whereas in Spanish, the nonreflexive pronoun has come to replace the almost exclusive use of the reflexive pronoun in PP contexts. This change is confirmed by historical patterns found in the Corpus del Español (Davies 2002). For example the prepositional phrase *hacia sí* 'toward himself' has steadily declined in frequency from the 18th century (2.3 occurrences per million example sentences) through the 20th century (1.1 occurrences per million example sentences); whereas the prepositional phrase *hacia él* 'toward him' has steadily increased from the 18th century (3.1 occurrences per million example sentences) to the 20th century (5.0 occurrences per million example sentences).

4.9.2 The Spanish patterns

Because of the difference in overall rates between the two languages, the subsequent analysis focuses on relative, not absolute percentages of reflexive and nonreflexive distribution. Thus, the important distributions for our purposes are A) the relative distribution of reflexive versus nonreflexive between the different prepositions tested and B) the relative distributions of reflexive versus nonreflexive between the different categorical spatial parameters. Concerning point A, the reflexive and nonreflexive Spanish pronouns, like their English counterparts, do not pattern equally among the nine different prepositions tested as shown in Table 4.10:

Spanish pronoun distribution by preposition

	através de 'through'	**al lado de** 'next to'	**alrededor de** 'around'	**debajo de** 'under'	**delante de** 'in front of'	**dentro de** 'inside'	**detrás de** 'behind'	**encima de** 'on top of'	**hacia** 'toward'	**Total**
Nonreflexive	0	1	0	0	15	10	16	6	18	66
Reflexive	3	0	11	4	16	145	15	6	120	320
Total	3	1	11	4	31	155	31	12	138	386

Spanish pronoun distribution by preposition

	através de 'through'	**al lado de** 'next to'	**alrededor de** 'around'	**debajo de** 'under'	**delante de** 'in front of'	**dentro de** 'inside'	**detrás de** 'behind'	**encima de** 'on top of'	**hacia** 'toward'	**Total**
Nonreflexive	0%	100%	0%	0%	48%	6%	52%	50%	13%	17%
Reflexive	100%	0%	100%	100%	52%	94%	48%	50%	87%	83%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Table 4.10- Count and Percentages of Spanish pronoun distribution according to head preposition

Several prepositions have very few examples with coreferential pronouns, inhibiting the application of a statistical model. However, we can still see, without statistical significance, that there is a differential patterning among the tested prepositions.

Table 4.10 demonstrates that certain prepositions such as *hacia 'toward'*; and *dentro de 'inside'* pattern at a relatively higher rate with the reflexive pronoun, 87% and 94% respectively; whereas prepositions *delante de 'in front of'*, *detrás de 'behind'*, and *encima de 'above/ on top of'* pattern at a lower rate with the reflexive pronoun; 52% or less of all examples included the reflexive pronoun for these prepositions.

We can first examine the small set of examples with the nonreflexive pronoun. Examples with the nonreflexive involve spatial contexts in which the referent is performing an action away from his body or simply possesses (literally or metaphorically) an object in an area near him. For English, we have characterized these events as semantically nonreflexive. The following examples in 4.88 and 4.89

show the nonreflexive pronoun in its canonical use in which the referent is performing an action away from his body:

(4.88) *Este hombre empezó a interrogarme todas las noches durante una semana* ***PRO****ᵢ poniendo delante de* ***él****ᵢ, en el buró, su pistola cuarenta y cinco …*(CREA)

This man$_i$ began to interrogate me every night for a week, PRO$_i$ putting his forty-five pistol in front of him$_i$ in the bureau …

(4.89) *Y la agarra a Nati, la* ***pro****ᵢ pone delante de* ***él****ᵢ y usa su brazo como espada …*(CREA)

And, he grabs Nati, he$_i$ puts her in front of him$_i$, and he uses his arm like a sword …

In 4.88 and 4.89, the referent puts an object in front of him, and the Spanish nonreflexive is used in the same semantic context as would be used the English nonreflexive. In examples 4.90-4.95, the nonreflexive is used, as in English, when the sentence is expressing a simple spatial relation between the referent and the object. These cases involve objects that are in front of, behind, or next to the referent:

(4.90) *Ahora mismo* ***pro****ᵢ tiene grandes problemas delante de* ***él****ᵢ …*(CREA)

Right now, (he$_i$) has big problems in front of him$_i$ …

(4.91) *El* ***bloque****ᵢ saharaui prointegración en Marruecos tendrá detrás de* ***él****ᵢ a un país de 30 millones de habitantes …*(CREA)

The Saharan pro-integration embargo$_i$ in Morroco has a country of 30 million people behind it$_i$ ….

(4.92) *…* ***un médico****ᵢ o un* ***artesano****ᵢ que se va deja detrás de* ***él****ᵢ un vacío que es difícil de llenar …*(CREA)

… a doctor$_i$ or an artesan$_i$ that goes leaves an empty space that is hard to fill behind him$_i$ …

(4.93) ***Don Quijote****ᵢ … deja detrás de* ***él****ᵢ sus libros, su biblioteca: su refugio.* (CREA)

Don Quixote$_i$ …leaves his books, library, his refuge, behind him$_i$.

(4.94) *…* ***pro****ᵢ cerraría la ventana detrás de* ***él****ᵢ después de arrojarse al vacío.* (CREA)

… he$_i$ would close the window behind him$_i$

(4.95) *…* ***el muchacho****ᵢ tiene al lado de* ***él****ᵢ otra botella, por la mitad.* (CREA)

The boy$_i$ has another bottle next to him$_i$, half full.

As in English, the nonreflexive is also found in contexts where the referent looks back to an area behind his body as in 4.96:

(4.96) ***pro***$_i$ *Miró detrás de* ***él***$_i$ *y los demás rieron.* (CREA)

He$_i$ looked behind him$_i$ and the others laughed.

In sum, the few uses of the nonreflexive pronoun in the Spanish PP tend to coincide with the semantic contexts of their English counterparts. That is, the nonreflexive *él* is often found in the PP of a similar context in which the nonreflexive *him* would be found. With that said, the low overall percentage of nonreflexive pronouns addressed earlier indicates that many events in Spanish, which would typically be described with the nonreflexive pronoun in English, are described in Spanish with the reflexive pronoun. In order to better characterize the Spanish data, the next four sections detail how spatial parameters interact with the use of the reflexive pronoun in the Spanish data.

4.9.3 Dual Position Occupancy

As in a small number of English examples, the Spanish reflexive pronoun *sí* is required to indicate that a divided referent is simultaneously located in two different locations (what I have labeled as "Dual Position Occupancy' in section 4.3). Examples such as 4.97 and 4.98, like their English counterparts, require the reflexive to ensure a coreferential reading:

(4.97) *Rilke celebra la sublimidad del canto terrestre, que después de enlazar y unir todos los frutos de la tierra* ***pro***$_i$ ***se eleva por encima de sí***$_i$ ***mismo*** *y lanza un mensaje al vacío cósmico.* (CREA)

Rilke celebrates the sublime nature of the earthly verse, which after intertwining and uniting the fruit of the land, **it**$_i$ **elevates above itself**$_i$ and captures the message of the empty cosmos.

(4.98) *Es un astro con luz propia, que* ***PRO$_i$ gira alrededor de sí$_i$ mismo****, completando una vuelta cada 25 días.* (CREA)

It's an asteroid with its own light, that **PRO$_i$ revolves around itself$_i$**, completing one turn every 25 days.

In 4.97, the message of the poetic verse elevates above itself as to simultaneously, metaphorically hold two positions, one on earth and the other in the cosmos. In 4.98, the asteroid as a whole is understood to be internally decomposable. One part of the asteroid, presumably the outer part, revolves around the other part of the asteroid, the inner part.

4.9.4 Directionality in Spanish

Examples that were coded to involve directional actions patterned at a slightly higher rate with the reflexive pronoun than did those coded as nondirectional ($\chi^2(1, N = 386) = 4.28, p < .05$). Table 4.11 below shows that 88% of directional examples patterned with the reflexive pronoun versus 80% of nondirectional examples patterned with the reflexive.

Spanish Pronoun Distribution by Directionality

	Directional	**Nondirectional**	**Total**
Nonreflexive	19	47	66
Reflexive	136	184	320
Total	155	231	386

Spanish Pronoun Distribution by Directionality

	Directional	**Nondirectional**	**Total**
Nonreflexive	12%	20%	17%
Reflexive	88%	80%	83%
Total	100%	100%	100%

Table 4.11- Spanish pronominal distribution according to directionality

The preposition *hacia* 'toward' accounted for almost all the directional examples in the Spanish data. (Out of a total of 155 directional examples, 137 contained *hacia*.) *Hacia* patterns much more frequently with the reflexive pronoun (86%) than does its English equivalent *toward(s)* (only 4% of the data with *toward(s)* pattern with the reflexive). Spanish complex predicates with verbs, not specified for directionality, and inward-directed PPs pattern with the reflexive pronoun. The reflexive pronoun marks these events as extrinsically reflexive like in English:

(4.99) *Pero entonces el odio y la vergüenza que* ***pro***$_i$ *sentía hacia* ***sí***$_i$ *mismo lo golpearon como golpea una tormenta violenta.* (CREA)

But then, the hate and shame that he$_i$ felt toward himself$_i$ hit him like a violent storm.

(4.100) *El control del ala se hace con el trapecio. Cuando el* ***piloto***$_i$ *tira del trapecio hacia* ***sí***$_i$*, el ala pica y cuando, por el contrario, empuja el trapecio, el ala se encabrita.* (CREA)

Controlling the wing is done with the trapeze. When the pilot$_i$ moves the trapeze toward himself$_i$, the wing moves in one direction, to the contrary, when the pilot pushes the trapeze, the wing moves in the other direction.

(4.101) *Estoy deshonrado! (*pro_i *Voltea hacia* ***sí***$_i$ *la pistola y dispara. Cae.)* (CREA)

I am dishonored! (He$_i$ turns the gun toward himself$_i$ and fires. He falls.)

In 4.99-4.101, we see verbs of unspecified canonical direction (*feel*, *move*, *turn*) paired with an inwardly directed PP marked with the reflexive pronoun. Unlike English however, many Spanish examples of intrinsically reflexive complex predicates also pattern with the reflexive pronoun as in 4.102-4.108:

(4.102) ***Blair***$_i$*, impulsado también por su proyecto, solo que en sentido diametralmente opuesto, ha traído hacia* ***sí***$_i$ *un enorme apoyo popular …* (CREA)

Blair$_i$, driven by his project, in the opposite sense, has brought enormous popular support toward himself$_i$ …

(4.103) *Mientras* ***Gorbachov***$_i$ *intenta atraer hacia* ***si***$_i$ *las voluntades occidentals …* (CREA)

While Gorbachev$_i$ tries to pull Western will toward himself$_i$ …

(4.104) *El animal va directamente al punto A cuando ve la fruta; al no poder alcanzarla con la mano, pone en relación el palo situado en B con el plátano, va a buscarlo y lo* ***pro****ᵢ atrae hacia* ***sí****ᵢ.*

(CREA)

The animal does directly to point A when it sees the fruit; upon not being able to reach it with its hand, it puts the stick sitting in B with the banana, it$_i$ goes to get it and pulls it toward itself$_i$.

(4.105) *Ansioso de* ***pro****ᵢ arrastrar hacia* ***sí****ᵢ toda la atención …* (CREA)

Anxious to PRO$_i$ pull all the attention toward himself$_i$

(4.106) … ***él****ᵢ la atrajo violentamente hacia* ***sí****ᵢ …* (CREA)

… he$_i$ pulls her violently toward himself$_i$ …

(4.107) ***pro****ᵢ Me atrajo hacia* ***sí****ᵢ para darme un fuerte abrazo fraternal.* (CREA)

He$_i$ pulled me toward himself$_i$ to give me a strong friendly hug.

(4.108) *Ésa es otra –contestaba Ignacio*$_i$, ***PRO****ᵢ acercando el cenicero hacia* ***sí****ᵢ* – (CREA)

That is another –replied Ignacio$_i$, PRO$_i$ pulling closer the ashtray toward himself$_i$ -

In 4.102-4.108, canonically inwardly directed verbs, such as *arrastrar*, *atraer*, and *acercar* (loosely translating to 'pull', 'attract/pull', and 'bring closer' respectively), with inwardly directed PPs are still marked with the reflexive pronoun. Thus, the corpus data suggest that inwardly directed complex predicates in Spanish, regardless of their canonical directionality, prefer the reflexive pronoun. However, in the corpus one does find examples of *hacia* 'toward' with the coreferential nonreflexive pronoun as in 4.109:

(4.109) ***Montoya****ᵢ tomó a María de una mano y la atrajo hacia* ***él****ᵢ.* (CREA)

Montoya$_i$ took María by the hand and pulled her toward him$_i$.

The existence of these examples suggests a shift from the reflexive prepositional pronouns to the coreferential nonreflexive pronoun. Remember, in earlier stages, these directional PPs would exclusively take the reflexive. Several Spanish speakers consulted for this thesis agree that the nonreflexive is more natural in these intrinsically reflexive complex predicates such as that in 4.109. A purely speculative argument can

be made that this shift to the nonreflexive is based on the intrinsic/extrinsic distinction relevant in the English cases. One can speculate that for these speakers, as in English, the reflexive may not be needed in 4.109 because the verb *atraer* 'attract/pull' inherently involves self-directed motion, thus indicating the predicate, as a whole, is interpreted as an intrinsically reflexive predicate.

Complicating the lack-of-necessity argument for Spanish PPs, however, is the fact that intrinsically reflexive simple predicates in Spanish such as *bañarse* 'bath' or *afeitarse* 'shave' are always marked with the reflexive clitic. This requisite marking suggests that a reflexive marker is nonoptionally paired with an intrinsically reflexive action. Therefore, in order to accept the lack-of-necessity argument, one must believe that Spanish speakers have decoupled the semantic function of the reflexive clitic from the semantic function of the reflexive prepositional pronoun. Interestingly, there is evidence to suggest this process of semantic distancing is occurring. The evidence comes from historical changes that have taken place from Old Spanish to Modern Spanish. In Old Spanish, the reflexive pronoun functioned as an indicator of a reflexive action on its own.

(4.110) *E* ***pro****ᵢ obligaron a sí e a sus bienes, …* (Pederson 2003: 204 #17)
and they mortgaged themselves and their belongings, …

This system has now been replaced by the required use of the reflexive clitic and the optional use of the pronoun accompanied by the emphatic *mismo/a(s)* (Pederson 2003: 206):

(4.111) (pro) *Se obligaron (a sí mismos) …*
They forced themselves …

(4.112) (pro) **Obligaron (a sí mismos) …*
They forced themselves …

Thus, simple predicates are marked for reflexivity with the reflexive clitic, but the role of the reflexive pronoun is more vague. In simple clauses its role (along with the emphatic *mismo*) is to provide added emphasis that the action is in fact reflexive in nature. But in other cases its role is to provide contrast marking as the English reflexive pronoun does. If these two reflexive markers, in fact, have different functions, it is possible that the reflexive pronoun, in contrast to the reflexive clitic, for some speakers, is only needed to mark extrinsically reflexive *complex* predicates. For these speakers, the reflexive pronoun is not directly linked to intrinsically reflexive predicates. This variety of Spanish replicates English: Spanish self-directed predicates do not require the reflexive pronoun because the predicate alone indicates the action is directed toward the referent's body; there is no incongruence to contrast between the canonical direction of the action and the actual direction of the action.

In sum, the Spanish directional data parallels the English directional data in many ways. The reflexive pronoun serves as a marker of extrinsic reflexivity in complex predicates as it does in English. Although many examples of intrinsically reflexive complex predicates are (redundantly) marked with the reflexive pronoun in the corpus data, their acceptability and naturalness may be lessening in certain dialects since these cases are dispreferred by at least a small subset of Cuban native speakers.

4.9.5 Containerhood in Spanish

The second group of data, involving Spanish prepositions that denote containment, supports the claim, shown for English, that a metaphorical action occurring inside the referent is more likely to pattern with the reflexive pronoun ($\chi^2(1, N = 386) = 21.89$, $p < .001$). 94% of all examples involving containment in Spanish patterned with the

reflexive pronoun as opposed to a proportionally lower percentage of reflexive uses (75%) in examples that did not involve containment.

Pronoun Distribution by Containerhood

	ContainerPrep	NonContainerPrep	Total
Nonreflexive	10	56	66
Reflexive	148	172	320
Total	158	228	386

Pronoun Distribution by Containerhood

	ContainerPrep	NonContainerPrep	Total
Nonreflexive	6%	25%	17%
Reflexive	94%	75%	83%
Total	100%	100%	100%

Table 4.12- Spanish pronoun distribution according to semantic environment

Distributionally mirroring the English trend, Spanish speakers in the corpus used the reflexive pronoun more frequently in examples whose action or state occurs inside the referent as shown in (4.113-4.115):

(4.113) *Jimenez es tierno, es culto a su manera,* ***pro****$_i$ encierra dentro de* ***sí****$_i$ una sabiduría falsamente popular.* (CREA)

Jimenez is tender, he is shy in his own way, he$_i$ encloses inside himself$_i$ a falsely popular knowledge.

(4.114) *La verdad es que cada* ***mexicano****$_i$ lleva dentro de* ***sí****$_i$ el gusto por su música nacional …* (CREA)

The truth is that each Mexican$_i$ carries within himself$_i$ a fondness for his national music …

(4.115) *El* ***individuo****$_i$ siente dentro de* ***sí****$_i$ el pleno querer de la voluntad …* (CREA)

The individual$_i$ feels inside himself$_i$ the desire of his own will …

At a 94% occurrence rate, the corpus data suggest that the reflexive pronoun is virtually required in these metaphorical contexts. Native speaker judgments confirm the naturalness of these examples as well. For example, if the reflexive in 4.115, is replaced with the nonreflexive, the example is dispreferred by native speakers:

(4.116) ???*El* ***individuo****$_i$ siente dentro de* ***él****$_i$ el pleno querer de la voluntad …* (CREA)

The individual$_i$ feels inside him$_i$ the desire of his own will …

4.9.6 Peri-personal space in Spanish

Just as in English, a higher use of the Spanish reflexive pronoun occurs when an action is performed inside the referent's peri-personal space ($\chi^2(1, N = 319) = 21.18$, $p < .001$).

Pronoun Distribution by Peri-personal Space

	Out Per Space	**In Per Space**	**Total**
Nonreflexive	13	38	51
Reflexive	15	253	268
Total	28	291	319

Spanish Pronoun Distribution by Peri-personal Space

	Out Per Space	**In Per Space**	**Total**
Nonreflexive	46%	15%	16%
Reflexive	54%	85%	84%
Total	100%	100%	100%

Table 4.13- Spanish pronoun distribution according to whether or not action is performed inside or outside the referent's peri-personal space

Because the Spanish examples that denote containment (like their English counterparts) automatically occur within the referent's peri-personal space, they potentially confound the effect of peri-personal space alone on the data. Table 4.14 shows that peri-personal space still remains as a significant correlate to a higher percentage of reflexive use even after the elimination of the container examples from the sample ($\chi^2(1, N = 161) = 7.85$, $p < .005$):

Spanish Pronoun Distribution by Peri-personal Space (NonContainer)

	In Per Space	Out of Per Space	Total
Nonreflexive	28	13	41
Reflexive	105	15	120
Total	133	28	161

Pronoun Distribution by Peri-personal Space (NonContainer)

	In Per Space	Out of Per Space	Total
Nonreflexive	21%	46%	25%
Reflexive	79%	54%	75%
Total	100%	100%	100%

Table 4.14- Spanish pronoun distribution according to whether or not action is performed inside or outside the referent's peri-personal space for *non-container* prepositions

79% of all noncontainer actions performed in the referent's peri-personal space use the reflexive pronoun as opposed to its use in just 54% of all examples involving actions performed outside the referent's peri-personal space. The reflexive is used in examples such as 4.117-4.119 in which the referent is performing an action on his body or in the areas immediately adjacent to his body. It should be noted that these examples too were judged to be natural by native speakers:

(4.117) *El **desconocido**$_i$ me cogió por la cintura, entonces, y me situó encima de **sí**$_i$, una de sus manos sostuvo su sexo mientras con la otra me ayudaba a introducirme en él.* (CREA)

The unfamiliar man$_i$ grabbed me by the waist, and then he$_i$ situated me on top of himself, one of his hands grabbed his organ while the other helped put me inside of him.

(4.118) ***Pro**$_i$ Arroja el cuchillo delante de **sí**$_i$.* (CREA)

He$_i$ stabs the knife in front of himself$_i$.

(4.119) *En tales condiciones para un investigador sin segunda intención alguna, el único medio de salvar la vida es, ante cualquier apaición, **PRO**$_i$ echar delante de **sí**$_i$ su muy modesto equipaje y levanter los brazos …* (CREA)

In the cases of the researcher without a second intention, the only way to save his life, before any apparition, is PRO$_i$ to throw his modest luggage in front of himself$_i$ and raise his arms …

The higher use of the reflexive prepositional pronoun in contexts within the referent's peri-personal space is not surprising. Not only does this spatial parameter yield higher rates of reflexive use in English, this spatial parameter is associated with the use of the reflexive clitic in Spanish as well. In Section 4.5, we examined data from Maldonado (1999) that clearly demonstrates that actions performed on the Initiator's body require the reflexive form of the verb. A logical assumption is that reflexive forms in general (of both verbs and pronouns) in Spanish are associated with actions performed inside the protagonist's peri-personal space.

Nonetheless, there are a few examples from the corpus in which the nonreflexive pronoun is found in the PP of a predicate performed within the referent's peri-personal space. These examples may be explained by the tendency alluded to earlier. That tendency is for an action that is most commonly self-directed to allow the nonreflexive pronoun. In certain examples of intrinsically reflexive complex predicates, such as in 4.120, the nonreflexive appears:

(4.120) … *tomó de la mano a la que le gustaba y la* ***pro****ᵢ sentó encima de* ***él****ᵢ, sin preocuparse de lo que pensaron los otros.* (CREA)

… he$_i$ took the hand of the one he liked and sat her on top of him$_i$, without worrying about what other people thought.

The act of sitting a person down most often occurs on one's own lap not on another person's lap, it is inherently a self-directed motion. The contrastive semantic quality of the reflexive pronoun is not necessary.

We've seen in this data that the Spanish corpus follows several trends established in the English corpus. First, canonically other-directed actions (e.g. *push* or *throw out*) and actions that don't seem to have a canonical direction (e.g. *move*) which are in fact self-directed in a given example show higher rates of occurrence with the reflexive pronoun than do actions which are inherently self-directed (e.g. *pull*). Actions which metaphorically occur inside the referent's body pattern almost exclusively with the reflexive pronoun, and actions performed within the referent's peri-personal space

occur more frequently with the reflexive pronoun than do actions performed outside of the referent's peri-personal space.

4.10 Conclusion

This chapter has served as evidence for four main take-away messages about the distribution of pronouns in the PP. First, spatial parameters in the given event influence the tendency toward the appearance of either the reflexive or nonreflexive pronoun in the PP. These spatial parameters, directionality, peri-personal space, and containherhood, are inter-related and the when the three parameters are present (as in an action that is self-directed, in the referent's peri-personal space, and when the body is construed as a penetrable container) as in 4.121, the likelihood of the reflexive pronoun is higher than it is in examples in which these parameters are absent:

(4.121) *John stabbed the knife into himself.*

The second take-away message provides an explanation for the aforementioned distribution. Syntactically complex events like the one described in 4.121, are schematically similar to syntactically simple events like 4.122:

(4.122) *John stabbed himself.*

The same trajector-landmark relations hold (in both 4.121 and 4.122 John is the landmark and a knife is the trajector), and the use of the reflexive indicates in both cases that a typically other-directed action is, in fact, self-directed.

There are theoretical implications for positing the inter-related predictor variables outlined in this chapter. A trend in current linguistics is to take findings such as those in this chapter and argue that grammar is probabilistic. This approach, championed by researchers such as Joan Bresnan (Bresnan and Nikitina 2007), argues that syntactic and/or semantic factors can proportionately influence the choice of one particular construction over the other, and the mere existence of this hypothesized paradigm is an argument against the traditionally argued categorical nature of grammar. Implicit within this hypothesis is the idea that the various predictor variables are additive in nature, that is, the addition of each subsequent predictor variable more accurately predicts the likelihood of one particular outcome. For example, Bresnan, Cueni, Nikitina, and Baayen (2007) argue that which dative alternative is used (the double object construction versus the prepositional object construction) is predictable based on discourse factors such as the post-verbal phrase's referential specificity, definiteness, animacy, and accessibility. Following this trend, one could argue that each semantic factor outlined above (directionality, peri-personal space, and containerhood) is associated with a numerical likelihood that the pronoun will occur as reflexive or nonreflexive. In this hypothesis, the likelihood sum of all three factors would be higher than any one factor alone. The distributional totals outlined in this chapter do support this idea. For example it is shown in Table 4.8 that PPs headed by the prepositions which combine both directionality and containment (mostly *into*, *out of*, and sometimes *in*), pattern with the reflexive pronoun 96% of the time, compared to 52% of the time in all the other container contexts. Nevertheless, we do not want, in any way, to imply that the speaker has implicit knowledge of these predictor variables. The argument made in this thesis is that the predictive nature of these spatial variables is a byproduct

of the speaker's schematic understanding of reflexive events. That is, the more schematically similar a syntactically complex reflexive event is to a syntactically simple reflexive event (which serves as the prototype), the more likely the speaker will use the reflexive pronoun to mark the reflexive nature of the event. This type of process is what is argued for in theories of Cognitive Grammar (Langacker 2002).

Inherent in my claim is the theory that grammar is analogical in nature. The speaker uses some type of best-fit apparatus when choosing the one pronoun over the other. The more closely the event in question resembles the prototypical reflexive event, the more likely the speaker is going to use the reflexive pronoun (if needed) to mark the event as such. Since intrinsically reflexive, syntactically simple events like bathing in English are not obligatorily marked with the reflexive, intrinsically reflexive and syntactically complex events like pulling an object toward oneself do not need to be marked with the reflexive either. Other events, however, that are typically other-directed, need the reflexive pronoun to ensure a semantically reflexive meaning.

The third message of this chapter is the assertion that the reflexive pronoun as a lexical item is semantically rich. It is not a vacuous grammatical marker of function only; rather it is a lexical item that shares a close connection to the concept of self. Its use often activates a very detailed and metaphorical understanding of the self, and, we therefore see it used in contexts which describe internal events within the referent.

The fourth and final take-away message of this chapter is the claim that the same spatial factors that influence a proportionately higher use of the reflexive in English are also at work in Spanish. Crosslinguistically, it is argued that speakers of both languages use their understanding of syntactically simple reflexive events and the corresponding schematics when choosing which pronoun to utter in syntactically complex events.

Along with the evidence from Norwegian, it is clear that different language data yields similar relative distributional patterns, hinting at the possibility that the spatial nature in which speakers understand reflexivity as a semantic domain is universal. It is the manner in which reflexivity is marked that is language specific.

The data in this chapter unquestionably support the claim that spatial schematics influence the choice of pronoun in the PP. Nevertheless, spatial schematics alone do not fully explain the corpus data. In this chapter two important questions have been left unanswered. Why are intrinsically reflexive complex predicates marked with the reflexive pronoun? And, for that matter, why are intrinsically reflexive simple predicates sometimes marked with the reflexive as well? The study now turns to a supplemental area of explanation. Space is of course not the only primitive in grammar. As it will be shown in the next chapter, the various discourse functions of the reflexive marker need to be understood in order to further explain the use of the reflexive in the English PP.

Chapter 5

The semantics of viewpoint

5.1 Introduction

Chapter 4 left two very important, related questions unanswered. Both are summarized in the following question. Why are many intrinsically reflexive complex (and simple) predicates in English (redundantly) marked with the reflexive pronoun as in 5.1[24]?

(5.1) *John$_i$ pulled the ashtray toward himself$_i$.*

Section 4.11 revealed that the nonreflexive pronoun is more common in these cases, and the subsequent analysis maintained that inherently reflexive events do not need to be marked as such with the reflexive pronoun. In a similar vein, there is a limited set of corpus examples in which the event is not body-internal, not in the referent's peri-personal space, and not semantically reflexive at all, yet the reflexive pronoun appears anyway as in 5.2:

(5.2) *John wanted as much attention around himself as he could possibly have.*

The focus of this chapter is to explain why the reflexive appears in the data when its function is not to mark the event as reflexive and to explain the non-spatial factors that influence pronoun choice in the PP.

[24] Redundant (or double) marking in language is of course a common phenomenon. For example many Romance languages require double negative marking within the pronominal system. Here, we call attention to this phenomenon not to imply that it is rare, but to try to give an explanation for it in the English PP cases.

5.2 Background

For a long time, researchers working in both syntactic and semantic traditions have argued that the cognitive perspective from which a sentence is to be interpreted should be modeled in the syntax (Sells 1987; Tenny 1996, 2003) or in the semantics of language (Fauconnier 1997; Langacker 2002; etc. (see Sanders and Spooren 1997 for a review of analyses)). A subset of these researchers looking specifically at anaphoric distribution in English, have argued that the reflexive pronoun is associated with cognitive viewpoint (Cantrall 1974; Kuno 1987; Zribi-Hertz 1989; Tenny 1996, 2003). Heretofore, we can refer to the notion of cognitive viewpoint as *metaphorical viewpoint*, the cognitive perspective from which the utterance is understood, as opposed to *literal viewpoint*, the vantage point from which an event is literally perceived. The idea of metaphorical viewpoint itself has its basis in physical perception. When we directly reference someone's opinion in English we use language from the domain of physical perception (Lakoff and Turner 1999: Chapter 12), such as in 5.3:

(5.3) a. *From John's perspective, Mike's actions were really mean.*

b. *If you assume Mike's vantage point, you can see how he was just trying to do the right thing.*

In these cases, the speaker is not referencing John or Mike's physical perception of an event, the speaker is referencing John and Mike's mental construal of an event.

Cantrall (1974) argues that the reflexive pronoun serves as a clue to the listener that the antecedent's metaphorical perspective is active. In the following example, Cantrall (1974: 44) shows that the use of the reflexive on the part of the speaker indicates that the utterance should be understood from John's (the referent's) perspective:

(5.4) *Mary represents to John$_i$ some hope for himself$_i$.* (Cantrall 1974: 44 #19)
(John's viewpoint is assumed)

Cantrall further argues that the reflexive is only licensed grammatically because John's viewpoint is available and serves as an anchor for the proposition. In a contrasting example, where the speaker's viewpoint intervenes, the reflexive is no longer grammatical:

(5.5) *Mary represents to John$_i$ what I see as some hope for him$_i$/(*himself$_i$).* (Cantrall 1974: 44 #21)
(The speaker's viewpoint is assumed)

Once the speaker's perspective on the proposition is taken, the coreferential, nonreflexive pronoun is the only grammatical option. Certain mental action verbs such as *understand* guarantee that the subject's perspective on the proposition is assumed by the listener, thus allowing for the reflexive to again be grammatical as in 5.6:

(5.6) *John$_i$ couldn't understand why the responsibility should fall on himself$_i$/him$_i$.*
(John's viewpoint is assumed) (Cantrall 1974: 44 #8)

However, if the referent is physically removed from the situation described in the utterance, so is his anchoring viewpoint, and the reflexive is rendered ungrammatical as shown in the contrast between 5.7 (my own) and 5.8 (Cantrall's):

(5.7) ?*Since John$_i$ admitted he was the driver in the accident, and said he$_i$ was drunk, the responsibility fell on himself$_i$.*

(5.8) **Though John$_i$ wasn't there, the responsibility fell on himself$_i$.* (Cantrall 1974: 44, #14)

Similarly, if the speaker's opinion is at all inserted onto the proposition vis-à-vis a modal such as *should*, the reflexive sounds ungrammatical as in 5.9:

(5.9) **The responsibility for John's$_i$ mother should fall on himself$_i$.* (Cantrall 1974: 44 #15)
(John's viewpoint is unavailable or overridden)

Cantrall argues that even when there is no overt perspective marker (like a modal), the use of the reflexive pronoun portrays the speaker's concern for John's interest in 5.10 as opposed to the use of the nonreflexive, which does not convey this special semantic property:

(5.10) *John$_i$ let that log roll too close to him$_i$/himself$_i$ before he jumped.* (Cantrall 1974: 46 #40)

Thus, Cantrall believes that the speaker can use the reflexive in 5.10 to signal that he empathizes with John, to signal that the speaker wants the listener to understand the situation through John's eyes, so to speak. Kuno (1987: 206) calls this semantic property 'empathy': "Empathy is the speaker's identification, which may vary in degree, with a person/thing that participates in the event or state that he describes in a sentence."

In her analysis of long-distance reflexives, Zribi-Hertz (1989) terms the point-of-view anchor the "subject of consciousness" (p. 705) (following Banfield 1979). She argues that the reflexive is licensed within a discourse-based point of view domain, when the referent of the reflexive serves as the subject of consciousness and within that domain there is no other intervening point-of-view anchor. She formally defines the Domain of Point of View as "a portion of discourse which involves one and only one narrative point of view; in other words, a DPV may not include a switch of narrative viewpoint (p. 713)." She formalizes this idea in 5.11:

(5.11) Minimal S(ubject) (of) C(onsciousness) + either (a) or (b):

a. the nearest available NP or combination of NPs (split antecedent) which occurs in discourse to the left of the pronoun, and is read as logophoric.

b. the speaker or addressee, or a group including either one or both (cf plural pronouns), whether or not explicitly mentioned in discourse.

Tenny (2003) takes the same ideas in Zribi-Hertz (1989) (and in turn Banfield 1979) and argues that a discourse point-of-view domain should be modeled in the syntax. Tenny, building on the work of Sells (1987), makes reference to five types of sentience roles, which the grammar can reference independently. Tenny explains that these roles of sentience can be organized hierarchically from top to bottom: Speaker/Hearer, Evaluator, Experiencer, and Perceiver. "The higher the role on the hierarchy, the more accessible it is as a point of view anchor for the relevant linguistic elements. The lower it is on the hierarchy, the more local its 'domain' seems to be. With the speaker and the hearer at the top of the hierarchy, and the evidential argument below it, the hierarchy reflects syntactic structure (p.12)." In summary, according to both Zribi-Hertz and Tenny, the "referentially deficient" (their understanding) reflexive pronoun is only licensed in peripheral syntactic contexts when its referent serves as the point of view anchor. As Tenny's hierarchy indicates, utterances in full are most likely understood to be anchored from the speaker's perspective; however, the speaker can insert a secondary point of view as in 5.4, and if this secondary point of view is not overridden as in 5.5, it can serve as the referent to a reflexive pronoun as in 5.4.

Secondarily, the link between the reflexive pronoun and the referent's point of view has a personalizing effect in the second person. Van Hoek (1997:190) gives the following examples to show this special effect the reflexive has:

(5.12) a. *People like you disgust me.* (Van Hoek 1997:190 #35a)

b. **People like yourself disgust me.* (Van Hoek 1997:190 #35b)

c. *With friends like yourself who needs enemies.* (Van Hoek 1997:190 #35c)

She explains, "The politeness effect can be attributed to the fact that the addressee's role as conceptualizer, and the construal relation connecting the (conception of the) addressee with the profile of the reflexive, are made more salient, as these are evoked to sanction the use of the reflexive marker. This implication of a heightened awareness or consideration of the addressee's point of viewpoint is felt to be more polite (p.190)." Therefore, since the use of the second person reflexive indicates politeness on the part of the speaker, it is infelicitous for the speaker to convey politeness when the speaker is actually implying insult, as in 5.12b. On the other hand, politeness and complement, together in 5.12c, go hand in hand.

The aim of this chapter is to evaluate the ideas presented above through the lens of the real-world corpus data. Does the reflexive pronoun appear in the PP more frequently when the utterance is conveying the referent/character's point of view as opposed to the speaker or narrator's point of view? And, can the reflexive's hypothesized link to viewpoint explain the seemingly 'redundant' reflexive marking of inherently reflexive events? Analysis at the macro-semantic level will shed some light on these questions. Black and white answers will not come from the statistical data presented below. As it will be explained below, the lack of clear, categorical answers to these questions may function more as a testament to the complexity of point-of-view as a grammatical phenomenon rather than a refutation of the importance of point-of-view semantics on the distribution of anaphora in PPs and the distribution of anaphora in general. The categorical nature of macro-semantic analysis has its limitations, and one

such limitation is the inability to follow changes in point-of-view over long chunks of narration. The next section will present the mixed results of various categorical parameters related to point-of-view that can be applied to the corpus data including person marking, narration viewpoint, verb type, and metaphorical status. We will try to understand these results in reference to the hypothesis that the reflexive pronoun is associated with the referent's metaphorical point-of-view.

5.3 Viewpoint and person marking

A short digression is needed before evaluating the hypothesis that reflexives mark point-of-view anchors. The semantic literature on viewpoint has not included examples of first person reflexives. Because the BNC corpus data includes examples from both first and third person pronouns, we will want to see if perspective marking with the reflexive is different between these two discourse categories.

In the narration of written texts, authors tend to choose between two types of narrative perspective: a third person narrative can be either an *objective* (Banfield 1982) narration of events in which the narrator is not able to "get into the head" of the character(s) or a *subjective* narration of events, which presents the consciousness, thought, perception, or private state of the character(s) in the narration. A first person narrative is inherently *subjective*, since the narrator is the character and has automatic access to his or her own thoughts. Often the two styles are mixed within a single longer text. Both styles show up in the written BNC corpus data, and as discussed, the overall corpus results presented thus far have conflated examples of both first and third person pronouns (reflexive and nonreflexive) since we have made the assumption that this discourse parameter is not relevant to the spatial factors discussed in Chapter 4. Person marking, however, is very relevant to the discussion of viewpoint semantics and

its relation to anaphoric distribution in the PP. The act of narration involves several differentiable discourse/narrative roles and assignments; of importance here are the role of 'narrator', the source to which the information is attributed, the role of 'character' (or 'protagonist'), the entity performing the narrated events, and finally, the assignment of perspective anchor, which can reside with either the narrator or the character. In all the BNC examples, the referent of the pronoun maps to the character; that is, the pronoun's referent fills the character role. This identity relationship is expected since the only case in which the referent would not map to the character would be if the narrator were directly referencing the reader with the second person pronouns *you* or *yourself.* Also, as expected, in first person narration, the referent is, by default the narrator. Thus, at first, one might expect that first person narration would be associated with the character's point-of-view more often than third-person narration because in first person narration the referent is the narrator as in 5.13 and 5.14:

(5.13) *I had always been appalled by the ease with which men seemed able to separate sex and love and now here was* ***I building a high wall around myself*** *and saying, "I will let you into my body but stay out of my mind."* (CB8 1147)

(5.14) ***Holding the open gown tightly around myself****, I rise to adjust the volume on the television, so that it will not disturb them.* (HGL 459)

In both 5.13 and 5.14, the reader understands the narrator to be the same person as the perspective-anchoring character, and both roles map to the real-world referent as in Figure 5.15:

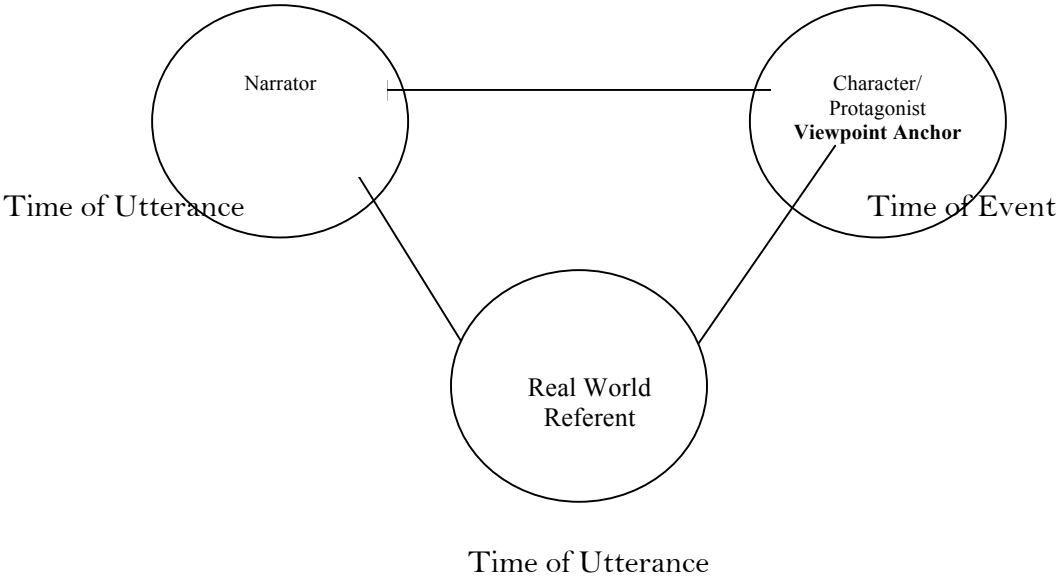


Figure 5.15 Character as viewpoint anchor

However, because the roles in narration are divisible and distinguishable, it is possible for the narrator and the character to map onto the same real-world referent while still being understood by the reader/listener as different entities. By looking at the details in first person example data we see that the complexities of narration allow the speaker to divide herself into narrator and character and narrate an event she has directly experienced from a distanced metaphorical perspective. Thus, in examples like 5.16, the pronoun's referent is the protagonist of the event, but the protagonist is not the perspective anchor for the event. In this case, the event is understood from the narrator's viewpoint, but not from the character's, even though both discourse roles link to the same person, the speaker/writer:

(5.16) ***I feel around me*** *and find a branch to sit on.* (APC 1105)

In 5.16, the narrator uses the historical present to describe to the reader the events that she, as a character experienced. In this type of narration, the speaker or writer separates the roles of narrator and character while simultaneously mapping on to both.

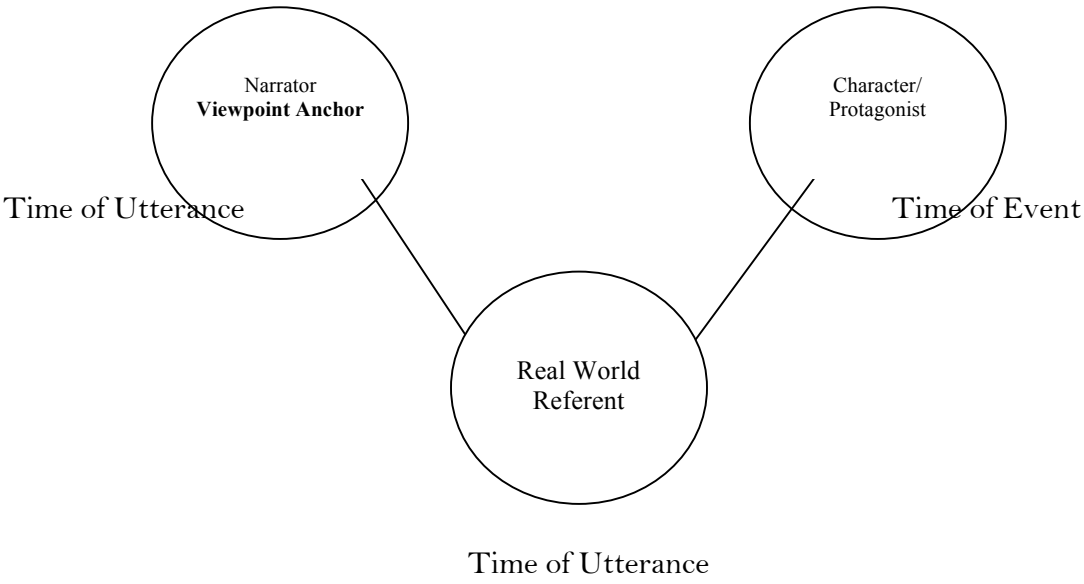


Figure 5.17 Narrator as viewpoint anchor

So, even though 5.16 is an example of first person narration, we see the same link between narrator and perspective anchor (versus character and perspective anchor) that is common in third-person narration. This alignment is important because it is, in some way, unexpected or surprising. It demonstrates a high level of complexity in the way speakers and listeners understand narration, and demonstrates an interesting and cognitively complex ability that speakers have: to 'step outside' themselves when describing events in which they have participated directly.

If, for now, we accept the hypothesis that the reflexive signals the character's perspective, we can look to see if the reflexive is used for this purpose within the two

different person-marking groups since it has now been shown that perspective marking is relevant to third and first person examples. Table 5.1 shows in fact no difference between first and third person example data ($\chi^2(1, N = 1667) = 0.01, p >$ or $= 1$):

Pronoun Distribution by Person

	First	**Third**	**Total**
Nonreflexive	344	1021	1365
Reflexive	77	225	302
Total	421	1246	1667

Pronoun Distribution by Person

	First	**Third**	**Total**
Nonreflexive	82%	82%	83%
Reflexive	18%	18%	17%
Total	100%	100%	100%

Table 5.1-Pronoun distribution according to person marking

In both first and third person example data, the reflexive appears in 18% of the examples. To reiterate, our point of this digression has been to show that perspective marking is relevant to examples with both first and third person pronouns. From now on, we will assume that any tie made between metaphorical perspective and pronoun choice holds for both grammatical contexts.

If the reflexive pronoun serves as a point-of-view indicator (as hypothesized in the semantic literature outlined above) in contexts in which often two viewpoint construals (character perspective vs. narrator perspective) are possible, a distinction should be present between the appearance of the reflexive when the narrator/speaker's point of view is assumed versus when the referent/character's point of view is assumed. The next section attempts to quantify this difference.

5.4 Directly testing for viewpoint

Chapter 4 illuminated a subset of corpus data that seemed to be redundantly marked with the reflexive pronoun. These were cases in which intrinsically reflexive complex predicates occurred with the reflexive pronoun. Chapter 4 also suggested that some nonreflexive events occur with the reflexive pronoun. For statistical purposes, it is hard to group these two types of data. They do not form a unique category that was coded in the initial data coding process, and because these data are rare in the corpus, they don't constitute a large enough group for any type of statistical analysis. For example, there were only 5 examples of the reflexive pronoun occurring in noncontainer, non-peri-personal, and non-directional contexts (versus 278 of nonreflexive). Nevertheless, we can conjecture that the occurrence of the reflexive pronoun may be explained in some of these cases by factors related to viewpoint if we can prove that the reflexive pronoun is a signal that the character serves as the viewpoint anchor.

The effects of viewpoint on pronoun distribution can be teased out of the pre-existing grouped data. Of all the spatial contexts, the reflexive pronoun patterns highest among the prepositions which denote containment. In Chapter 4 it is argued that this spatial context evokes a metaphorical understanding of the self that would yield a high percentage of reflexive pronouns independent from any type of perspective effect. This metaphorical understanding seems a possible confounding factor, raising the following question: how can one test for the effects of perspective marking in a data set in which the reflexive has an 'independent' reason to occur? The answer and justification for the selection of this data group is twofold. First, in order to make sure that perspective effects are targeted within this data set we can look for relative

differences; that is, we expect an even higher percentage of container examples to use the reflexive pronoun if the character/referent's perspective is marked (as opposed to container examples in which the narrator's perspective is assumed). Second, by virtue of the fact that the referent's internal space is always evoked in the container data, many of these examples describe the referent's internal feelings and thoughts. These are contexts in which the reader is often guided to taking on the character's (as opposed to the narrator's) perspective.

Again, the hypothesis referenced in the semantic literature predicts that the reflexive will occur more frequently when reader is assumed to channel the referent/character's perspective during the narration than when the reader is assumed to take on the narrator/speaker's perspective during the narration. Coding for this type of difference in viewpoint is not an exact science, and thus about 20% of the examples were judged to be indeterminate as to whose perspective was most likely intended. Viewpoint assignment was determined based on my own interpretation of whose perspective was intended, and to further limit the data, only third-person examples were looked at.

My own diagnostic for referent viewpoint included passages (a minimum of two sentences) in which the referent's senses, emotions, or mental state are described as in 5.18 and 5.19 and examples in which the referent's memory introduces the propositions as in 5.20:

(5.18) *Jack could see hard frost gleaming on the windows and* ***he$_i$ could feel the bleak coldness of the beds inside himself$_i$.*** (AF5 539)

(5.19) ***He$_i$ had learned to live in himself$_i$*** *and be self-sufficient. Instead of feeling Kee's presence as an intrusion, upsetting his concentration at a critical stage, he was inspired by her.* (CBN 628)

(5.20) ***He$_i$ remembered*** *sitting at his desk in the bothy before Bicker arrived,* ***PRO$_i$ trying to tear the story out of himself$_i$*** *and filling pages with blizzards and killing.* (GWF 1208)

On the other hand, examples such as 5.21, in which a generic protagonist serves as the pronoun's referent, were coded to represent the narrator's perspective:

(5.21) ***The researcher$_i$ who is really at home as one of the natives must seek inside himself$_i$*** *to create his `remote areas'. (E. Ardener 1987).* (A0K 81)

Likewise, periphrastic commentary, when used to indicate the narrator's judgment (e.g. *It seems* …), as in 5.22, served as a diagnostic for narrator perspective:

(5.22) ***It is beginning to look as if he$_i$ also carries within him$_i$*** *a seed of self-destruction.* (CEP 4673)

Other examples required more context to show that the reader was intended to understand the viewpoint as coming from a perspective other than the antecedent character's, the perspective of another character (named *Berdichev* in this case) as in 5.23:

(5.23) *So it was with the boy. He would be the means through which the Seven would be destroyed; not as Berdichev had imagined, from without, but from within. The Seven would be the agents of their own destruction. For* ***the boy$_i$ carried within him$_i$*** *not a ball of wax, but an idea. One single, all-transforming idea. DeVore sat back. Yes, and Li Yuan would fight to preserve the boy, for he honestly believed that he could control him.* (G04 3740)

After coding for viewpoint just among third-person container examples (again, to limit the time-intensive coding process), the reflexive pronoun was found to occur more frequently in contexts in which the antecedent character's perspective is assumed as opposed to contexts in which the narrator's perspective (or another character's perspective) is assumed ($\chi^2(1, N = 147) = 20.45, p < .001$):

Pronoun Distribution in Container Preps by Viewpoint for 3rd Person Examples

	Narrator VP	Character VP	Total
Nonreflexive	49	12	61
Reflexive	37	49	86
Total	86	61	147

Pronoun Distribution in Container Preps by Viewpoint for 3rd Person Examples

	Narrator VP	Character VP	Total
Nonreflexive	57%	20%	41%
Reflexive	43%	80%	59%
Total	100%	100%	100%

Table 5.2-Pronoun distribution based on narrator vs. antecedent-character perspective

The reflexive pronoun occurred 80% of the time when the proposition was judged to come from the antecedent character's perspective as opposed to 43% of the time when the proposition was judged to come from the narrator or another character's perspective. (We can assume that the reflexive shows up in the later group because of other spatial schematic reasons (detailed in Chapter 4).) The corpus confirms that the reflexive pronoun patterns more often when the character as opposed to the narrator/speaker serves as the perspective anchor; however, the distribution is not as black and white as is implied by the researchers discussed in section 5.2. The nonreflexive pronoun still occurred in 20% of the cases judged to represent the antecedent-character's perspective, indicating that the reflexive pronoun should not be thought of as a requirement to mark character perspective. Other language tools and overt description of perspective can, of course, carry the same function. Thus, the data confirms a *tendency*, not a rule, for the reflexive pronoun to be associated with the character's point-of-view. (See reference to the theoretical implications of distributional tendencies discussed in Chapter 4.)

As discussed earlier in regard to the container examples, emotive contexts often trigger the reader/listener to take the character's perspective. Because emotions are metaphorically located in the container-construed body (Lakoff and Johnson 1980), there is an assumed understanding that individuals (as opposed to outsiders) have unique access to their own emotional states, and only the individual can validate the accuracy of any description of his feelings. In the next section, this idea will be explored through the categorization of lexical verb type and its effect on anaphoric distribution. The goal of this section is to support the reflexive-perspective link by identifying indirect variables in the corpus data that buttress the argument.

5.5 Indirect indicators of a link between pronoun choice and metaphorical viewpoint

Beyond direct coding for viewpoint anchoring, there are indirect indicators in the BNC data to support the notion that one purpose of the reflexive is to signal character viewpoint. In this section two indirect measures of this parameter will be explored: verb semantics and metaphorical context. Certain categories of events (such as mental actions) are more likely to be understood via metaphor. Such is the case with emotive verbs. Emotive verbs comprise a category of verbs that describe the emotional state of the referent. Human emotion is a concept structured by multiple metaphors including EMOTIONS ARE FLUIDS IN THE BODY (Lakoff and Johnson 1980), resulting in descriptions of emotions like anger that come from the source domain of hot liquids: *boiling*, *brimming*, *blow your top*, etc. Human thought is also a concept structured by metaphor. Thought is not described by most or understood literally as neuronal firing (as it is neuroscience), but rather metaphorically vis-à-vis the source domain of physical

movement through a landscape (the mind). (See Lakoff and Turner 1999 for details of this complex set of metaphors.) As we alluded to in the previous section, accurate detailed description of emotion and thought is most often assumed to come from an internal perspective. That is, listeners/readers understand detailed description of these mental processes by assuming the perspective of the character, not the perspective of the narrator. Likewise, the narrator evokes the character's perspective of events by describing how the character thinks and feels about events within the storyline. Therefore, description of emotion and thought can be used as an indirect indicator of character over narrator perspective. And because these descriptions are often metaphorical, metaphor activation, secondarily serves as an indirect indicator of character perspective.

5.6 Verb types

Although not directly implicated as a potential governing factor for pronoun choice in the literature, the English corpus was initially coded along the parameter of lexical verb type. The impetus for this parameter was to examine if broad semantic divisions within the verb alone had any effect on pronoun distribution. The data outlined in Chapter 4 indicate that the semantics of the complex predicate as a whole is a predictor of pronoun distribution. By separating the data into four categories of verb type, it is possible to test the idea that lexical semantics alone or lexical semantics in tandem with spatial semantics may factor into the appearance of one pronoun versus the other in the adjoining PP. These four categories (Emotive, Sensory, Stative, and Motion) are extremely broad classifications, each encompassing a myriad of semantically distinct subcategories. The motivation for the creation of only four general categories is due to

practical factors. With a limited set of data, each category must include a substantial number of examples in order to do cross-category statistical comparisons. Having many small semantic categories in which only a few examples fit would make statistical analysis unreliable.

The first semantic category coded for is labeled "Sensory" (SEN). In these examples the verb refers to the human sensory system and the perception of objects; included in this general category are verbs of perception (*see* verbs, *sight* verbs, and *peer* verbs) as categorized in Levin (1993: Ch. 30) and other sensory actions such as hear and feel. The following sentences have been put into the SEN category:

(5.24) *He$_i$ failed to notice the borrowed things that the girls wore,* ***PRO$_i$ looking around him$_i$*** *instead in dumb bafflement: it was a wedding day, a shining moment in his life, and, except for the dressed children, it could be any ordinary day.* (A6N 1071)

(5.25) *As he waits at a stop-light somewhere out beyond the freightyards* ***he$_i$ drums his fingers on the steering wheel and gazes in front of him$_i$****, thinking.* (G0F 3330)

(5.26) *Soon the mist grew thicker and* ***he$_i$ could see only a few feet in front of him$_i$****.* (H9U 863)

(5.27) ***He$_i$ felt arms around him$_i$****, lifting his body up, easing him off the hook.* (ALJ 972)

It should be noted that Sensory verbs have been put into this category even when the verb is used metaphorically; for example verbs of visual perception are often used metaphorically for psychological introspection (KNOWING IS SEEING (Lakoff and Johnson 1980)). The general process of categorization outlined here is meant to organize the examples by verb choice alone, the categorization is not sensitive to the verbs exact meaning due to the context in which it appears.

The second semantic category coded for is labeled "Movement" (M). This category is by far the most general, including a multitude of distinct subcategories. The only central criterion used for inclusion in this category is that the verb, in its basic

sense (Levin 1993: 22), involves some type of motion on the part of the antecedent. This motion can be long distance and lacking in detail as in a verb which implies the protagonist has, at some point, physically changed locations as in *find*, or the motion can be self-contained and specific as in *reach.* An unexhaustive list of verbs included in this category are verbs of putting (e.g. *put*, *place*), verbs of exerting force (e.g. *push*, *pull*), verbs of throwing (e.g. *throw*, *toss*), verbs of contact by impact (e.g. *bump*, *hit*), verbs of creation and transformation (e.g. *make*, *build*, *change*), meet verbs (e.g. *meet*, *gather*), verbs of gestures (e.g. *wave*), some verbs of change of state (e.g. *become*), verbs of modes of being involving motion (e.g. *walk*, *run*), verbs of body internal motion (e.g. *reach*), verbs of assuming a position (e.g. *sit*, *stand*), and general verbs of motion (e.g. *move*) as categorized in Levin (1993). The following sentences are examples which have been put into the M category:

(5.28) ***He$_i$ retreated into himself$_i$*** *or, rather, into the past.* (BNN 1896)

(5.29) *He$_i$ sat back in the bed,* ***PRO$_i$ pulling the sheets up around him$_i$.*** (EVG 2356)

(5.30) ***He$_i$ wraps around himself$_i$*** *the dusty patchwork cloak of his invisibility and the heralds gallop hither and yon in vain.* (ECU 2345)

(5.31) *Often he$_i$ seemed desperate, violent, as if* ***PRO$_i$ thrashing around inside himself$_i$*** *for certainties that weren't there.*(F99 62)

Just as described for examples which fall into the Sensory category, many examples containing verbs included in the Movement category use movement verbs to describe metaphorical actions which may not involve physical movement. Nevertheless, these examples have still been categorized in the Movement category based on the basic semantics of the verb in its literal sense.

The third semantic category coded for is called "State" (S). In these examples the basic semantics of the verb describe a stative event that does not involve movement.

An unexhaustive list of verb types included in this category are Levin's verbs of spatial configuration (e.g. *surround*), measure verbs (e.g. *count*), and fit verbs (e.g. *fit, belong*). Typical stative verbs not included in this category, but rather included in their own category are some psych verbs or verbs of cognition such as *like.* The following sentences are examples from the English corpus which have been put into the S category:

(5.32) *Despite that* ***he$_i$ has managed to keep a group of people around him$_i$*** *who have a loyalty both to him and to the company.* (K59 1132)

(5.33) *While the forge is always warm and inside is the loving, childlike* ***Joe$_i$ who$_i$ has goodness all around him$_i$*** *and this is particularly shown when he looks after Pip when he has a fever and the power of that scene.* (KA1 1091)

The fourth and final semantic category coded for is called "Emotive" (E). In this limited number of examples the semantics of the verb describe a psychological state involving the emotions. Included in this category are admire verbs (e.g. *like*), marvel verbs (e.g. *wonder*), and verbs of desire (e.g. *want, need*) as categorized in Levin (1993). The following have been put into the E category:

(5.34) *And* ***he$_i$ began to think within himself$_i$*** *what had happened, and of that breath which had passed through him, and how the leper was not there.* (ASW 528)

(5.35) ***He$_i$ debated, within himself$_i$****, whether to be oblique or direct with Peter, whose face seemed to him both shuttered and vulnerable.* (CMJ 1661)

The results of the statistical comparison (shown in Table 5.3) demonstrate that the only verbal category which yields statistically significant higher rates of the reflexive pronoun are verbs of emotion ($\chi^2(1, N = 1667) = 33.76, p < .001$).

Pronoun Distribution by Semantic Category of Verb

	Emotive	Motion	Stative	Sensory	Total
Nonreflexive	9	699	349	308	1365
Reflexive	15	156	75	56	302
Total	24	855	424	364	1667

Pronoun Distribution by Semantic Category of Verb

	Emotive	Motion	Stative	Sensory	Total
Nonreflexive	38%	82%	82%	85%	83%
Reflexive	62%	18%	18%	15%	17%
Total	100%	100%	100%	100%	100%

Table 5.3-Pronoun distribution based on semantic category of verb

The reflexive pronoun appears in 62% of the examples which involve verbs of emotion. In the other semantic categories (motion verbs, stative verbs, and sensory verbs), the reflexive patterns in the same ways as it does in the overall total distribution; that is, these semantic categories do not yield higher percentages of reflexive pronouns. The fact that the verbs of emotion comprise a smaller group of data should be noted. Although this group is large enough for a statistical comparison with the other more sizable groups, its relative size indicates that these results should be viewed with caution. Nevertheless, this skewed distribution is what one would expect. As we discussed at the opening of this section, verbs of emotion are used in contexts in which the character's feelings and opinions are being described. Quite often it is assumed that this mental description is coming from the character's perspective. Thus, if accepted as a representative sample, the results shown in Table 5.3 serve as a further clue that the reflexive is associated with the referent's viewpoint.

Emotive contexts are often metaphorical as well. The next section will explore the role of metaphorical construal as indirect support for the reflexive-perspective link. The activation of metaphors may also serve as indirect evidence in favor of the notion

that the character's perspective is associated with the a higher occurrence of the reflexive pronoun since emotion and thinking are described via metaphor.

5.7 A link between metaphor and viewpoint

As discussed in section 4.3, all the container examples in the corpus activate a metaphorical understanding of the human referent as a container which houses his internal self along with his emotions. Likewise, included among many of these examples, is the common metaphor that THE MIND IS A CONTAINER (Lakoff and Johnson 1980) (a subpart of the "Conduit Metaphor" for communication (Reddy 1979)). The activation of these types of metaphors is what allows the listener/reader to understand examples such as the following in 5.36-5.39:

(5.36) *In the last chapter, in prison, he says that without the crime* ***he$_i$ would not have found within himself$_i$*** *such questions, desires, feelings, needs, strivings, and development.* (A18 201)

(5.37) *If* ***I find in myself*** *a desire for which no experience in this world can satisfy, the most probable explanation is that I was made for another world.* (ARG 1474)

(5.38) *So far as* ***Andrew$_i$ was able to think beyond himself$_i$****, Nicandra's beauty overpowered him.* (H7H 1820)

(5.39) *I knew* ***I needed some place to go inside myself*** *-- frighten people off -- so I shoved three tapes and my Walkman in the middle of the bag beside the big fat envelope where I'd squeezed all my dreams.* (BMS 231)

The metaphorical understanding of a person's mind as a container for his thoughts and emotions involves the notion that other people do not have the same access to these internal "contents" to which the owner himself has. The body serves as an exterior barrier and outsiders cannot feel or understand what goes on "inside" another person's body. Because of this common conception, the description of internal thoughts and emotions is understood to come from the possessor himself. Thus, in many of the metaphorical contexts evoked in the corpus data, the activation of the metaphors

involved coincides with the assumption that the propositions described are coming from the character's perspective as shown in the same examples above in 5.36-5.39. On the other hand, simple spatial descriptions which are not metaphorical do not require that the propositions described are coming from the character's perspective. The examples in 5.40-5.44 do not evoke metaphors for the self and are not metaphorical in nature at all, but are rather construed as simple, literal spatial configurations:

(5.40) *Then he$_i$ heard a sound behind him$_i$.* (EF1 1044)

(5.41) *He$_i$ rushed from the cell, PRO$_i$ shutting the door behind him$_i$.* (FAB 3273)

(5.42) *He$_i$ was smiling, but holding a gun out in front of him$_i$.* (HTY 2134)

(5.43) *He$_i$ put the opened bottle down next to him$_i$ and smelled the top.* (CA3 641)

(5.44) *He$_i$ could have stuck pins into himself$_i$ and it would have taken ten seconds for his body to complain.* (FSP 2109)

On the gross semantic level, it can be assumed, then, that metaphorical contexts are to some extent correlated with the referent's perspective, and the activation of metaphors in general serves as an indirect indicator that the sentence is assumed to come from the referent's perspective. The results of coding third-person examples as "metaphorical", i.e. activating a Conceptual Metaphor, or "nonmetaphorical," i.e. interpreted literally and not activating a Conceptual Metaphor, is shown in Table 5.4:

Pronoun Distribution in 3rd Person Contexts by Metaphor Activation

	Metaphorical	**Nonmetaphorical**	**Total**
Nonreflexive	245	776	1021
Reflexive	165	60	225
Total	410	836	1246

Pronoun Distribution in 3rd Person Contexts by Metaphor

	Metaphorical	**Nonmetaphorical**	**Total**
Nonreflexive	60%	93%	82%
Reflexive	40%	7%	18%
Total	100%	100%	100%

Table 5.4- Pronoun distribution in metaphorical and nonmetaphorical contexts

In metaphorical contexts, the reflexive pronoun occurs at a much higher rate ($\chi^2(1, N = 1246) = 203.27, p < .001$), occurring 40% of the time as opposed to only 7% of the time in nonmetaphorical contexts. If the referent's perspective is being channeled more frequently in metaphorical contexts, as argued above, then this set of data serves as indirect support for our reflexive-perspective hypothesis.

5.8 Summing Up

In sections 5.3 -5.7, we have shown direct and indirect evidence to support the claim that metaphorical viewpoint plays a role in the distribution of anaphora in the PP. The evidence given is mixed and the results of the various coding procedures in this section are far from 100 percent conclusive. But, each motivating factor we have tested for (narrator vs. referent perspective, verb type, and metaphorical context) provides new evidence in support of previous researchers' hypotheses as outlined in 5.2. Taken as a whole, the data supports a tendency for the reflexive to occur when the antecedent-character's viewpoint is assumed.

Previous research on this hypothesis provides considerable data to support it, and the corpus data discussed in this chapter add to that collection. What seems to be missing from the analysis as a whole is an adequate explanation for why the reflexive pronoun serves as a signal of the antecedent-character perspective. We have now established that metaphorical perspective is linked to a higher probability of the reflexive pronoun, and certain spatial schematics are linked to a higher probability of reflexive use. The following question now arises: are space and perspective linked, and if so how? In the next section we propose an answer to this question. Once the link

between space and perspective is understood, a natural explanation arises as to why the reflexive is used to signal referent viewpoint.

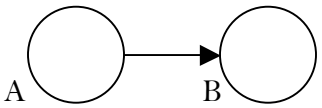
5.9 Viewpoint, the PP, and the core data

In Chapter 4 we saw that syntactically simple reflexive constructions share many spatial characteristics with the PP constructions that yield high percentages of reflexive pronouns. We categorized reflexive events into two types: intrinsically marked reflexive events and extrinsically marked reflexive events, and showed how these two groupings encompassed both syntactically simple (basic clauses) and syntactically complex (clauses with prepositional phrases) constructions in a relational network of the type proposed in theories of cognitive grammar (Langacker 2002; VanHoek 1997). In this section, we will outline how syntactically simple reflexive constructions share viewpoint semantics with the PP data. The link between space and metaphorical perspective comes in the form of a proposed set of participant roles. I claim that these schematized, abstracted roles often conflate the two dimensions discussed thus far: energy transfer and metaphorical perspective, and therefore, serve as an explanation as to why the reflexive is associated with aspects of these two features.

As discussed in Section 4.6 and 4.7, reflexive events have been characterized as a movement of energy, involving one participant only; whereas, nonreflexive events have been characterized as a transfer of energy, an action chain, from one participant to another (Kemmer 1993, Langacker 2002, Maldonado 1999, Talmy 2000). In cognitive theories of grammar, event structure is often schematized in image format, and the very basic structure of what often appears in a transitive clause can be represented schematically as done by Kemmer (1993: 50). She presents the action chain diagram in

5.45 to represent a prototypical transitive event in which the 'initiator', A, initiates the event to a second participant, B, who is the target or 'endpoint' of that event. The transitive event action chain differs from a reflexive event in that, in a nonreflexive action, two separate entities fill the roles of A and B:

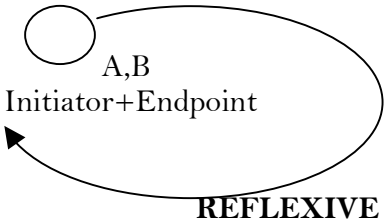
(5.45)



Initiator, Endpoint

TRANSITIVE EVENT

(5.46)



REFLEXIVE EVENT

Within the semantic literature, many different names have been given to the semantic role exhibited by A: source (Kuno 1987), agent (Payne 1997), actor (VanValin 2004), Initiator (Kemmer 1993, Maldonado 1999), etc. Likewise, different names have been assigned to the semantic role of B: target (Kuno 1987), patient (Larson 1984), undergoer (VanValin 2004), Endpoint (Kemmer 1993, Maldonado 1999). In the case of active transitive, (and non-specificational) clauses, all researchers, regardless of semantic theory, would agree that there is an asymmetry between the roles of A and B in the clause. Across different verb types, at a gross semantic level, we should notice a common quality among all the A participants in the following examples and among all the B participants in the following set (which differ in their syntactic structure):

(5.47) *John kicked Bill$_j$/him$_j$.*

(5.48) *John loves Bill$_j$/him$_j$.*

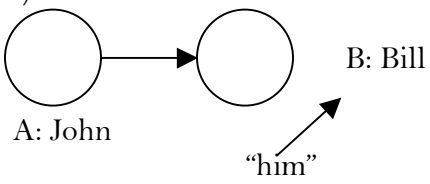
(5.49) *John hates Bill$_j$/him$_j$.*

(5.50) *John stuck needles into Bill$_j$/him$_j$.*

(5.51) *John wrapped the rope around Bill$_j$/him$_j$.*

In every case in 5.47-5.51 John is understood as a source of the action which is ultimately affecting Bill, who serves as the target of the action, or the recipient or endpoint of the energy transfer. When Bill is referenced as *him*, as is often the case, the nonreflexive pronoun references the B participant and is associated with the target/recipient/endpoint role. Thus, when the nonreflexive pronoun is used in direct object position, it references the event participant who fills the target role in the energy transfer.

(5.52)



On the other hand, in the prototypical reflexive construction, the reflexive pronoun is associated with the A participant. The same sentences in 5.47-5.51 are repeated as reflexive actions in 5.53-5.57:

(5.53) *John kicked himself.*

(5.54) *John loves himself.*

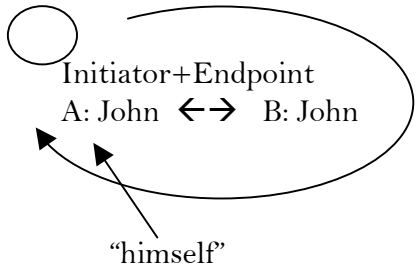
(5.55) *John hates himself.*

(5.56) *John stuck needles into himself.*

(5.57) *John wrapped the rope around himself.*

In these cases, the reflexive pronoun functions to ensure that the target of the action will be construed to be the same entity as the source of the action, as shown in 5.58.

(5.58)



The reflexive pronoun lies in an object position, just as the nonreflexive pronoun does in 5.47-5.51, but the reflexive *himself* references the source of the action, which in the cases above, sits in the subject position. In these reflexive events there is no separate entity which serves as the target of the action because the same participant fills both roles. We have now established a clause-internal difference between reflexive and nonreflexive pronouns. Reflexives reference the source (and target) in the energy transfer dynamic, and the ultimate reference establisher is the source of the action; whereas nonreflexives can only reference targets in the energy transfer dynamic of a finite clause. To be clear, nonreflexives, in other syntactic configurations can and do reference semantic sources. Mikkelsen (personal communication) points out that a nonreflexive can reference an energy source in an embedded non-finite clause such as in 5.59a or an energy source in another clause as in 5.59b:

(5.59) a. *I watched him kick the chair.*

b. *Mike$_i$ kicked Mary. I was so mad that I hit him$_i$ back.*

Although the nonreflexive does reference an energy source, an A participant, the source of kicking in 5.59a, it simultaneously, and obligatorily references an energy target, a B participant, of a different event (the target of watching in 5.59a). The same holds true

for a nonreflexive that has an extra-sentential referent in 5.59b. For example even though Mike is an energy source of the kicking event in the first sentence, he is an energy target of the hitting event in the second sentence. Based on these patterns of reference, I argue here that the reflexive pronoun has a stronger intra-sentential semantic link to the A participant than does the nonreflexive pronoun; the nonreflexive has a strong intra-sentential semantic link to the B participant. By introducing this hypothesized link between energy sources and the reflexive pronoun, we can now further hypothesize that the semantic specification of the reflexive includes the role of 'source'. The nonreflexive pronoun would not be heavily linked to the 'source' role. In cases where it does reference the source of one event, it also must reference the target of another, and it must reference a target in order to be introduced into the discourse as an object. (The deictic use of the nonreflexive seems to neither reinforce a relationship with event targets nor event sources.)

If the reflexive is primarily linked to sources and the nonreflexive always references a target (even if it simultaneously references a source), one should be able to test for an asymmetrical association between the two pronouns' reference resolution capacities. That is, if speakers associated the reflexive at a deep level with energy sources, and the nonreflexive, at a deep level with energy targets, this association should show up when listeners are forced into reference resolution based on ambiguous contexts in which either pronoun can reference either role.

This evidence (for the link between sources and reflexives) is suggested in reference resolution experiments performed on picture noun phrases containing either a reflexive or nonreflexive pronoun (Kaiser, Runner, Sussman, and Tanenhaus 2004).

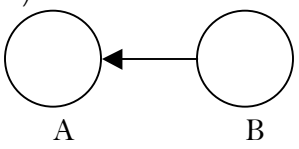
The contrasting semantics in the following sentences tested in these experiments demonstrate the link between the reflexive and the source role:

(5.60) a. **Peter** *told* Andrew about the picture of **himself** on the wall. (Kaiser et al 2004:5)
(source)
(subject)

b. Peter *heard* from **Andrew** about the picture of **himself** on the wall. (Kaiser et al 2004:5)
(subject) (source)

In 5.60a Peter is both the *source* of the information in the predicate and fills the subject position in the sentence; however in 5.60b, Peter fills the subject position of the sentence but is the *recipient* of the information. When subjects were asked to infer the referent of the reflexive pronoun in the sentences in 5.60, they exhibited an overall preference to interpret Peter as the referent in 5.60a, where he serves as the source of the information, as opposed to 5.60b where he serves as the recipient of the information. In each condition, the subjects linked the referent to the subject position much more readily than to they did to the nonsubject position; however, they were more likely to do so when the referent was both the subject and the source of the information as in 5.60a. These results indicate that the semantic specification of reflexive pronouns is not only characterized by a strong tendency to indicate coreference with the grammatical subject in these cases, but is secondarily characterized as the 'source' of the action. For the large majority of transitive verbs, such as *tell*, the subject position will coincide with the source of the energy transfer, thus strengthening this connection. Some verbs, however, such as *hear* or *hear from*, display an incongruent alignment of subject and source. In these cases, the following diagram (represented more elaborately in Talmy 2000: 115) shows the situation:

(5.61)



In 5.61, the energy transfer initiates with the B participant and the A participant serves as the endpoint. For verbs like *hear from* as in 5.60b, the nonsubject position aligns with the source role, thus creating a conflict between the grammatical position and the semantic role most typically associated with that position. Subjects are most often sources, but the configuration in 5.60b violates the prototype.

If the nonreflexive pronoun is associated with the 'recipient/target' role (the B participant in the prototypical transitive structure (in 5.47-5.51)), then subjects should more readily conclude that Andrew is the referent of the pronoun in 5.62a and Peter is the referent of the pronoun in 5.62b:

(5.62) a. Peter *told* **Andrew** about the picture of **him** on the wall. (Kaiser et al. 2004: 5)
(recipient)
(nonsubject)

b. **Peter** *heard* from Andrew about the picture of **him** on the wall. (Kaiser et al. 2004: 5)
(recipient) (nonsubject)

In this second condition, these predictions do hold true. The subjects were more likely to interpret Andrew as the referent of the pronoun in 5.62a, where Andrew sits in an object position and serves as the *recipient* of the information than they were to interpret Andrew as the referent of the pronoun in 5.62b, where Andrew sits in an object position but serves as the *source* of the information.

Although their experiments were not designed to test energy relations among event participants (but rather to help explain (a-theoretically) reference resolution

strategies), Kaiser et. al.'s experimental results provide further evidence that the reflexive pronoun is more robustly linked to the source of energy transfer and the nonreflexive pronoun is more robustly linked to the target of energy transfer. A likely reason for this association is that the reflexive pronoun, in reflexive events, is coreferent with the A participant who serves in the majority of events as the source of the transfer, and the nonreflexive pronoun, in the prototypical use is coreferent with the B participant who serves in the majority of events as the target or recipient of the energy transfer.

The 'source' category, like many abstracted linguistic categories, seems to have a gradient structure. Talmy (2000a: 275) unknowingly provides evidence for this idea when he hypothesizes that a reflexive form suggests a more direct connection between two references and a single entity. He uses the following examples to make his case:

(5.63) a. *The Pharaoh$_i$ built a pyramid for himself$_i$/*him$_i$.* (Talmy 2000a: 275)

b. *The Pharaoh$_i$ had a pyramid built for himself$_i$/him$_i$.* (Talmy 2000a: 275)

c. *The Pharaoh$_i$ had his subjects build a pyramid for *himself$_i$/him$_i$.* (Talmy 2000a: 275)

d. *The Pharaoh$_i$ told the King about his plans for the building of a pyramid for himself$_i$.*

He argues that there is more immediacy between the initiator-outcome relationship in 5.63a with the reflexive, whereas the nonreflexive form in 5.63c suggests a more distant connection between two references to a single entity: a greater causal distance between the initiator and the final outcome. Stating that there is no syntactic justification for the use of either pronoun in 5.63b, he concludes that 5.63b is evidence for the gradient nature of the link. Another way to characterize Talmy's findings is to claim that in 5.63a the Pharaoh is the source of the energy transfer, the source of the pyramid-building event (licensing the use of the reflexive pronoun), but in 5.63c, his subjects are

the source of the energy transfer and the Pharaoh is not the direct source of the event. 5.63d (my own example) reminds us that the reflexive is permitted in a position far away from its antecedent, as long as the antecedent (*The Pharaoh*) is still understood to be the source of the building event. Thus, it is not the case that the antecedent in 5.63c is too far away from the reflexive to license the reflexive; we argue here that the reflexive is not licensed because its potential antecedent is not the direct source of the building event. Under the assumption that the reflexive is felicitous when the Pharaoh is understood to be the builder, 5.63b is grammatical with the reflexive since the Pharaoh is still somewhat the source of the building event because he instigates the event and no other direct source is implicated. These examples as a set illustrate the gradient quality of the abstracted source role.

The special status reflexives share with energy sources provides an explanation for certain pronoun patterns in PPs embedded in NPs as well. Nichols (2006) demonstrates that only reflexives are grammatical inside complex NP objects of verbs of creation (i.e. *paint, write, make*, etc.) as in 5.64:

(5.64) a. *Mary$_i$ painted a picture of *her$_i$/herself$_i$.* (Nichols 2006: 2)

b. *John$_i$ wrote a book about *him$_i$/himself$_i$.* (Nichols 2006: 2)

(5.65) *Mary$_i$ saw a picture of her$_i$/herself$_i$.* (Nichols 2006: 2)

(5.66) *Mary$_i$ saw a picture of her$_{??i/j}$/herself$_i$ that she$_i$ had taken the other day.*

Nichols argues that when the antecedent is responsible for the act of creation in 5.64, the short-distance (nonreflexive) pronoun is disallowed. The semantic context of creation (in 5.64a and 5.64b) highlights the subject position as the source of the product. The energy transfer starts with the painter or writer and is directed toward the painting

or the book. The nonreflexive pronoun does not have a robust connection to sources, rather its primary link is to semantic targets, and thus its use in 5.64 implies its referent is not the source of the created product, thus warranting the disjunctive (non-coreferential) reading. Therefore, the only way to get a coreferential reading is to use the reflexive pronoun. In 5.65, we do not know who is responsible for the creation of the picture. The nonreflexive is allowed because its use has no implication; that is, Mary could be the subject of a picture created by somebody else, the sentence doesn't carry enough information to implicate Mary as the creator. On the other hand, in 5.66 (my own example), when given the information that Mary is in fact the creator of the picture, the nonreflexive, on first reading seems to refer to another woman as the subject of the picture since the nonreflexive is linked to targets. When one learns that Mary is the source of the picture, a coreferential interpretation of the nonreflexive seems strange since its use, in comparison to the reflexive, does not reinforce Mary as the source of the creation event, the energy transfer from creator to creation.

We can now move on to explain why the reflexive is associated with the referent's viewpoint by demonstrating a final association between the source role and the viewpoint anchor. Sources of actions are also likely anchors for viewpoint interpretation. Among emotive verbs and verbs of mental actions, such as *think*, *love*, *like*, *admire*, etc., the source of the mental energy transfer is interpreted as the viewpoint anchor (if there is no intervening viewpoint to override the interpretation), since these verbs express metaphorical viewpoint. If the reflexive is associated with the source role in typical transitive clauses, then it can also be associated with the participant who serves as the viewpoint anchor. In simple reflexive clauses these roles often overlap, and the reflexive is associated with both semantic aspects (source and viewpoint anchor) of

the A participant. In this way, the core uses of the reflexive (in simple transitive clauses) are semantically linked to syntactically peripheral uses of the reflexive, including the examples of the reflexive in the PP. The core use of the reflexive in simple reflexive events can thus be seen as the impetus for the semantic associations carried over into the cases of the reflexives in less frequent constructions such as the PP data, picture nouns, and many constructions that utilize long-distance reflexives. This hypothesis is supported in the following argument based on uses of the long-distance reflexive.

One unique case, apart from simpler clauses (like *John thinks about himself*) where the source role and the viewpoint anchor align is introduced in Grimshaw (1990) under a different premise having to do with argument structure. The examples she offers (some borrowed from Pesetsky 1987) are a useful source of evidence for the idea that the reflexive pronoun is associated with these two grammatical primitives. In both 5.67a, an example of "cataphora" in which the pronoun precedes the antecedent, and 5.67b, an example of "long-distance" anaphora, the reflexive pronoun is found in an uncommon position, grossly in violation of the Condition A of Binding Theory:

(5.67) a. ?*Stories about herself generally please Mary.* (Grimshaw 1990:165 #31b)

b. ?*It worried them that it seemed necessary for pictures of each other (of themselves) to be on sale.* (Grimshaw 1990:165 #33a)

In both examples, the reflexive is judged to be grammatical (even if marginally so). In each case, the reflexive, always in an object position of course, is contained within the noun phrase that functions as the source of the energy transfer: the source of the pleasing in 5.67a and the source of the worrying in 5.67b. Likewise, the reflexive is coreferent with the viewpoint anchor from which the sentence is understood. Both

verbs, *please* and *worry*, express the psychological or mental state of the referent. 5.67a and 5.67b contrast with 5.68. In 5.68 the reflexive is still contained within the NP that serves as the source of the energy transfer, but the reflexive no longer aligns with the point of view anchor since the verb does not imply that the proposition should be understood from Mary's point of view:

(5.68) *?*Stories about herself generally describe Mary accurately.* (Grimshaw 1990:165, #33b)

Example 5.68 shows how important the semantics of point of view are in licensing the reflexive in non-prototypical constructions, and the examples as a unit provide further evidence for a semantic correspondence between sources, point of view anchors, and the reflexive pronoun.

A final piece of evidence for this tripartite association comes from data and description given in Kuno (1987). Kuno (p.164) argues for an "awareness condition" on picture noun reflexives: "Use a picture noun reflexive if, at the point in time that the sentence refers to, the referent of the reflexive perceived, perceives, will perceive the referent of the picture noun as one that involves him. Use a picture noun non-reflexive pronoun otherwise." He gives the following minimal pair to illustrate the condition:

5.69 a. *John$_i$ knows that there is a picture of himself$_i$ in the morning paper.*
(Kuno 1987:164 #3.14a)

b. **John$_i$ still doesn't know that there is a picture of himself$_i$ in the morning paper.*
(Kuno 1987:164 #3.15a)

In 5.69a John's awareness of the picture licenses the reflexive, and in 5.69b, John's ignorance of the picture rules out the use of the reflexive. As Kuno later suggests (p.167), the semantic property of awareness is inextricably linked to ideas about

discourse viewpoint. That is, the impossibility of the reflexive in 6.60b can be explained by the fact that John cannot serve as a viewpoint anchor in 6.60b since he has no knowledge of the picture's existence. The listener would assume that the proposition should be interpreted from the speaker's point of view. In 5.69b, negation thus functions to remove John from the role of viewpoint anchor.

In another set of examples (based on sentences produced by Kuno for his argument about the awareness condition) negation functions to remove the referent, not necessarily from the position of viewpoint anchor, but to remove the referent from the source position in the energy transfer dynamic. In 5.70, John serves as the source of the energy transfer to the painting:

(5.70) *John$_i$ painted that portrait of himself$_i$ on horseback.*

If the sentence is negated, however, John no longer serves as the source of the energy transfer to the painting, someone else does:

(5.71) **John$_i$ didn't paint that portrait of himself$_i$ on horseback …(Mary did).*
(Adapted from Kuno 1987: 163 #3.11a)

When John is not the source of the energy transfer, the reflexive is not licensed in the picture NP because reflexives share a semantic link with sources. In these cases negation has the effect of removing John from the source role thus ruling out the reflexive.

The above cases should not be confused with the reflexive's basic function, which is to signal event reflexivity. When the use of the reflexive pronoun is to signal the

event as semantically reflexive, negation, or a lack thereof, doesn't play a factor in licensing the reflexive. The reflexive is mandatory in 5.72a and 5.72b:

(5.72) a. *John$_i$ kicked himself$_i$.*

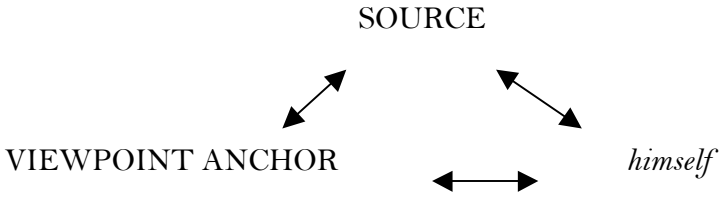
b. *John$_i$ didn't kick himself$_i$.*

The fact that the reflexive is obligatory in both 5.72a and 5.72b indicates that negation in simple clauses such as these (with no added context) does not necessarily function to remove John from the source role but rather to negate the occurrence of the reflexive event; that is John still serves as the source of a potential energy transfer that did not occur. It remains unresolved as to why negation works differently with reflexives in simple clauses like 5.72 and reflexives in peripheral syntactic domains like 5.71.

5.10 Summary

The above analysis serves as an explanation for a tripartite association between sources (in the energistic understanding) of events, viewpoint anchors, and the reflexive pronoun as illustrated in Figure 5.73:

(5.73)



Apart from its use to extrinsically mark an event as reflexive, the reflexive is often used in the BNC data when its antecedent is the source of the energy dynamic and/or the viewpoint anchor from which the event construal should be filtered. We now have a plausible explanation as to why the reflexive shows up in the data when it "is not needed"; that is, when the event is intrinsically reflexive as in 5.74 or when the event is not reflexive in nature as is 5.75.

(5.74) *John sat at the bar and began to think about the meeting. He pulled the ashtray toward himself.*

(5.75) *John constantly wants attention around himself.*

In both cases, John is the source of the energy and the anchor of the metaphorical perspective from which the sentence should be understood.

5.11 Conclusion

As a whole, this chapter has served as supplemental explanation of the pronoun patterns found in the English corpus. Evidence in this chapter affirms the notion that the grammar of space and spatial relations in the event structure alone do not predict the distributional tendencies of anaphora in the PP. Just as in other syntactically peripheral domains such as complex NPs, the use of the reflexive in the PP signals different clues to the overall construal of the event. Not only do speakers use the reflexive pronoun to signal the metaphorical viewpoint from which the event should be understood, but the reflexive pronoun itself is associated with the semantic role of 'source' in the energy transfer between event participants. Importantly, this chapter has provided evidence that the specificational qualities of the reflexive pronoun include a complex set of

discourse and event markers. These associations play a separate and supplemental role in explaining the overall pronominal distribution found in the PP.

Chapter 6

Repercussions for syntactic theory

6.1 Introduction

Chapters 4 and 5 combine to provide a new picture of anaphoric distribution. The analysis put forth crucially relies on event semantics to explain pronoun use in the prepositional phrase as part of a syntactically complex construction, and the analysis extends to explain the use of the reflexive to mark reflexive semantics, to reference energy sources, and to signal viewpoint construal. The above-mentioned spatial characterization of reflexive events assumes that a reflexive event is defined by referential identity between the source and target of the action. The person or entity that initiates the event is the same person or entity that is affected by the event. In English, there are several ways to mark this symmetry. Rarely it is marked on the verb as in 6.1:

(6.1) *Some narcissists self-loathe unconsciously.*

Sometimes it is present in a description:

(6.2) *John is a self-hater.*

And, as discussed in Chapter 4, intrinsically reflexive events show this symmetry with no required marking at all:

(6.3) *John shaved this morning.*

(6.4) *John pulled the book toward him.*

Finally, of course, and most commonly, the shared identity between source and target is marked with the reflexive pronoun as either the object of the verb or of the preposition in an extrinsically reflexive event:

(6.5) *John stabbed himself.*

(6.6) *John stabbed the knife into himself.*

These example show, crucially, that a reflexive event may be marked with the reflexive pronoun, but that marking is not a requirement for the event to be construed as reflexive. Needless to say, the reflexive semantics in the examples given above affect the event's construal. In this chapter it will be shown that our characterization of semantic reflexivity interacts with the syntactic model of anaphora. The data explored in Chapters 4 and 5 lead to interesting repercussions for Binding Theory as well as provide new insights into syntactically puzzling data involving pronouns. This chapter outlines several consequences and applications of the aforementioned semantic tendencies.

6.2 As for syntax …

With our new semantic characterization of anaphoric patterns, it is appropriate to examine how the proposed explanation fits in with the syntactic models reviewed in Chapter 2. In Chapter 2, Büring (2005) and Hestvik's (1991) similar syntactic models for anaphoric distribution were outlined. Although they differ on their specific implementation, both researchers maintain that the general principles of Binding Theory can extend to the PP domain. In contrast, Safir (2004) and Reinhart and

Reuland (1993) appeal to discourse factors similar to those outlined in Chapter 5 of this thesis as partially responsible for anaphoric distribution in the PP, but they, nevertheless, adhere to binding principles when discussing core anaphoric distribution in syntactically simple clauses. Accordingly, Safir (2004; 3) concludes: "Binding Theory remains at the heart of most current approaches [to anaphora] in generative grammar".

One of the syntactic arguments for the need for Binding Theory comes from examples of ellipsis. In syntactically simple clauses, binding has been used to explain why the reflexive semantics of Clause 1 carry over to the *vp*-elided reconstruction of Clause 2 as in 6.7:

(6.7) *John kicked himself and then Larry did too.*
(Sloppy Reading: Larry kicked Larry.)
(Strict Reading: Larry kicked John.)

In the sloppy reading Larry kicked Larry not John. The syntactic explanation of the sloppy interpretation is that both the structure and identity relationships of the subject and object in Clause 1 are copied to Clause 2 (along of course with the specified verbal semantics). Then, when *Larry* fills the Spec IP position of Clause 2, his identity automatically fills the object position as well, thus providing the semantically reflexive interpretation that Larry kicks himself. In the "strict" reading of 6.7, Larry kicks John. The syntactic explanation of the strict interpretation is that only the structural relationship of the subject and object in Clause 1 is copied to Clause 2. The identity relationship between subject and object positions is not copied; therefore the assignment of John as Patient in Clause 1 is carried over to the interpretation of Clause 2. In Clause 2, Larry fills the Agent role, but John remains as the Patient; the reflexive

interpretation of Clause 1 is not carried over to Clause 2. Safir (following Fiengo and May 1994) calls this type of reading an example of 'vehicle change' (2004:133).

Our patterns found in the PP data display interesting properties when tested with syntactic copying. Context free VP ellipsis tests of the PP data indicate that for many examples, one interpretation (strict or sloppy) is more "available", more natural, than the other. The data in this section suggest that binding between antecedent and pronoun is affected by reflexive semantics, which are in turn understood (as we have established) by spatial schematics. In order to demonstrate this intriguing interaction, the PP data has been divided below by pronoun type.

We can first examine binding relationships through Principle A: the reflexive must be bound within the relevant minimal domain. The analysis of Chapter 4 pointed to several reasons why the reflexive would appear in the PP. One reason is to signal a syntactically complex event is in fact reflexive in nature: a canonically outward directed action is directed toward the referent's body, as in 6.8:

(6.8) *John$_i$ pushed the box toward himself$_i$.*

A second reason the reflexive might occur is to reflect the metaphorical understanding of the human body as a container as in 6.9:

(6.9) *John$_i$ found the truth within himself$_i$.*

A third reason the reflexive might occur is to signal that the action is performed within the referent's peri-personal space as in 6.10:

(6.10) *John$_i$ put the blanket around himself$_i$.*

In all of these cases, *John* and *himself* are co-indexed and their relationship follows the definition of binding: *John* binds *himself* if John and *himself* are co-indexed and *John* c-commands *himself* within the relevant minimal domain, the IP. We can speculate that Büring and Hestvik would conclude that the relevant minimal domain is the IP, not the PP, in these cases since the PPs in 6.8-6.10 are complement-like in nature. In fact, the *VP ellipsis* test performed on 6.8-6.10, confirms the special relationship between the antecedent and the reflexive pronoun, in each case whatever spatial relationship is set up between antecedent and pronoun in the first clause holds for the second:

(6.11) *John$_i$ pushed the woman toward himself$_i$, and Larry did too.*

(Most available reading: Larry pushed the woman toward Larry.)

(6.12) *John$_i$ found the truth within himself$_i$, and Larry did too.*

(Most available reading: Larry found the truth within Larry.)

(6.13) *John$_i$ put the blanket around himself$_i$, and Larry did too.*

(Most available reading: Larry put the blanket around Larry.)

In 6.11-6.13, the sloppy reading is most available; it is assumed that Larry performs the same actions in reference to his own body, not in reference to John's body. Example 6.11 shows that the reflexive semantics established in Clause 1 of 6.11 have a strong tendency to carry over to Clause 2. Example 6.12 shows that the protagonist's body-internal actions in Clause 1 carry over to the reconstructed event in Clause 2. And, example 6.13 confirms that a self-directed action in Clause 1 is reconstructed as a self-directed action in Clause 2. It should not be surprising that these facets of semantic

reflexivity from Clause 1 are carried over to Clause 2 if we postulate that the syntactic binding relationship between antecedent and reflexive is also copied.

The more surprising results come from the canonical uses of the nonreflexive in the PP. Chapter 4 illuminated corresponding reasons why the nonreflexive would occur. In one case the nonreflexive occurs because the syntactically complex event is already reflexive in nature, the action is canonically performed toward the referent's body and the reflexive pronoun is not needed to signal reflexivity as in 6.14:

(6.14) *John$_i$ pulled the woman toward him$_i$.*

A second reason the nonreflexive might occur is in a case in which the action is literal and not metaphorically performed within the referent's body as in 6.15:

(6.15) *John$_i$ put the book next to him$_i$.*

And, a third reason the nonreflexive might occur is to signal that the action is not necessarily performed within the referent's peri-personal space as in 6.17:

(6.17) *John$_i$ constructed the fortress around him$_i$.*

Of interest in this set is example 6.14. It is example 6.14 that seems to conflict with the principles of binding when subject to an ellipsis test. Here, the relevant question for a theory of syntax is, does the nonreflexive pronoun gain its reference through binding or through independent coreference? The nonreflexive, according to Principle B, should be free in the minimal domain (IP); that is, the nonreflexive can be co-referent with the subject but its reference should be attained independently; it should not be bound by the

subject. Yet, the same VP ellipsis tests performed above in 6.11-6.13 to show that the reflexive is bound, clearly show here in 6.18 that the nonreflexive can also be bound within the same syntactic domain:

(6.18) *John$_i$ pulled a woman toward him$_i$, and Larry did too.*

(Most available reading: Larry pulled a woman toward Larry.)

In 6.18, the first and most available interpretation is the sloppy reading in which Larry is performing the same exact action as John; the action is performed in reference to Larry's body (not John's), and in this interpretation, the same spatial relationships between the referent and the pronoun set up in the first clause hold in the second. In Clause 1 of this sloppy reading, the subject position c-commands the PP object position and they are co-indexed, therefore the pronoun is bound. From a syntactic perspective, this structure-identity relationship is copied to Clause 2 yielding the reflexive interpretation of Clause 2. Example 6.18 suggests that canonical uses of the coreferential nonreflexive in a reflexive event may trigger a "bound" relationship between antecedent and pronoun more easily than its use in a nonreflexive event. So, for example, in 6.19 below, in which the coreferential nonreflexive is used in a spatial context that is not reflexive, the "unbound" (strict) interpretation is equally available with the sloppy interpretation. In the strict interpretation *John* and the pronoun are co-referent, but in the "reconstructed" event in the elided *vp*, John is assumed to still be the spatial landmark for the event, not Larry:

(6.19) *John$_i$ pushed a box away from him$_i$, and Larry did too.*

(Possible interpretation: Larry pushed a box away from John.)

(Possible interpretation: Larry pushed a box away from Larry.)

To show the contrast in which we are interested, this event can be switched to a reflexive event by changing the verb and preposition to convey an inwardly directed action. When this change is made, the unbound (strict) interpretation becomes less felicitous as in 6.20:

(6.20) *John$_i$ pulled a box toward him$_i$, and Larry did too.*

(Unlikely interpretation: ??Larry pulled a box toward John.)

(Likely interpretation: Larry pulled a box toward Larry.)

What these examples show us is that the interpretation of an elided VP is sensitive to the event's semantics and whether or not the event is reflexive in nature. The reflexive semantics of a reflexive event, even one that is not marked with the reflexive pronoun, will likely be reconstructed under VP ellipsis as a reflexive event. This example set, as a whole, support the idea that binding, as a structural relationship between pronoun and referent, may not be necessary to explain why reflexive semantics in first clauses carry over to "reconstructed" reflexive semantics in second clauses. We can make this conclusion because, within the appropriate semantic context, the "bound" interpretation is possible with either the reflexive or the nonreflexive pronoun. The data here can thus be added to a group of research, including Hardt (2003), that argues against a pure syntactic explanation of copying phenomena.

If a "bound" interpretation is sensitive to canonical spatial relationships and the conditions on binding are not the defining factor determining which pronoun will occur in a complex predicate, then an important question is raised as to Binding Theory's explanatory power for the complementary distribution of the two English pronoun

types in syntactically simple predicates. Does Binding Theory explain the core distribution? As Safir (2004: 15) puts it, are binding conditions simply "stipulations" on domains or are they actually "about" the reflexivity of predicates? But, any potential answer to Safir's question begs the larger more fundamental question, which is: how should the semantic understanding of reflexivity be modeled in the grammar of English? Is Binding Theory needed to model the distribution of reflexive and nonreflexive pronouns? How can the difference between a reflexive and a nonreflexive interpretation, namely the fact that there are two interpretations, be explained without referencing the syntax?

The answer to this question lies in one's beliefs about the fundamental primitives of grammar. If those primitives are based on spatial relations, energy dynamics, and abstracted schematic simulations, then the answer is relatively simple. We can now return to the core data in 6.7, repeated here in 6.21:

(6.21) *John kicked himself and then Larry did too.*
(Sloppy Reading: Larry kicked Larry.)
(Strict Reading: Larry kicked John.)

In Clause 1, John performs a reflexive action where he kicks his own body, and that action is simulated as such. Clause 2 is more open to interpretation because its lacks essential semantic information. In one case the actor, Larry, is simulated performing a reflexive action himself; thus Larry carries out the same kicking action that John does on Larry's own body. Larry, like John in Clause 1, is both the source and target of the action. In another case, Larry is simulated to kick John. John has been identified as the target of the kicking in Clause 1 and so he can serve as a possible target of kicking in

Clause 2 because not enough information is provided to rule out that simulation. Clause 2 allows for two possible energistic relationships and thus two possible simulations of the event.

6.3 Binding as a model of reflexivity

In Chapter 2 we touched on the fact that syntactic approaches to binding always reference the way in which the theory cleanly models the core, simple clause complementarity of reflexives and nonreflexives. However, within this literature there is no explanation given as to why. With our enriched model of semantic reflexivity we can now give a semantic explanation as to why Binding Theory predicts complementarity in syntactically simple clauses. In syntactically simple English clauses, event targets appear in verbal object position. In these simple clauses the reflexive pronoun must be used to signal the event is semantically reflexive, that the target of the action is also the source of the action. The reflexive then sits in verbal object position, the position that is c-commanded by the subject position. The reflexive references the source of the action which appears in subject position, so by default the reflexive pronoun is not only coreferential with the subject, but it's position is also c-commanded by the subject position.

We have seen from our data in Chapters 4 and 5 that the reflexive pronoun is just one of several signals of a reflexive event. It is argued here that binding should be decoupled from semantic reflexivity, that, to use Safir's words, the conditions on binding are not "about" reflexivity per se, but rather denote an ad hoc structural relationship that exists within the English data. The reason an event is construed as reflexive is not because the reflexive pronoun is dependent on its antecedent for reference, but because

the reflexive pronoun signals a source-target alignment in the construal of the event. As we have seen, this same source-target alignment can also occur even with the use of the nonreflexive pronoun when a syntactically complex predicate is canonically inwardly directed.

Apart from the attempt to explain reflexive semantics with Binding Theory, the theory has also been implemented to explain the locality/nonlocality effects of the two pronoun types in the domain of reference resolution strategies. Within this domain as well, schematic reflexivity interacts in an interesting way with the listener's implementation of syntax-based resolution strategies. This topic is the subject of the next section.

6.4 Locality

The schematic characterization of reflexivity and the corresponding spatial data call into question the usefulness of Binding Theory as a predictor of reference resolution. Traditionally, binding has been implicated as an explanation of the general notion that the reflexive pronoun (in it's non-logophoric use) will take the most local possible antecedent within a given domain. Conversely, the theory predicts that the nonreflexive pronoun will not take the most local possible referent to fill its semantic value. Two specific claims will be made in this section in reference to these generalizations on locality and Binding Theory. The overall claim is that binding itself is not the best predictor of locality effects. Specifically, the first sub-claim is that nonreflexive pronoun resolution is dependent on schematic, semantic reflexivity, which overrides "non" locality. The second sub-claim, also made by various researchers for other arguments (see Zribi-Hertz 1989), is that the logophoric (or 'exempt' (Safir 2004))

use of the reflexive pronoun needs to be separate from its use to signal that an event is reflexive. We will begin the discussion with the latter claim.

Returning to the data in Chapter 2, we can recall that Binding Theory was used by Chomsky (1995) to show how the reflexive pronoun depends on its governor within the most local domain. He used the NP domain to show these locality effects; the data is repeated here in 6.22 and 6.23:

(6.22) *John$_i$ likes Bill's$_j$ pictures of himself$_i$.*

(6.23) *John$_i$ likes Bill's$_j$ pictures of him$_{i/*j}$.*

In 6.22, the reflexive pronoun must refer to Bill because (within the NP domain) this DP position c-commands the (object of the preposition) position in which the pronoun sits. Because the nonreflexive pronoun must be free from Binding in 6.23, it can only refer to the Spec-VP position; it can only reference John. As a result the reflexive pronoun takes the more local antecedent and the nonreflexive references the less local possible referent. These locality effects disappear when one of the two possible referents is absent from the sentence as in 6.24 and 6.25:

(6.24) *John$_i$ likes the pictures of him$_i$.*

(6.25) *John$_i$ likes the pictures of himself$_i$.*

In either case, in 6.24 or 6.25, the pronoun can refer to John because there is no other more local possible binder. Putting aside the lengthy discussion in Chapter 5 about Source/Target role referencing and its effect on the use of the reflexive pronoun, a simple conclusion from this basic data raised by Chomsky is that when there are two

possible referents in a given sentence, the nonreflexive pronoun will reference the less local of the two. And, importantly, in these cases the reflexive is not used to signal a reflexive action (there are no reflexive semantics conveyed in 6.22), but is used to indicate that the pronoun's referent is the more local of two possible referents.

The corpus data illuminates an interesting property about semantic reflexivity and its effects on locality. The tendency for the nonreflexive to reference the "least" local referent is overridden by event simulation based on the understanding of canonical event schematics. In the following data, the sentence in 6.26 has the potential to describe an inwardly directed reflexive event. Whereas, the sentence in 6.27 describes an outwardly directed, nonreflexive event. In both cases the nonreflexive pronoun has two possible referents, yet the reference resolution outcome is not based on locality effects, but rather it is based on the semantics of the predicate:

(6.26) *John pulled Bill's chair toward him.*

(More likely: John pulled Bill's chair toward John.)

(6.27) *John pushed Bill's chair toward him.*

(More likely: John pushed Bill's chair toward Bill.)

In 6.26, the more likely interpretation is that the nonreflexive refers to John because the act of pulling is canonically self-directed, constituting a semantically reflexive event. On the other hand, in 6.27, the more likely interpretation is that the nonreflexive refers to Bill because the act of pushing is canonically outwardly directed, inhibiting the interpretation that the event is semantically reflexive. The examples in 6.26 and 6.27 are of course syntactically different from the NP domain examples above (6.22 and 6.23) since the referent and pronoun are not in the same minimal domain (not both in the NP

–Bill's chair)). Nevertheless, the pair as a unit share the same syntactic structure indicating that, when implemented in semantically reflexive data, reference resolution for the nonreflexive pronoun is dependent on the semantics of the predicate, not just dependent on locality effects.

6.5 Interim Summary

In this section it has been argued that the schematic understanding of semantic reflexivity has an obvious effect on event simulation. For the cases discussed in this section, reflexive semantics serves as a predictor of reference resolution, and a thorough understanding of event semantics trumps the predictions made by Binding Theory; that is, when the data becomes syntactically complex, Binding Theory performs inadequately in predicting pronoun use.

A final point to make on the repercussions of our hypotheses from Chapters 4 and 5 has to do with the nature of English pronouns and how pronouns themselves should be represented in grammar. This discussion is the subject of the next section.

6.6 Should the Syntax of Anaphora be the Semantics of Anaphora?

In his seminal work on the topic of anaphora, *The Syntax of Anaphora* (2004), Ken Safir argues for a competition theory of anaphoric resolution. Implicitly referencing popular theories of grammar, such as Optimality Theory (Prince and Smolensky 1993), that involve competition amongst competing forms in language, Safir suggests that the complementary distribution of nonreflexive and reflexive pronouns in simple English clauses is based on both the syntactic position of the pronoun and its ranking on a hierarchical scale of dependence. Safir posits that all types of referential expressions

exist on a scale of referential dependency, in which reflexive pronouns are more dependent than nonreflexive pronouns, which are, in turn, more dependent than referential NPs (proper names, descriptions, epithets, etc.). This scale exists in the grammar and interacts with the syntactic apparatus of the grammar. Safir uses the basic assumptions of c-command to capture his theory of reference resolution in the "Form to Interpretation Principle: If x c-commands y and z is not the most dependent form available in position y with respect to x, then y cannot be directly dependent on x (Safir 2004: 50)." The following sentences demonstrate the interpretation principle at work:

(6.28) a. *John hit himself.*

b. *John hit him.*

In 6.28a, the reflexive pronoun *himself*, z, is the most dependent form available among referential expressions. *John*, in position x, c-commands the object position, y, and thus, the reflexive pronoun gains its referential value directly from position x; *himself* refers to John. Contrastively, in 6.28b, the nonreflexive pronoun him, z, is not the most dependent for available among the possible referential expressions. Again, *John*, in position x, c-command the object position, y, but because the nonreflexive pronoun is not the most dependent form available (the reflexive is), the object position, y, in which the pronoun sits cannot be directly dependent on the subject, x, and, therefore, *him*, must not reference John. In this case, coreference between x and y is not expected, and x and y form an obviative pair.

Safir's analysis introduces two questions: how are pronouns represented in the lexicon, and how should reflexive events be represented in the grammar? First,

addressing the former question, it is clear that Safir, like most syntactic researchers working on binding, assumes that English pronouns are semantically vacuous, their only inherent semantic value includes gender and number specification. We have seen evidence in Chapter 4 for a much richer lexical specification for the reflexive pronoun, arguing that the reflexive shares a close semantic connection with the word *self* and concept self, and the use of the reflexive pronoun simultaneously activates the various metaphorical understandings associated with the self. This different representation of the reflexive pronoun moves it closer to the category of other referential NPs, such as definite descriptions and names.

In addressing the second question, we have concluded that, in simple clauses, reflexive pronouns are markers of reflexive actions. Their job, so to speak, is to encode a direct identity between the Source and Target of an event. This characterization constitutes what a reflexive action truly is. It is posited here that reflexive pronouns are simple markers of reflexive actions, and if what constitutes a reflexive action (and all the spatial schematics involved) is sufficiently understood, then the distribution of nonreflexive pronouns does not need to be understood in order to understand the distribution of reflexive pronouns.

We have seen plenty of examples in the PP data in which the nonreflexive pronoun is clearly assumed to be co-referent with the subject even though the reflexive is available and not used. A competition theory ignores the important roles of the semantic understanding of reflexive events. By basing a theory of pronoun use on referential dependency, the important spatial factors that help dictate pronoun choice are ignored. The referential hierarchy and the competition theory based on closed-class

paradigmatic contrasts Safir posits are epiphenomenal to the English pronominal system and not necessary to include in the grammar.

The first person pronoun serves as further evidence against a competition theory of anaphoric complementarity. The nonreflexive, *me*, can only refer to the identity of the speaker. Among first person pronouns, unlike third person pronouns, there is no referential ambiguity as to whom the nonreflexive refers. Within the first person paradigm, the pronoun can only reference the speaker's (or narrator's) identity. In this paradigm both the reflexive and nonreflexive pronoun's identities are dependent (in Safir's meaning) on the speaker. Safir's theory seems to predict that reflexive actions marked in the third person are processed differently from reflexive actions marked in the first person, yet there is no a priori reason to believe that this is the case.

In 6.28b above, *him* must reference an identity other than John's. That our theory predicts this fact is not because the reflexive is available and not used, but because the action is not reflexive and thus simulated as a nonreflexive event in which the Source of the action cannot be the Target of the action. Our theory hypothesizes that, just like third person reflexives, first person reflexives in simple clauses mark symmetry between the Source and Target of an event; however, unlike the third person nonreflexive, the first person nonreflexive has to refer to the identity of the speaker, but cannot be used to mark symmetry between the Source and Target of an action, thus if the nonreflexive is used as in 6.29, a special construal is assumed in which the pronoun actually refers to a separate entity with a shared identity (a 'proxy' reading):

(6.29) *I just kicked* ***me***.

(Appropriate if I (the human) kicked me (a statue of myself.))

We predict that the way in which an event is simulated is a phenomenon separate from the link between a pronoun and a participant's identity. As the first person examples show, the fact that two referential items refer to the same identity in a syntactically simple clause, does not necessarily equate to semantic reflexivity. As discussed in our schematic model of reflexive events in Chapter 4, when the subject is singular, a semantically reflexive event involves only one participant, who fills both Source and Target event roles. In example 6.29, there are two event participants, (the kicker (the speaker) and the kicked (the statue)) but only one identity, which is shared by both participants. Again, shared identity does not equate to semantic reflexivity.

6.7 Summary

The theory of semantic reflexivity put forth in this thesis suggests that reference resolution is dependant on a schematic understanding of the semantics of the described event. One job of the reflexive pronoun is to signal that an event is semantically reflexive. The interpretation of the denoted event as reflexive is not achieved because the listener, when processing (or simulating) the event, is (subconsciously) aware of the fact that the nonreflexive could be used instead, as Safir would argue. Rather, the reflexive interpretation is achieved because the event is simulated as semantically reflexive; shared identity between pronoun and referent is a secondary association registered by the listener. Along these same lines, we posit here that even a subconscious registry of shared identity between pronoun and antecedent is not necessary to construe an event as semantically reflexive since intrinsically reflexive events need no pronoun at all to be simulated as self-directed actions.

Chapter 7

Conclusion

The goal of this thesis has been to summarize well-known existing, both syntactic and semantic, models of anaphora in and out of the PP; provide a clear picture of anaphoric distributional trends within naturally occurring data; explain those trends by appealing to spatial relationships among event participants and objects; and finally, connect the distribution of anaphora in the PP to the occurrence of anaphora in syntactically simple clause structure. Beyond trying to meet these objectives, we have concluded several important take-away messages about anaphoric distribution in the PP.

First, the distribution of reflexive and nonreflexive pronouns in the PP is not haphazard. Distributional trends can be explained by motivating forces within the grammars of English and Spanish. These motivating forces are based on the perceived directionality and location of the event with respect to the protagonist's body. In both languages, events that occur inside the referent are more likely to pattern with the reflexive pronoun than are events that occur outside the referent's body. Specific instances of described events that are self-directed, that is directed toward the referent, especially ones that are canonically directed away from the referent (such as the act of pushing away), obligatorily pattern with the reflexive pronoun. And, finally, events that are performed within the peri-personal space of the referent are more likely to pattern with the reflexive pronoun than are events performed outside of the referent's peri-personal space.

We attributed these distributional trends to a range of factors that play a role in the semantic specifications and associations of the pronouns themselves. First, it was argued that the high rate of reflexive pronouns in events that are metaphorically located in the body is justified by the reflexive pronoun's close semantic association with the concept of self, a metaphorical body-internal entity. Second, we claimed that the reflexive pronoun is triggered by events which are performed on the body (in the referent's peri-personal space), and also triggered by canonically other-directed events which are in fact self-directed. Cases of these types were explained by demonstrating a schematic, semantic parallelism between syntactically complex reflexive events and syntactically simple reflexive events. We contended in both cases, the reflexive pronoun signals a contrastive element. In both syntactically complex cases, the PP examples (e.g. *John pushed the box toward himself*), and syntactically simple cases, those with basic clause structure (e.g. *John kicked himself*), the reflexive is used to signal that the direction of the event is counter to the direction of expectation. By taking this approach, we are able to provide a reason for why certain reflexive events (e.g. bathe, or pull something toward you) do not have to, and most often, do not occur with the reflexive pronoun.

Supplemental to our discussion of the influence of spatial parameters on the distribution of anaphora, we supported the claim that character (as opposed to narrator/speaker) viewpoint is associated with a higher occurrence of the reflexive pronoun. We partially attributed this assertion (that the reflexive is a perspective marker) to a link between space and metaphorical perspective. A binary set of participant roles: Source and Target were introduced, and we argued for an association between the reflexive pronoun and the Source role. We claimed that these schematized,

abstracted roles often conflate two dimensions present in most events: energy transfer and metaphorical perspective, and we posited that this associative conflation is an explanation as to why the reflexive is associated with metaphorical viewpoint.

By providing more data for a semantic characterization of reflexivity, we finish our account by rejecting the explanatory power of Binding Theory, concurring that its role as a model of anaphoric distribution is separate from its ability to explain the principles of which it is made. That is, using the words of Safir (2004: 15), (contrary to the conclusions he gives), our observations indicate that Binding Theory *is* merely a "stipulation" on domains, but not actually "about" the phenomenon of reflexivity.

Of course answering (or trying to answer) some questions raises others. This whole thesis has been about *trends* in the use of the reflexive versus the nonreflexive pronoun in the PP. All of our data point to levels of frequency in corpus data, which we translate to likelihood of usage by an individual speaker. One question this methodology raises and does not answer is the following: if we actually understood every single piece of what constitutes a speaker's linguistic apparatus and what dictates the choice of one pronoun or the other (e.g. detailed spatial schematics, intended meaning, perspective marking, syntactic priming effects, memory, etc.), would we be able to predict with certainty which pronoun would be chosen in a given discourse environment? Or, is our data representative of a linguistic phenomenon which is truly unpredictable, meaning is the distribution of pronouns in the PP an area of grammar that is best characterized by theories which incorporate probabilities of occurrence, but not absolute prediction?

Along the same lines, the research provided here could be used as a beginning of a cross-linguistic study on the topic. We have speculated that our hypothesis, which

puts spatial primitives at the explanatory center of anaphoric distribution, is a linguistic universal. Of course there are at most three languages explored in this dissertation to support that claim. Based on anecdotal evidence, we should believe that similar finding are to be discovered in other Germanic and Romance languages and Japanese. Certainly, our claims can be tested in any language which shows a split between reflexive and nonreflexive pronominal forms.

Lastly, we haven't definitively addressed the enterprise of space in syntax. In Chapter 4 we touched on the theory of some syntacticians (Svenonius 2004 a, b; Tungseth 2003) who claim that spatial relations can be modeled in syntax. We concluded that these types of models are too limited by discrete categorization to accurately model the data. A question we did not discuss was the validity of the premise that spatial relations between sentence participants can be part of the syntax at all. We leave questions such as these to future "anaphorists".

References

Aitchison, J. 1994. *Words in the mind: An introduction to the mental lexicon.* Cambridge MA: Blackwell Publishers, Inc.

Baker, C. L. 1995. Contrast, discourse prominence, and intensification, with special reference to locally free reflexives in British English. *Language* 71: 63-101.

Banfield, A. 1979. Où l'épistémologie, le style et la grammaire rencontrent l'histoire littéraire: le développement de la parole et de la pensée représentées. *Langue français* 44. 9-26.

Banfiled, A. 1982. *Unspeakable Sentences: Narration and representation in the language of fiction.* Boston: Routledge & Kegan Paul.

Belletti, A. and Rizzi, L. 1988. Psych verbs and theta theory. *Natural Language and Linguistic Theory* 6: 291-352.

Bennett, D. C. 1975. *Spatial and Temporal Uses of English Prepositions: An Essay in Stratificational Semantics.* London: Longman.

Bergen, B. 2005. Mental simulation in literal and figurative language. In *The Literal/Non-Literal Distinction*, S. Coulson and B. Lewandowska-Tomaszczyk (eds.), 255-278. Franfurt: Peter Lang.

Borsley, R. 1999. *Syntactic Theory, a Unified Approach.* London: Arnold.

Büring, D. 2005. *The Syntax and Semantics of Binding Theory.* Cambridge: Cambridge University Press.

Bresnan, J., Cueni, A., Nikitina, T., and Baayen, R.H. 2007. Predicting the dative alternation. In *Cognitive Foundations of Interpretation*, G. Boume, I. Kraemer, and J. Zwarts (eds.), 69-94. Amsterdam: Royal Netherlands Academy of Science.

Bresnan, J. and T. Nikitina. 2007. The gradience of the dative alternation. In *Reality Exploration and Discovery: Pattern Interaction in Language and Life*, L. Uyechi and L. Hee Wee. Stanford: CSLI Publications.

The British National Corpus, version 3 (BNC XML Edition). 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: <http://www.natcorp.ox.ac.uk/>

Cantrall, W. 1974. *Viewpoint, Reflexives, and the Nature of Noun Phrases.* The Hague: Mouton.

Chomsky, N. 1981. *Lectures on Government and Binding.* Dordrecht: Foris.

Chomsky, N. 1995. *The Minimalist Program.* Cambridge, Mass.: MIT Press.

Cohen, D. 1999. Toward a unified account of intensive reflexives. *Journal of Pragmatics* 31: 1041-1052.

Dabrowska, E. 1997. *Cognitive Semantics and the Polish Dative*. New York: Mouton de Gruyter.

Davies, M. 2002. Corpus del Español. < http://www.corpusdelespanol.org>

Dodge, E. and Lakoff, G. 2005. Image schemas: From linguistic analysis to neural grounding. In *From Perception to Meaning: Image Schemas in Cognitive Linguistics*, B. Hampe and J. E. Grady (eds.), 57-92. New York: Walter de Gruyter.

Dubinsky, S. and Hamilton, R. 1998. Epithets as antilogophoric pronouns. *Linguistic Inquiry* 29: 685-694.

Feldman, J. 2006. *From Molecule to Metaphor: A Neural Theory of Language*. Cambridge, Mass.: MIT Press.

Fiengo, R. and May, R. 1994. *Indices and Identity*. Cambridge, Mass.: MIT Press.

Fillmore, C. J. 1968. The case for case. In *Universals in Linguistic Theory*, E. Bach and R. Harms (eds.), 1-88. New York: Holt, Rinehart, and Winston.

Fillmore, C. J. 1992. "Corpus linguistics" vs. "computer-aided armchair linguistics", In *Directions in Corpus Linguistics: Proceedings from a 1991 Nobel Symposium on Corpus Linguistics*, J. Svartvik (ed.), 35-66. Stockholm: Mouton de Gruyter.

Freud, A. 1986. *The Essentials of Psycho-Analysis: By Sigmund Freud: Selected and Introduced by Anna Freud*. London: Hogarth Press and the Institute of Psycho-Analysis.

Gahl, S. and Garnsey, S.M. 2004. Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language* 80(4): 748-775.

Gahl, S. and Garnsey, S. M. 2006. Knowledge of grammar includes knowledge of syntactic probabilities. *Language* 82(2), 405-410.

Gallese, V. and Lakoff, G. 2005. The Brain's concepts: the role of the Sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology* 22(3-4): 455-479.

Gensler, O. 1977. *Non-Syntactic Antecedents and Frame Semantics*. Proceedings of the 3rd Annual Meeting of the Berkeley Linguistics Society. Berkeley CA: Berkeley Linguistics Society.

Gordon, P. and Hendrick, R. 1997. Intuitive knowledge of linguistic coreference. *Cognition* 62(3): 326-370.

Grimshaw, J. 1990. *Argument Structure.* [Linguistic Inquiry Monograph Eighteen]. Cambridge Mass.: MIT.

Hardt, Daniel. 2003. Sloppy identity, binding, and centering. In *Proceedings of SALT 13.* Seattle, Washington.

Hellan, L. 1988. *Anaphora in Norwegian and the Theory of Grammar.* Dortrecht: Foris Publications.

Herskovits, A. 1986. *Language and Spatial Cognition: An Interdisciplinary Study of the Prepositions in English.* Cambridge: Cambridge University Press.

Hestvik, A. 1991. Subjectless Binding Domains. *Natural Language and Linguistic Theory* 9(3): 455-496.

Heycock, C. 1995. Assymetries in Reconstruction. *Linguistic Inquiry* 26: 547-570.

Hirose, Y. 2006. The conceptual basis for reflexive constructions in Japanese." The Fourth International Conference on Construction Grammar. September 1-3, University of Tokyo, Japan.

Hornstein, N. 1996. Control in GB and Minimalism. *Glot International* 2(8): 3-6.

Hornstein, N. 1999. Movement and Control. *Linguistic Inquiry* 30(1): 69-96.

Huang, C.T.J. 1983. A Note on Binding Theory. *Linguistic Inquiry* 13(3): 554-561.

Huettig, F. & Altmann, G.T.M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition* 96(1): 23-32.

Johnson, M. 1987. *The Body in the Mind. The Bodily Basis of Meaning, Imagination and Reason.* Chicago: The University of Chicago Press

Jolly, J. 1991. *Prepositional Analysis within the Framework of Role and Reference Grammar.* New York, Peter Lang.

Kaiser, E., Runner, J.T., Sussman, R. and Tanenhous, M. 2004. Susceptibility to discourse/semantic factors: An experimental approach to short-distance pronouns and logophoric reflexives. *Workshop on Semantic Approaches to Binding Theory ESSLLI,* Nancy.

Kang, B. 2001. The grammar and use of Korean reflexives. *International Journal of Corpus Linguistics* 6(1): 134-150.

Kayne, R. 2002. Pronouns and their antecedents. In *Derivation and Explanation in the Minimalist Program,* S. D. Epstein and T. D. Seely (eds.),133-158. Malden MA: Blackwell.

Kemmer, S. 1993. *The Middle Voice.* Philadelphia: John Benjamins Publishing Company.

Kiparsky, P. 2002. Disjoint Reference and the Typology of Pronouns. In *More than Words* I. Kaufmann and B. Stiebels (eds.), *Studia Grammatica* 53: 179-226. Berlin: Akademie Verlag.

Kjellmer, G. 2004. Yourself: a general purpose emphatic-reflexive. *Advances in corpus Linguistics: Papers from the 23rd International Conference on English Language Research on Computerized Corpora (ICAME 23) Göteborg 22-26 May 2002.* Aijmer, K. and B. Altenberg (eds.), 267-278. Amsterdam/New York: Rodopi.

Konig, E. and Siemund, P. 2000. The development of complex reflexives and intensifiers in English. *Diachronica* 18(1): 39-84.

Kuno, S. 1972a. Pronimonalization, reflexivization, and direct discourse. *Linguistic Inquiry* 3: 161-195.

Kuno, S. 1972b. Functional sentence perspective. *Linguistic Inquiry* 3: 269-320.

Kuno, S. 1987. *Functional Syntax: Anaphora, Discourse, and Empathy.* Chicago: University of Chicago Press.

Laitinen, M. 2002. HE and THEY in indefinite anaphora in written present-day English. *International Journal of Corpus Linguistics* 7(2): 137-164.

Lakoff, G. 1996. Sorry, I'm Not Myself Today: The Metaphor System for Conceptualizing the Self. In *Space Worlds and Grammar.* G. Fauconnier and E. Sweetser. (eds.), 91-123. Chicago: University of Chicago Press.

Lakoff, G. and Johnson, M. 1980. *Metaphors We Live By.* Chicago, University of Chicago Press.

Lakoff, G. and Johnson, M. 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought.* New York: Basic Books.

Langacker, R. W. 1999. *Grammar and Conceptualization.* New York: Mouton de Gruyter.

Langacker, R. W. 2002. *Concept, Image and Symbol: The Cognitive Basis of Grammar.* New York: Mouton de Gruyter.

Larson, M. L. 1984. *Meaning-based translation: A guide to cross-language equivalence.* Lanham, MD: University Press of America.

Larson, R. K. 1988. On the double object construction. *Linguistic Inquiry* 19(3): 335-391.

Levin, B. 1999. Objecthood: An even structure perspective. *CLS* 35: volume 1: *The Main Session.* Chicago: University of Chicago.

Levinson, S. and Meira, S. 2003. 'Natural Concepts' in the spatial topological domain – Adpositional meanings in crosslinguistics perspective: An exercise in semantic typology. *Language* 79: 485-516.

Lidz, J. 2001. Condition R. *Linguistic Inquiry* (32)1: 123-140.

Lindstromberg, S. 1998. *English prepositions explained.* Philadelpha: John Benjamins.

Liu, H. 2003. Unpacking meaning from words: A context-centered approach to computational lexicon design. In *Modeling and Using Context, 4th International and Interdisciplinary Conference, CONTEXT 2003, Stanford, CA, USA, June 23-25, 2003, Proceedings*, P. Blackburn, C. Ghidini, R.M. Turner, and F. Giunchiglia (eds.), 218-232. Springer: Berlin.

Lødrup, H. 1999. Inalienables in Norwegian and Binding Theory. *Linguistics* 37(3): 365-388.

Lødrup, H. 2006. Do simple and complex reflexives have different binding domains? Annual Meeting of the Linguistics Society of America. January 6, 2006.

MacDonald, J. E. 2004. Spanish Reflexive Pronouns: a Null Preposition Hypothesis. In *Proceedings of the 23rd West Coast Conference on Formal Linguistics.* V. Chand, A. Kelleher, A. Rodríguez, and B. Schmeiser (eds.), 528-540. Somerville, MA: Cascadilla Press.

MacDonald, J. E. 2005. Motion verbs and reflexive pronouns in Spanish. Annual Meeting of the Linguistic Society of America, January 4th, 2005, Oakland CA.

Maldonado, R. 1999. *A Media Voz: Problemas Conceptuales del Clítico* se. Mexico: Universidad Autónoma de México.

Malkiel, Y. 1990. *Diachronic Problems in Phonosymbolism, Volume I.* Philadelphia: John Benjamins Publishing Company.

Malkiel, Y. 1992. *Diachronic Studies in Lexicology, Affixation, Phonology, Volume II.* Philadelphia: John Benjamins Publishing Company.

McEnery, T. and Wilson, A. 1997. *Corpus Linguistics.* Edinburgh: Edinburgh University Press.

Meyer, C. 2002. *English Corpus Linguistics: An Introduction.* Cambridge: Cambridge University Press.

Mylne, T. 1999. A feature-based system for classifying semantic roles. *Proceedings of the 1999 Conference of the Australian Linguistic Society.* <http://www.linguistics.uwa.edu.au/about/research/als99>

Nichols, L. 2006. The referential properties of pronouns and anaphors inside noun phrases. Manuscript, University of California, Berkeley.

Nida, E. 1975. *Componential Analysis of Meaning: An Introduction to Semantic Structures*. The Hague: Mouton.

Noguchi, T. 1997. Two types of pronouns and variable binding. *Language* 73(4): 770-798.

Oaks, M. P. 1998. *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University Press.

O'Dowd, E. M. 1998. *Prepositions and Particles in English: A Discourse Functional Account*. Oxford: Oxford University Press.

The Oxford English Dictionary. 2nd ed. 1989. OED Online. Oxford University Press. 2006. <http://dictionary.oed.com>.

Payne,T. E. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.

Pederson, E. W. 1991. Subtle Semantics: Universals in the Polysemy of Reflexive and Causative Constructions. *Linguistics*. Berkeley: University of California, Berkeley.

Pesetsky, D. 1989. Wh-in-situ: Movement and unselective binding. In *The Representation of (In)definiteness*, E. Reuland and G.B. Alice ter Meulen, (eds.), 98-129. Cambridge: MIT Press.

Pederson, J. 2003. Reflexive intensification in Spanish: A historical perspective. The XVIth International Conference on Historical Linguistics, Copenhagen.

Pollard, C. and Sag, I. 1992. Anaphor in English and the scope of Binding Theory. *Linguistic Inquiry* 23: 261-303.

Prince, A. and Smolensky, P. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University Center for Cognitive Science Technical Report 2.

Pycha, A., Milbrath, C. and Eyre, S. 2005. Anaphora in African-American English. Annual Meeting of the Linguistics Society of America. January 8, 2005, Oakland CA.

Reddy, M. 1979. The Conduit Metaphor. In *Metaphor and Thought*, A. Ortony, (ed.), 284-324. Cambridge: Cambridge University Press.

Regier, T. 1995. A model of the human capacity for categorizing spatial relations." *Cognitive Linguistics* 6(1): 63-88.

Reinhart, T. 1976. The Syntactic Domain of Anaphora. Doctoral Dissertation, MIT, Cambridge, Mass.

Reinhart, T. and Reuland, E. 1993. Reflexivity. *Linguistic Inquiry* 24(4): 657-720.

Rini, J. 2003. The origin of Spanish *entre tú y yo* 'between you and me': A typological parallel to English *between you and I*?" *Diachronica* (20)1: 139-165.

Runner, J.T. and Kaiser, E.. 2005. Binding in Picture Noun Phrases: Implications for Binding Theory. HPSG05 Conference, Department of Informatics, University of Lisbon August 23–24, 2005.

Runner, J.T., Sussman, R. and Tanenhaus, M. 2006. Processing reflexives and pronouns in picture noun phrases. *Cognitive Science* 30: 193-241.

Safir, K. 2004. *The Syntax of Anaphora.* New York: Oxford University Press.

Sanders, J. and Spooren, W. 1997. Perspective, subjectivity, and modality from a cognitive linguistic perspective. In *Discourse and Perspective in Cognitive Linguistics*, W.A. Liebert, G. Redeker and L. Waugh (eds.), 85-114. Philadelphia: John Benjamins Publishing Company.

Sells, P. 1987. Aspects of logophoricity. *Linguistic Inquiry* 18(3): 445-479.

Schalley, A. 2004. *Cognitive Modeling and Verbal Semantics: A Representational Framework Based on UML.* New York: Mouton de Gruyter.

Sinclair, J. M., Jones, S. and Daley, R. 2004. *English Collocation Studies: The OSTI Report.* London: Continuum.

Slobin, D. 1996. Two ways to travel: Verbs of motion in English and Spanish. In *Grammatical Constructions : Their Form and Meaning*, M. Shibatani and S. A. Thompson (eds), 195-221 .. New York : Oxford University Press.

Speas, M. & C. Tenny. 2003. Configurational properties of point of view roles. In *Asymmetry in Grammar*, A.M. di Sciullo (ed.), 315-344. Amsterdam: John Benjamins.

Stojanovic, D. 2003. 'I shave just like man': intrinsic and extrinsic reflexive constructions in child English. In *Proceedings of the 2002 Annual Canadian Linguistic Association Meeting*, S. Burelle and S. Somesfalean (eds.), 312-325. Montréal: Cahiers Linguistiques de l'UQAM.

Stern, N. 2004. The semantic unity of reflexive, emphatic, and other *–self* pronouns. *American Speech* (79)3: 270-280.

Strahan, T. 2001. The distribution of reflexive pronouns in Norwegian. *Australian Journal of Linguistics* 21(1): 159-169.

Svenonius, P. 2004a. Adpositions, particles, and the arguments they introduce. Manuscript, University of Tromsø, Norway.

Svenonius, P. 2004b. Spatial P in English. Manuscript, University of Tromsø, Norway.

Stubbs, M. 2001. *Words and Phrases: Corpus Studies of Lexical Semantics.* Oxford: Blackwell Publishers.

Talmy, L. 2000a. *Toward a Cognitive Semantics Volume I: Concept Structuring Systems.* Cambridge MA: MIT Press.

Talmy, L. 2000b. *Toward a Cognitive Semantics, Volume II: Typology and Process in Concept Structuring.* Cambridge MA: MIT Press.

Tenny, C. 1996. Short distance pronouns and locational deixis. Archives of On-Line LINGUIST Conference on Geometric And Thematic Structure In Binding: October 14 to November 4, 1996.

Tenny, C. 2003. Short distance pronouns, argument structure, and the grammar of sentience. Manuscript, Carnegie Mellon University.

Tungseth, M. 2003. Two structural positions for locative and directional PPs in Norwegian motion constructions. *Nordlyd* (31)2: 473-487.

Tyler, A. and Evans, V. 2003. *The semantics of English Prepositions.* Cambridge: Cambridge University Press.

van Hoek, K. 1995. Conceptual reference points: A cognitive grammar account of prominal anaphora constraints. *Language* 71(2): 310-340.

van Hoek, K. 1997. *Anaphora and Conceptual Structure.* Chicago, University of Chicago Press.

Van Valin, R.D. 2004. Semantic macroroles in Role and Reference Grammar. In *Semantische Rollen*, R. Kailuweit and M. Hummel (eds.), 62-82. Tübingen: Narr.

Wechsler, S. 1997. Prepositional phrases from the Twilight Zone. *Nordic Journal of Linguistics* 20 (2):127-154.

Zribi-Hertz 1989. Anaphor binding and narrative point of view: English reflexive pronouns in sentence and discourse. *Language* 65(4): 695-727.

Appendix 1

Overview of Corpora

In this appendix, information is provided on the composition of the two corpora consulted for this project. The information provided here comes from the corpora's respective websites (British National Corpus (BNC) http://www.natcorp.ox.ac.uk/index.html; Real Academia Española http://www.rae.es).

British National Corpus

The BNC is comprised of 100,106,008 words. The corpus covers 4,124 texts, of which 863 texts (5%) are transcribed from spoken conversations or monologues. Of the spoken texts, 762 (6,154,248 words) are transcriptions of spoken texts collected from a pre-defined set of domains or contexts: educational, business, institutional, leisure, and unclassified. A total of 153 texts (4,211,216 words) are transcriptions of spoken texts collected by recruits identified by demographic sampling of the UK population, and the speakers represent different ages, socials classes, regions of the UK, and sexes. Each text has been segmented into orthographic sentence units. Because of copyright laws, full published works are not included in the Corpus (Meyer 2002); rather small excerpts of published materials reside in the database. The Corpus was completed in 1994, and mostly includes texts from 1975-1994 and a few texts dating back to 1964. 75% of the written texts were chosen from *informative* writings: of which roughly equal quantities were chosen from the fields of applied sciences, arts, belief and thought, commerce and finance, leisure, natural and pure science, social science, world affairs. 25% of the written texts are *imaginative*, that is, literary and creative works. The type of published

material included (the *medium*) is comprised of 60% books, 25% periodicals (newspapers and magazines etc.), between 5 and 10% miscellaneous published material (brochures, advertising leaflets, etc), and between 5 and 10% unpublished written material such as personal letters and diaries, essays and memoranda, etc. I, however, did not specify my searches, and thus my search findings include written and spoken data from all mediums and topics represented in the corpus.

Real Academia Española

All the Spanish data used in this study comes from Real Academia Española's *Corpus de Referencia del Español Actual, (CREA)* (Reference Corpus of Modern Spanish). The online corpus (www.rae.es) is composed of 150,778,934 words divided into three categorical mediums: books, 53%, press, 45%, and miscellaneous (which includes transcription of spoken Spanish), 2%. The origin of the source data comes from Spain, 56%, Latin America, 44%, and "Other" places, <.01%, and includes materials from the following countries: Argentina, Bolivia, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, Spain, United States, Uruguay, and Venezuela. The corpus includes texts dated from 1977-2002. Searches can be specified for country of origin as well as informational topic such as politics, tourism, or anthropology; as in the BNC, I did not specify my searches, and thus my search findings include data from all mediums, topics, and countries represented in the corpus.

Appendix 2

Search strings and total hits

The following two tables report the search strings entered into the BNC and CREA respectively along with the total hits produced for each search string.

Pronoun distribution by preposition

	next to	beneath	behind	in front of	before	below	above	beyond	on top of	on	toward	around	out of	into	in	inside	within	Total
NR	191	137	1759	650	827	96	181	62	44	2148	881	922	220	217	994	135	87	9551
R	0	0	2	2	10	2	9	7	2	147	5	21	22	35	168	29	71	532
Total	191	137	1761	652	837	98	190	69	46	2295	886	943	242	252	1162	164	158	10083

Table X.1-Total hits of third person nonreflexive and reflexive according to head preposition -English

Pronoun distribution by preposition

	next to	beneath	behind	in front of	before	below	above	beyond	on top of	on	toward	around	out of	into	in	inside	within	Total
NR	144	28	546	350	277	50	78	77	31	1602	314	354	157	125	766	154	58	5111
R	0	0	0	0	6	0	4	2	0	72	5	5	13	9	85	8	16	225
Total	144	28	546	350	283	50	82	79	31	1674	319	359	170	134	851	162	74	5336

Table X.2-Total hits of first person nonreflexive and reflexive according to head preposition -English

Spanish pronoun distribution by preposition

	a través de	al lado de	alrededor de	debajo de	delante de	dentro de	detrás de	encima de	hacia	Total
Nonreflexive	208	36	61	38	240	375	333	131	1232	2451
Reflexive	5	0	15	8	20	315	26	10	262	661
Total	213	36	76	46	260	690	359	141	1494	3112

Table X.3-Total hits according to head preposition -Spanish