SPLIT ANTECEDENTS AND PRONOMINAL EXPRESSION IN SPANISH

by

WILLIAM SCOTT PENNINGTON III

(Under the Direction of Lewis C. Howe)

ABSTRACT

Because the syntactic and/or thematic salience of a particular discourse referent is relative to that of co-occurring discourse referents, relying on syntactic and/or thematic salience to resolve anaphora is beneficial only when there are competing referents for a given anaphoric reference. Data from a production experiment and a processing experiment suggest that the accessibility of plural discourse referents in Spanish, as measured by form of referring expression, does not necessarily depend on the way in which those referents are constructed within the discourse. Non-split antecedents were found to be no more accessible than split antecedents in the absence of competing plural referents for a given plural anaphoric expression.

INDEX WORDS: Split antecedents, Pronominal expression, Form of referring expression, Anaphora, Centering, Accessibility, Spanish

SPLIT ANTECEDENTS AND PRONOMINAL EXPRESSION IN SPANISH

by

WILLIAM SCOTT PENNINGTON III

B.A., Furman University, 2004

A Thesis Submitted to the Graduate Faculty of the University of Georgia in Partial Fulfillment of

the Requirements for the Degree

MASTER OF ARTS

ATHENS, GEORGIA

2008

SPLIT ANTECEDENTS AND PRONOMINAL EXPRESSION IN SPANISH

by

WILLIAM SCOTT PENNINGTON III

Major Professor: Lewis C. Howe

Committee: Diana L. Ranson
Hildebrando Ruiz

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2008

DEDICATION

I would like to dedicate this work to my wife and best friend, Janna, who for the past two years has graciously humored my interest in Spanish pronouns while devoting herself to much more important work.

ACKNOWLEDGEMENTS

I would like to sincerely thank my major professor, Chad Howe, for his patience, responsiveness, and constant support. His insight and enthusiasm have served as an inspiration for most of my work as a graduate student. I would also like to thank Diana Ranson, Hildebrando Ruiz and Gary Baker for their helpful and perceptive reviews of this work. I extend a special thanks to my adoptive family in Cochabamba, Bolivia, who opened their home to me while I was away from my own. Finally, I thank Javier Lluch and Anne Wesserling for their contributions to the development of the test sentences used in the experiments developed for this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1   Introduction

The problem to be addressed by an analysis of anaphora can be summarized by two

principal questions: The first, related to production, asks why a speaker chooses one form of

referring expression over another (e.g., a pronoun instead of a lexical NP). The second, related to

interpretation, asks how a particular form of referring expression is interpreted by the

interlocutor (Blackwell 1998: 606). In (1b), for example, what prompts the Spanish-speaker to

use a null pronoun (Ø) instead of an explicit pronoun or a full NP? Conversely, to whom (or

what) does the interlocutor assign the null reference, and why?

(1) (a)   Juan$_i$ fue a la tienda$_k$ para María$_j$.
     (b)   Ø$_{i/k/j}$ Compró unos refrescos para la fiesta.

In order to address these questions, it is important to note that interpretation and

production are closely related: the successful interpretation of anaphora is possible only if the

production of anaphora is predictable – that is, if it depends on a set of identifiable constraints. It

is not surprising, then, that the identification of these constraints has taken center stage in the

discussion of anaphora (cf. Beaver 2004).

For reasons I will explain, however, much of the literature which seeks to identify the

constraints on the production and/or resolution of anaphora focuses primarily on singular

anaphora. Additionally, most analyses of anaphora stop short of explaining how the relationship between referents and their antecedents affects discourse coherence. In this regard, Centering Theory (cf. Grosz et al. 1995) is particularly valuable because it applies a set of constraints on anaphoric production and/or interpretation in order to predict the relative coherence of a given segment of discourse.

Thus, the primary goal of the current study is to address the problems which arise when frameworks dealing with singular anaphora are applied to an analysis of plural anaphora. Specifically, based on the results of a production experiment and a processing experiment, I will discuss the extent to which the production and/or resolution of plural anaphora is influenced by the structural representation of plural antecedents in Spanish. Finally, I will offer a preliminary consideration of how split antecedents might fit within a theoretical model which predicts the relative coherence of a given segment of discourse.

1.2   Review of the Literature

There are three main points which need to be developed in this section. First, all of the literature which I will review acknowledges, whether explicitly or implicitly, that the production and/or interpretation of anaphora is governed ultimately by pragmatic constraints. Second, within this body of literature, relatively few frameworks (with Centering Theory being one notable exception) attempt to apply the constraints on anaphoric production and interpretation in order to make predictions about relative discourse coherence. Third, for reasons which I will explain, there has been very little discussion regarding the production and/or interpretation of plural anaphora (especially regarding plural anaphora which resolves to plural antecedents).

1.2.1   Cognitive Frameworks: Accessibility Theory and the Givenness Hierarchy

Accessibility Theory (Ariel 1991) and the Givenness Hierarchy (Gundel et al. 1993) relate the form of a referring expression to the cognitive status of its referent. Cognitive statuses, as proposed by Gundel, represent "information about location in memory and attention state" (1993: 274). Within both frameworks, the cognitive status of a referent is determined by the interaction of syntactic and pragmatic factors.

Ariel (1991) maintains that the accessibility of a referent is determined by three factors: 1) distance between the referring expression and its antecedent, 2) the prominence of the referent, and 3) competition. Distance, according to Ariel, is not necessarily measured by number of clauses and/or words, but by being "embedded within the same frame, paragraph, etc." (1991: 445). Prominent referents include those discourse entities which are "discourse and sentence topics, discourse participants, important people or things in the lives of the discourse participants" (1991: 445). This second criterion alludes to the importance of the syntactic role of the antecedent, since discourse and sentence topics typically occupy the subject position (cf. Bolinger 1979; Reinhart 1981; Blackwell 1998). By competition, Ariel means the number of "other antecedents that can potentially serve as antecedents for the said NP" (1991: 445). Thus, higher prominence, smaller distance, and less competition all contribute to the greater accessibility of a particular discourse entity. Finally, Ariel suggests that the accessibility of a given referent is reflected in the form of the referring expression which evokes it: the greater the semantic and phonetic content of the referring expression, the lower the accessibility of the referent, and vice versa.[1] For example, the use of a null pronoun in Spanish would indicate a

---

[1] This relationship is expressed explicitly in the following scale (taken from Ariel 1991: 449):

highly accessible referent, while the use of a lexical NP would indicate a relatively inaccessible referent.

Likewise, Gundel et al. (1993) propose six cognitive statuses which correspond with certain forms of referring expressions (1993: 275). Similar to Ariel's Accessibility scale, within the Givenness Hierarchy, the most salient referents (those in the 'in-focus' set) allow for referring expressions with little or no semantic and phonetic content, while the least salient referents (those which are only identifiable by type) require greater linguistic specification.[2]

Also similar to Accessibility Theory, within the Givenness framework, the cognitive status of a referent is determined by both syntactic structure and pragmatic factors. As stated by Gundel et al., "subjects and direct objects…are highly likely to bring a referent into focus, whereas this is not the case for elements in subordinate clauses and prepositional phrases. However, inclusion in the 'in-focus' set depends ultimately on pragmatic factors" (1993: 279-

LOW ACCESSIBILITY
        Full name + modifier
        Full name
        Long definite description
        Short definite description
        Last name
        First name
        Distal demonstrative (+ Modifier)
        Proximal demonstrative (+ Modifier)
        Stressed pronouns + Gesture
        Stressed pronouns
        Unstressed pronouns
        Zeros
HIGH ACCESSIBILITY

[2] The Givenness Hierarchy for Spanish (as appears in Gundel et al. 1993: 284):

IN FOCUS	ACTIVATED	FAMILIAR	UNIQUELY IDENTIFIABLE	REFERENTIAL	TYPE IDENTIFIABLE
Ø él	ÉL éste ése aquél este	ése N aquel N	el N		Ø N un N

280). The following examples serve to clarify this interaction (adapted from Gundel et al. 1993:

280):

(2)	(a)	My neighbor's **bull mastiff** bit a girl on a bike.
	(b)	**It**'s the same dog that bit Mary last summer.

(3)	(a)	Sears delivered new siding to my neighbors with the **bull mastiff**.
	(b)	#**It**'s the same dog that bit Mary last summer.

(4)	(a)	The government is looking for a manager for **a large wind energy project**.
	(b)	I'm going to see the man in charge of **it** next week.

In (2b), the use of the expression *it* – which, in English, designates a referent which is highly

salient or 'in-focus' – is justified in part because of the syntactic prominence of the antecedent,

which occupies the subject position in (2a). When the same antecedent is moved to a more

oblique syntactic position in (3a), subsequent anaphoric reference reflects a lesser degree of

salience. Thus, the corresponding use of the semantically and phonetically weak expression *it* in

(3b) is considered infelicitous. However, in (4a-b), although *a large wind energy project*

occupies an oblique syntactic position, its referent is brought into focus "because of its

importance in this context" (1993: 280). Therefore, the use of *it* in (4b) is justified.

1.2.2 Pragmatic Frameworks:

Huang (1991) and Blackwell (1998) adopt a pragmatic approach based on a neo-Gricean

framework in order to account for the interpretation of anaphoric expressions. Huang's analysis

deals primarily with intra-sentential anaphora and claims that "the interpretation of reflexives,

pronouns, zero anaphors and lexical NP's can largely be determined by the systematic interaction

of [Levinson's] M- and I-principles" (1991: 316-317). Briefly, Levinson's I-principle states that "minimally informative forms [such as null pronouns] favor or 'I-implicate' a co-referential interpretation in the absence of any information to the contrary" (Blackwell 1998: 614). Conversely, the M-principle states that more informative forms favor or M-implicate a non-co-referential interpretation.

This interaction, however, is additionally subject to a "world-knowledge based Disjoint Reference Presumption (DRP), which assumes that "the arguments of a predicate are intended to be disjoint, unless marked otherwise" (Huang 1991: 317). By specifying that this presumption is based on "world-knowledge," Huang means that world-knowledge, rather than linguistic surface structure, may serve to mark the arguments of a predicate as co-referential.

Blackwell, while acknowledging that "grammatical constraints define what is and is not co-referentially possible at the sentence level," also maintains that syntactic structure cannot account for "pragmatic rules of inference and the interaction of language with world knowledge" (1998: 611). In other words, where a language's grammar permits anaphora which may be resolved to more than one referent, that resolution is governed by the following: 1) semantic constraints (such as lexical entailment), 2) pragmatic constraints (such as the interaction of Levinson's M- and I-implicatures), 3) cognitive factors (such as antecedent salience), and 4) extra-textual knowledge constraints (such as world-knowledge, cultural knowledge, or any knowledge shared by both the speaker and the interlocutor).

Both Blackwell (1998) and Huang (1991) identify factors which aid in the interpretation of anaphora where a language's grammar permits more than one linguistic form. Blackwell goes a step further in expanding upon Huang's "world-knowledge-based DRP constraint," explaining that any shared knowledge between a speaker and interlocutor, such as "understanding of each

other's attitudes, individual interests, family relations, and mutual assumptions about what is

socially or culturally prescribed" may facilitate the resolution of otherwise ambiguous anaphora

(1998: 614).

1.2.3   What About Discourse Coherence?

In accordance with the relevant questions identified in the introduction, the goal of the

studies just discussed is to identify a set of constraints which explains the way anaphora is

produced and/or interpreted in language. However, because these constraints are largely

pragmatic in nature, what they predict is essentially a "usage preference" (Huang 1991: 318). In

other words, most of the constraints identified in these studies allow for the possibility (however

unlikely) that a speaker may deviate from them for one reason or another.

For example, the preference for a co-referential interpretation of the 3SG object pronoun

in the following Chinese sentence (1) is justified by Huang's pragmatics-based analysis of

anaphora:

(3)   **Xiaoming$_i$** de huai piqi gei **ta$_{i/j}$** dai lai le xuduo mafan
         '**Xiaoming's$_i$** bad temper has brought a lot of trouble to **him$_{i/j}$** (Huang 1991: 319)

Specifically, as explained by Huang, "Since *ta* and *Xiaoming* are not clausemate core-arguments,

the DRP does not operate here. Consequently, the use of *ta* will I-implicate a preference for local

co-referentiality" (1991: 319).

However, as indicated by the subscript notation for the pronoun, the grammar readily

permits either co-reference or non-co-reference. Although Huang's analysis predicts a strong

preference for co-reference, it is at least grammatically possible that a speaker use the pronoun *ta* non-co-referentially in this case.[3]

The question arises, then, as to what happens when a speaker deviates (where grammatically possible) from the constraints which govern anaphoric production. More specifically, how does such a deviation affect discourse coherence? For example, how does one explain the difference in perceived coherence (if there is such a difference) between (4c) and (5c), below?[4]

(4) (a)   Juan$_i$ lavó el coche con Pedro$_j$.
    (b)   Ø$_i$/Él$_i$ Limpió afuera mientras que Pedro$_j$ limpió adentro.
    (c)   Ø$_i$ Enceró el capó.

(5) (a)   Juan$_i$ lavó el coche con Pedro$_j$.
    (b)   Ø$_i$/Él$_i$ Limpió afuera mientras que Pedro$_j$ limpió adentro.
    (c)   Ø$_j$ Pasó la aspiradora por los asientos.

Based on general world-knowledge, the interlocutor may successfully resolve the null pronoun in (5c) to Pedro, since it has been established that he is cleaning the inside of the car, and since car seats are located on the inside of a car. However, it remains unclear, based solely on analyses which identify world-knowledge as a constraint on anaphoric interpretation, whether or not the application of world-knowledge to resolve otherwise ambiguous anaphora results in increased processing difficulty or decreased coherence.

---

[3] A non-co-referential usage, although it deviates from the preferences proposed by Huang, does not necessarily render the referent irretrievable: the resolution of anaphora may also be facilitated (or constrained) by various semantic and/or contextual cues. This will be addressed more explicitly in my discussion of Centering Theory (cf. Walker et al. 1998) below.

[4] Example adapted from Walker et al. 1998: 6-7.

1.2.4   Centering Theory: An Application of Constraints

All of the analyses discussed until this point stop short of predicting how discourse is affected when a speaker makes linguistic choices which, although grammatically acceptable, are somehow less than optimal in terms of relative coherence. In this respect, Centering Theory (cf. Walker et al. 1998) is especially valuable. Not only does it delineate a set of constraints which guides the interpretation of anaphora, but it also makes predictions about differing levels of perceived discourse coherence which arise when the most likely (or preferred) interpretation is incorrect or impossible.

Like Accessibility Theory (cf. Ariel 1991) and the Givenness Hierarchy (cf. Gundel et al. 1993), Centering Theory establishes a relationship between the form of a referring expression and the salience of its referent/antecedent. This relationship is then applied to make predictions about the relative level of coherence of inter-sentential anaphora. Before proceeding further, a brief overview of Centering Theory is necessary.

The objective of Centering Theory, as stated by Walker et al. (1998), is to explain how the relationship between the form of an anaphoric expression and the salience of its referent can be applied to predict the level of coherence of a given utterance. As such, within the centering model, it is essential that the potential antecedents of a given utterance (called the forward-looking centers, or Cf set) can be definitively ranked in terms of salience. This ranking of forward-looking centers then serves to determine how the most salient anaphoric reference of a subsequent utterance (Cb), is resolved.

The centering model ranks members of the Cf set for a particular utterance in terms of syntactic role,[5] so that a subject is more salient than an object, which is more salient than a

---

[5] For a justification of using syntactic role in order to evaluate the salience of a particular discourse entity, see Bolinger 1979; Brennan et al. 1987; Reinhart 1981.

discourse entity which occupies an oblique syntactic position: Subject(s) > Object(s) > Other (Walker et al. 1998: 7).[6] Thus, the set of forward-looking centers (or Cf set) in (6a), below, is ranked as follows: *Juan > María > el coche*. Since *Juan* is the most salient entity presented in the utterance, he is the "preferred center" (Cp) of the utterance. Also, if we assume that (6a) is the initial utterance of a particular discourse, then it follows that none of the references therein is anaphoric. Thus, (6a) contains no backward-looking center, or Cb.

(6) (a)   **Juan$_i$** ayudó a **María$_j$** a lavar **el coche$_k$**.
                Cf: [JUAN, MARÍA,COCHE] Cp: [JUAN] Cb: ?

    (b)   Luego, **ella$_j$ lo$_i$** invitó a tomar **una cerveza$_l$**.
                Cf: [MARÍA, JUAN, CERVEZA] Cp: [MARÍA] Cb: [MARÍA]

Likewise, assuming that (6b) is a continuation of the discourse initialized by (6a), the entities in (6b) may be ranked as *ella > lo > una cerveza*, or [María] > [Juan] > *una cerveza*. In this case, the most salient anaphoric reference is [María], since it occupies the subject position. Thus, [María] is the Cb of (6b). Additionally, because [María] happens to occupy the subject position, [María] is also the Cp of (6b).

Once the discourse entities of a given utterance have been ranked in terms of relative salience, the relationship between these entities and their corresponding referring expressions may be applied in order to predict discourse coherence. Within the centering model, there are four levels of relative discourse coherence defined by the relationship between a given

---

[6] This ranking is for discourse entities in English. Languages which tend to employ different or additional syntactic elements (e.g., the "experiencer" locus for psychological verbs in Spanish) require Cf hierarchies which acknowledge those elements. For example, the Cf ranking developed by Taboada for Spanish is the following: EXPERIENCER > SUBJECT > ANIMATE INDIRECT OBJECT > DIRECT OBJECT > OTHER > IMPERSONAL/ARBITRARY PRONOUNS (Taboada 2005: 14).

antecedent and subsequent anaphoric reference to that antecedent, as demonstrated in Table 1, below.

Table 1.1 Transition Types within Centering Theory

	$Cb(U_i) = Cb(U_{i-1})$ OR $Cb(U_{i-1}) = [?]$	$Cb(U_i) \neq Cb(U_{i-1})$
$Cb(U_i) = Cp(U_i)$	CONTINUE	SMOOTH-SHIFT
$Cb(U_i) \neq Cp(U_i)$	RETAIN	ROUGH-SHIFT

These transition types are also ranked in terms of preference or relative discourse coherence. The CONTINUE transition is the most coherent, and therefore preferred to all other transition types. Specifically, as stated by Walker et al.: "The CONTINUE transition is preferred to the RETAIN transition, which is preferred to the SMOOTH-SHIFT transition, which is preferred to the ROUGH-SHIFT transition" (1998: 4).

For example, the transition between an utterance (U_{i-1}) and a subsequent utterance (U_i) is defined as CONTINUE when two criteria are met: First, either the backward-looking center (Cb) must be the same from (U_{i-1}) to (U_i), or there must be no backward-looking center in the initial utterance (U_{i-1}). Second, the Cb of (U_i) must also be the preferred center (Cp) of (U_i). The transition from (7a) to (7b), therefore, is defined as CONTINUE.

(7) (a)   **Juan$_i$** ayudó a **María$_j$** a lavar **el coche$_k$**.
   (b)   **Ø$_i$** Hizo muy buen trabajo.
   (c)   **Ella$_j$** agradeció la ayuda.

Based on the salience rankings discussed for (6), it is established that *Juan* is the most salient entity, or the Cp of (7a). If we assume for explanatory purposes that (7a) is the first utterance of the discourse, then Cb(7a) = [?], meaning that there is no backward-looking center for (7a). Also, since the most salient anaphoric reference in (7b) (in this particular case, there is only one anaphoric reference) is expressed by the null subject pronoun, we know that this null pronoun represents Cb(7b).[7] Furthermore, since Cb(7b) also occupies the most salient position of (7b), we see that Cb(7b) = Cp(7b). Thus, the transition from (7a) to (7b) is a CONTINUE transition because Cb(7a) = [?] and Cb(7b) = Cp(7b).

However, applying the same criteria, we must define the transition from (7b) to (7c) as a SMOOTH-SHIFT. First, we see that Cb(7b) refers to *Juan*, while Cb(7c) refers to *María*. Thus, Cb(7c) ≠ Cb(7b). Second, we see that Cb(7c) occupies the most salient position in (7c). Thus, Cb(7c) = Cp(7c).

1.2.5   What About Plural Anaphora?

Another area which has received little attention in much of the literature dealing with anaphora is the relationship between plural anaphora and their referents. One of the possible reasons for this lack of attention is that plural referents can be constructed differently than single referents (cf. Albrecht and Clifton 1998; Gordon et al. 1999; Koh and Clifton 2001; Moxey et al. 2004; Koh et al. 2005). A singular anaphoric expression necessitates a one-to-one correspondence with the discourse referent to which it is resolved. For example, the discourse referent for any given utterance of the subject pronoun *ella*, by definition, must be exactly one

---

[7] This assumption is based on what Walker et al. (1998) call the Pronoun Rule, which states the following: First, if there are multiple pronouns in an utterance, then one of them must be the Cb. Second, if there is only one pronoun, then that pronoun must be the Cb. The Pronoun Rule "captures the intuition that pronominalization is one way to indicate discourse salience, and that backward-looking centers, Cb's, are often deleted or pronominalized" (1998: 5).

singular entity. However, this is not always the case with plural anaphora. While some plural

anaphora may be resolved to a single instantiation of a plural entity, as in (8b), others may be

resolved to two or more singular entities (split antecedents), as in (9b), below.

(8) (a)		**Los chicos$_i$** fueron a la tienda$_j$.
(b)		Ø$_i$ Iban a comprar unos refrescos.
(9) (a)		**Juan$_i$** fue a la tienda$_k$ para **Pedro$_j$**.
(b)		Ø$_{i+j}$ Iban a dar una fiesta.

The fact that the referent of a plural anaphoric expression may be composed of two or

more singular discourse entities means that those entities may be presented in the discourse with

disjoint syntactic and thematic roles. The referent for the null reference in (9b), for example,

consists simultaneously of the subject/agent (*Juan*) and the prepositional object / beneficiary

(*Pedro*) of the preceding utterance. This presents problems for several of the analyses already

discussed (i.e., Accessibility Theory (cf. Ariel 1991); The Givenness Hierarchy (cf. Gundel et al.

1993); Centering Theory (cf. Walker et al. 1998)), which rely partly on the syntactic prominence

of a given antecedent in order to make predictions about the preferred form of a corresponding

referring expression. In the case of some plural anaphora, the referent may be composed of both

a syntactically prominent and a syntactically oblique antecedent.

Ariel (1991), one of the few previously mentioned authors to address this particular

phenomenon (but see also Cameron 1997; Koh and Clifton 2001; Koh et al. 2005), states quite

simply that "a split antecedent is an inferior antecedent…it is not as salient as a single

antecedent" (1991: 457). Within Accessibility Theory (cf. Ariel 1991), Ariel's

cognitive/pragmatics-based analysis of anaphora, the fact that a split antecedent is not salient

means that corresponding referring expressions tend to require more semantic and phonetic content. Thus, a null expression, which delivers little semantic information and is phonetically void, should be unacceptable when referring to a split antecedent. In order to justify this conclusion, Ariel cites the following example from Hebrew, where the null reference which evokes the split antecedent composed of *Talila* and *Itamar* is unacceptable:

(10) * **Talila$_i$** amra le-**itamar$_j$** se- Ø$_{i+j}$ hiclixu.
      'Talila said to-Itamar that [they] succeeded.'
      (cited in Ariel 1991: 457)

However, as previously demonstrated in (9a-b), null reference to a split antecedent in Spanish is quite possible in at least some cases. Cameron (1997) investigates this phenomenon in a study which compares the accessibility of split antecedents to non-split antecedents (like those in (8a)) in Spanish. Based on results which indicate that non-split antecedents are no more likely to yield null plural anaphora than split antecedents, Cameron concludes that "[Ariel's] structural conception of antecedent inferiority, as distinguishes split from non-split antecedents, does not hold." (1997: 48).[8] Put another way, split antecedents were found to be just as accessible as non-split antecedents. Crucially, this conclusion assumes that the form of a referring expression (in this case, a null vs. an overt subject pronoun) is the only reflection of an antecedent's accessibility.

However, there is evidence within related psycholinguistic literature which suggests that an antecedent's accessibility may be reflected in other ways, namely the relative processing difficulty of a corresponding anaphoric expression. Moxey et al. (2004) demonstrated, not

---

[8] An analysis of the methodology and results of this particular study will be developed in detail in Chapter 2.

surprisingly, that processing difficulty[9] increased for anaphoric expressions which resolved to "an unfocused antecedent" (cited in Sanford et al. 2004). Specifically, plural anaphora which resolved to two syntactically and thematically disjoint antecedents, such as in (11), were significantly more difficult to process than plural anaphora which resolved to two antecedents which shared syntactic and thematic roles, such as in (12), below:

(11) (a)   **Mary**$_i$ cleared up the garden$_k$ for **John**$_j$.
     (b)   **They**$_{i+j}$ enthusiastically made a bonfire$_l$.

(12) (a)   **Mary**$_i$ and **John**$_j$ cleared up the garden$_k$.
     (b)   **They**$_{i+j}$ enthusiastically made a bonfire$_l$.
                            (cited in Sanford et al. 2004)

This difference in processing difficulty suggests that some types of split antecedents in English are more accessible than others, despite the fact that they may be evoked by the same form of referring expression. This, in turn, means that the form of a referring expression may not be a subtle enough measurement by which to determine the accessibility of a split antecedent. Thus, while Cameron (1997) demonstrates that in Spanish, too, the same form of referring expression is used to evoke both split antecedents and non-split antecedents, his results may not justify the conclusion that both types of antecedent are equally accessible.

---

[9] Processing difficulty was measured using an eye-tracker which recorded the amount of time spent reading a particular anaphoric expression and the amount of time spent re-reading previous discourse once an anaphoric expression was introduced.

1.3   Conclusion and Objectives of the Present Study

In summary, if we acknowledge that the production and interpretation of anaphora is governed largely by pragmatic constraints,[10] we must also acknowledge that it is possible for speakers to choose referring expressions with varying degrees of acceptability. In light of this, Centering Theory is valuable because it makes predictions about how a speaker's choice of referring expression affects discourse coherence. However, very little work in Centering deals specifically with anaphora in Spanish. Furthermore, both within Centering literature and the larger volume of literature dealing with anaphora, there has been little focus on plural anaphora – and even less on plural anaphora which resolves to split antecedents. Precisely because plural anaphora may resolve to split antecedents, many of the existing frameworks – including Centering – have yet to propose any constraints which govern their production and interpretation.

Based on the preceding literature review, then, the questions which I will attempt to address in the present study are as follows: Do different types of plural antecedents (non-split antecedents and various types of split antecedents) in Spanish correspond with varying levels of salience or accessibility? And, if so, how might these varying degrees of accessibility affect discourse coherence? Regarding the first question, the primary objective of the present study is

---

[10] Here, pragmatic constraints should be distinguished from constraints which are syntactic or semantic in nature. By syntactic constraints, I mean those which are stipulated by a language's grammar, such as binding Principle B, which states that co-arguments of a predicate are intended to be disjoint unless explicitly marked to the contrary (from Beaver 2004: 15). In example (1), for instance, Principle B stipulates that *John* and *him* are non-co-referential.

    (1)   John$_i$ shot him$_j$.

By semantic constraints, I mean types of constraints which, for example, prevent the resolution of a subject pronoun to an inanimate object in Spanish (as in (2)) or the joining of animate and inanimate objects to form a plural referent (as in (3)).

    (2)   *El puente fue$_i$ construido rápidamente, pero él$_i$ funciona muy bien.

    (3)   Juan$_i$ pasa mucho tiempo en su sillón$_j$ favorito. ??Ø$_{i+j}$ Son muy buenos amigos.

to address the potential problems which arise when frameworks designed to explain the production and/or interpretation of singular anaphora are applied to an analysis of plural anaphora.

Regarding the second question, the Centering model may prove especially useful because it is designed specifically to make predictions about discourse coherence based on the relationship between anaphoric expressions and their referents. Thus, the secondary objectives of the present study are as follows: 1) to consider an expansion of Centering Theory's salience ranking for singular referents which would account for split referents, and 2) to consider whether or not a salience ranking for plural referents may be applied to make predictions about discourse coherence. These objectives may be achieved in part through an analysis of the results of two different experiments, which will be discussed in detail in the following chapters.

CHAPTER 2

THE ACCESSIBILITY OF PLURAL ANAPHORA: A PRODUCTION EXPERIMENT

2.1   Introduction

One of the objectives of the second chapter is to narrow the focus from a discussion of

anaphora in general to a discussion of plural anaphora in Spanish. In doing so, I will offer an

explanation of the seemingly incongruous results of two studies which deal with plural anaphora,

namely Cameron (1997) and the production experiment conducted by Moxey et al. (2004). (I

have selected these two studies in particular because they focus exclusively on the ways in which

speakers access plural antecedents). The discussion of the results and methodologies of the two

studies will lead to a justification of the production experiment conducted for the current

analysis, and will bring to light some of the problems which arise when frameworks designed to

explain the production and/or interpretation of singular anaphora are applied to an analysis of

plural anaphora.

2.2   Cameron 1997 vs. Moxey, Sanford, Sturt and Morrow 2004

As previously discussed, Cameron's 1997 study investigates the accessibility of split vs.

non-split antecedents in Spanish. Before proceeding further, it is important to understand exactly

how Cameron defines these two types of antecedents. Non-split antecedents are "those

antecedents to plural subjects which occur in non-conjoined noun phrases" (1997: 43).

Conversely, antecedents to plural subjects which occur in any other context are considered split

antecedents. Thus, the plural NP *las mujeres* in (1a), below, is defined as a non-split antecedent,

while *él* and *Diego* in (2a) are defined as a split antecedent.

(1) (a)   Y eso era la la devoción de **las mujeres**$_i$ antes.
  (b)   Era…todos los sábados, **ellas**$_i$ cogían un paño con con aceite y…
      (Cameron 1997: 37)

(2) (a)   Sí, de de niños porque **él**$_i$ le andaba siempre en – mucho con **Diego**$_j$
  (b)   y **ellos**$_{i+j}$ salieron mucho los dos juntos.
      (Cameron 1997: 44)

Based on Ariel's claim that "a split antecedent is an inferior antecedent…it is not as

salient as a single antecedent" (1991: 457), and on the idea that the accessibility of a particular

antecedent should be reflected in the form of a corresponding referring expression (cf. Ariel

1991), Cameron predicts the following:

> The frequency of pronominal subject expression should be significantly
> higher in those contexts wherein a plural personal subject takes as an
> antecedent, a split-antecedent pair [such as in (2)], than in those contexts
> wherein a plural subject takes a non-split, single antecedent [such as in (1)].
> (Cameron 1997: 37)

However, data from the speech of 20 native Spanish-speakers (10 from Madrid and 10

from San Juan, Puerto Rico) indicated that the frequency of plural subject expression was

roughly the same for both split and non-split antecedents. Cameron's conclusion, therefore, was

that split antecedents and non-split antecedents are equally accessible.

These results are perhaps surprising, not only in light of Ariel's claim about the relative

inaccessibility of split antecedents (1991: 457), but also considering the seemingly contradictory

results of Moxey et al.'s (2004) study, which deals with the way speakers access various types of split antecedents in English. In the production portion of the study, participants were asked to provide continuations to sentences which presented three different types of split antecedents, represented by the following examples:

(3) Jack and Jill painted the lounge.

(4) Jack painted the lounge with Jill.

(5) Jack painted the lounge for Jill. (Moxey et al. 2004: 348)

In (1), *Jack* and *Jill* share both syntactic and thematic roles (as subjects and co-agents). In (2), although *Jack* and *Jill* fulfill different syntactic roles (as a subject and a prepositional object, respectively), they "are likely to be interpreted as playing common (co-agent) thematic roles" (Moxey et al. 2004: 348).[1] In (3), *Jack* and *Jill* differ both syntactically and thematically, with *Jack* serving as a subject/agent and *Jill* serving as a prepositional object/recipient (2004: 348).

According to the authors, the sharing of syntactic and/or semantic roles by two or more discourse entities should facilitate the creation of what is referred to as a "Complex Reference Object," which "can be thought of as a sum of parts, or as a set of atomic tokens" (for example, the referent represented singularly by *Jill* in (2) would be considered an atom, whereas the plural referent created by the sum of *Jack and Jill* in (2) would be considered a Complex Reference Object) (Moxey et al. 2004: 346). Furthermore, based on the work of Barker (1992), Eschenbach

---

[1] Moxey et. al also discuss the possibility that "*with-* phrases can introduce instruments (e.g., *with a hammer*) while *for-* phrases can introduce a temporal interval (e.g., *for three hours*) (2004: 351). However, these interpretations would be extremely unlikely in sentences where the objects of *with* and *for* are human referents. In Moxey et al.'s experiment, only 5 responses reflected such an interpretation (e.g., *Kevin mowed the lawn with Mary. He should have used a lawnmower*), and no responses in the current study reflected such an interpretation.

et al. (1989) and Kamp & Reyle (1993), the authors maintain that the creation of a Complex

Reference Object is necessary in order to license subsequent plural reference (Moxey et al. 2004:

346). Thus, Moxey et al. predicted that split antecedents such as in (1) would yield the highest

frequency of plural references, while split antecedents such as in (3) would yield the lowest

frequency of plural references.

The results of the production experiment confirmed these predictions. For the conjoined

(*and*) condition (such as in (1)), roughly 60% of the responses contained a plural reference. For

the co-agent (*with*) condition (such as in (2)), the frequency dropped to approximately 40%,

while only around 5% of responses to the benefactive (*for*) condition (such as in (3)) contained

plural references.[2]

In summary, while Cameron's (1997) results demonstrated that two disjoint entities such

as *él* and *Diego* in (2a) yield roughly the same frequency of null plural pronouns as plural NP's

such as *las mujeres* in (1a), Moxey et al.'s (2004) results demonstrate that various types of split

antecedents in English vary significantly in terms of they ways in which they are accessed in

discourse.

2.3   Cameron 1997 vs. Moxey et al. 2004: An Analysis of the Results

The first difference between the two studies which needs to be addressed is that

Cameron's study compares split antecedents to non-split antecedents, while Moxey et al.'s study

compares only different types of split antecedents. However, this discrepancy becomes of

---

[2] These percentages are based on a graphical representation (cf. Moxey et al. 2004: 350). The exact percentages
were not explicitly stated in the text.

minimal theoretical importance when it is examined in light of Moxey et al.'s proposed constraints on the accessibility of split antecedents.

As already stated, in Moxey et al.'s (2004) production experiment, the accessibility of split antecedents was determined to increase according to the extent to which the individual components of that antecedent shared syntactic and/or thematic roles (2004: 347). For example, if the components of a non-split antecedent are unidentified as singular entities, then every entity represented by that antecedent is inextricably linked to the same syntactic and thematic role. In other words, since without additional information we cannot identify the individual components of a plural NP, neither can we separate those components in terms of syntactic or thematic role. Consequently, all else being equal, non-split antecedents should be the most accessible form of plural antecedent available. For example, in (6), every entity represented by *los chicos* is at once the grammatical subject and the agent of the action expressed by the verb *cortaron*. This, according to Moxey et al.'s results, should render *los chicos* as a highly accessible plural referent. Conversely, in (7), since the entities represented by *Juan* and *María* differ both grammatically and thematically, the plural referent composed of *Juan+María* should be relatively inaccessible.

(6)   ***Los chicos*** *cortaron el césped.*

(7)   ***Juan*** *cortó el césped para **María**.*

The question arises, then, as to why Cameron's results did not reflect any significant difference in accessibility between non-split and split antecedents, represented by (6) and (7), respectively?

One possible explanation, already discussed briefly in Chapter 1, is that the form of a referring expression may not be a subtle enough indicator of accessibility to reflect the difference between split and non-split antecedents. In other words, even if there is a difference in relative accessibility between split and non-split antecedents (as the constraints proposed by Moxey et al. suggest), this difference might not be great enough to necessitate a significant differentiation in the form of referring expressions (i.e., the use of an explicit plural pronoun for split antecedents and the use of a null plural pronoun for non-split antecedents). This may explain why Moxey et al. supplemented the results of their production study with an on-line processing experiment (discussed in detail in Chapter 3), which was designed specifically to capture the difference in relative accessibility between various types of split antecedents.

Besides differing slightly in terms of focus – Cameron's (1997) study compares split antecedents to non-split antecedents while Moxey et al.'s (2004) study compares only different types of split antecedents – the two studies also differ methodologically. Cameron's analysis is based on data from naturally occurring discourse in spoken Spanish, while Moxey et al.'s analysis is based partly on a carefully controlled production experiment in written English. This produces at least two noteworthy consequences. First, because the Moxey et al. production experiment was conducted in English, respondents were limited in terms of their choice of referring expression for the entities presented in the test sentences. Since English does not typically allow for null pronouns, plural reference to the antecedents presented in the test sentences could be made only with an explicit pronoun.[3] If, as suggested previously, the form of a referring expression is not a subtle enough indicator of accessibility to reflect differences

---

[3] The instances of plural reference reported by Moxey et al. (2004) consisted entirely of the explicit pronominal forms such as *they* or *them*. While it is theoretically possible that respondents could have made plural reference with demonstratives or other anaphoric expressions such as *these people* or *the two*, no such responses were reported.

between split and non-split antecedents in Spanish (where there is a greater number of possible forms for a given expression), then it will be even less likely to reflect such a difference in English (where there is a smaller number of possible forms for a given expression).

Thus, what we may be seeing in the results of the experiment is not a reflection of different levels of accessibility which correspond with different types of split antecedents, but rather a reflection of what the respondents interpreted the test sentences to be about (e.g., a plural entity vs. a singular entity). In other words, the extent to which two discourse entities are presented as sharing syntactic and thematic roles may affect not only the relative accessibility of those entities as a plural antecedent, but also (and perhaps to a greater extent) the perceived theme of a particular utterance.[4]

Not surprisingly, respondents typically interpreted the theme of sentences with conjoined subjects (such as in (3)) to consist equally of each conjoined entity, which is reflected by the fact that responses to these types of sentences contained the highest frequency of plural reference (Moxey et al. 2004). Conversely, sentences in which the two entities were syntactically and thematically disjoint (such as in (5)) were generally interpreted as being about either one entity or the other, reflected by the fact that responses to these types of sentences contained the lowest frequency of plural reference (and the highest frequency of singular reference[5]) (Moxey et al. 2004). Since plural reference in both cases could only be manifested in the form of an explicit pronoun (*they* or *them*), the results of the production experiment are somewhat inconclusive with respect to the relationship between the accessibility of plural antecedents and the form of a

---

[4] This observation alludes to the effects of the Conjunction Cost on anaphoric reference (cf. Albrecht and Clifton 1998) and to the relationship between the perceived theme of an utterance and the syntactic constituents of that utterance (cf. Blackwell 1998; Bolinger 1979; Brennan et al. 1987; Gundel et al. 1993; Reinhart 1981).

[5] 62% of singular pronominal references in the *for* condition evoked the second discourse entity mentioned – i.e., the oblique argument (Moxey et al. 2004: 350).

corresponding referring expression: plural reference could either be made via the use of an explicit pronoun, or could not be made at all.

A second major consequence of the differing methodologies has to do with the fact that Cameron's (1997) study analyzes naturally occurring discourse, while Moxey et al.'s (2004) production study analyzes readers responses to a single, "out-of-the-blue" test sentence. This means that the context in which anaphoric reference was analyzed differs substantially between the two studies. In the Moxey et al. production experiment, context was determined almost entirely by the content of a given test sentence (any other context would have had to be contributed by extra-textual knowledge possessed by the respondent). In Cameron's study, the context surrounding a given anaphoric expression was determined by all previous and related utterances in the discourse. Before discussing further the effects of these differing contexts, it is helpful to return to the factors which contribute to the accessibility of a given referent, as stated within Accessibility Theory (cf. Ariel 1991).

As discussed briefly in Chapter 1, Ariel proposes three main criteria for determining the accessibility of a particular discourse entity: 1) distance between the antecedent and its corresponding referring expression, 2) the relative salience of the antecedent within the discourse situation (partly determined by syntactic function, and partly determined by what the speaker and/or interlocutor consider "important"), and 3) competition (or, "how many other antecedents can potentially serve as antecedents for the said [referring expression]") (Ariel 1991: 445). In Moxey et al.'s (2004) production study, the distance criterion was held constant: reference to any of the antecedents presented in the test sentences occurred within the next utterance (produced in this case by the respondent). Distance, therefore, was not a variable factor in determining the accessibility of the antecedents presented. Cameron, too, controlled for distance by analyzing

only those references which occurred within 5 clauses of a corresponding antecedent. This leaves us with two remaining criteria: namely, relative salience and competition, which, in light of the aforementioned methodological differences, may have contributed to the dissimilar results of the two studies.

With regard to competition, the number of antecedents which could potentially serve as antecedents for a particular anaphoric expression was held constant in the Moxey et al. production experiment. Because there were only two animate entities presented in each test sentence, a respondent who chose to make plural reference was left with only one possible antecedent for that reference: the Complex Reference Object (cf. Moxey et al. 2004: 346) formed by the sum of those two entities. This means that in every case, a plural antecedent was entirely free of competition for a subsequent plural reference. In the Cameron study, however, the competition criterion was not taken into account. This means that for any given plural reference analyzed by Cameron, there could have been varying types of plural referents which "competed" to serve as the antecedent. First, there could have been more than two singular entities presented within the discourse preceding a plural reference, which means that a plural expression could refer variably to any combination of two or more entities within the preceding discourse. Second, there could have been other plural entities (plural NP's, or non-split antecedents) within the previous discourse which "competed" to serve as the antecedent for a given plural anaphoric expression. Either of these scenarios, according to Ariel (1991: 445), would have decreased the accessibility of a particular plural antecedent (whether split or non-split), increasing the likelihood that plural reference would take the form of an explicit pronoun instead of a null pronoun.

The final criterion identified by Ariel as being instrumental in determining the accessibility of an antecedent is that of relative salience. The salience of an antecedent may affect the resolution of a corresponding referring expression in two main ways. The first of these, which I will discuss in further detail in Chapter 3, is related to the fact that the salience of a given discourse entity is largely determined by the syntactic function of that entity (i.e., subject, object, or oblique). The second, which I will discuss here, is related to the fact that salience is also determined to some extent by what the speaker and/or interlocutor consider important within a particular discourse situation. As Ariel states, "[the] speaker, addressee, [and] important people or things in [the] addressee's life (spouse, children), etc., are more easily retrievable than other representations" (1991: 445). This observation alludes to the importance of shared knowledge in the resolution of anaphora. According to several pragmatics-based analyses of anaphora (cf. Blackwell 1998; Ranson 1991; Huang 1993), knowledge shared by discourse participants, whether it comes via explicit mention in the discourse or via more general or world-knowledge, acts as a constraint which may serve to facilitate the resolution of (otherwise ambiguous) anaphoric references.

For example, Ranson (1991) discusses a specific situation in which knowledge shared by the speaker and interlocutor (provided by information stated explicitly in the discourse) serves to clarify the referent of the null subject for the verb *cogió*:

(8)   Pero el niño$_i$ ya es un hombre…como yo más o menos, con gafas. Porque
        cuando pequeño $\varnothing_i$ tuvo el sarampión…$\varnothing_{i+j}$ estaban acostados, él$_i$ y
        María$_j$. $\varnothing_{i+j}$ Se levantaron, le dio el aire, y entonces a causa de aquello,
        pues, **$\varnothing_i$** cogió una endeblez en el nervio del ojo derecho.
        (Ranson 1991: 143)

As Ranson explains, "the knowledge provided in the text itself that the speaker's son, and only

his son, had *sarampión*, allows the hearer to assign the verb *cogió* to the speaker's son, rather

than to his daughter…which was the last [explicit] reference made" (1991: 143).

Another example which demonstrates how shared knowledge may contribute to the

resolution of anaphora is taken from Blackwell (1998).

> (9)　　L: No te digo que no, digo: pero mujer, no tengo:, aunque no tenga: bañera::
> 　　　　C: Te lavás! Claro. Natural. Natural. Natural.
> 　　　　L: Pues cada::=
> 　　　　C: =Lo que haga falta!
> 　　　　L: =Pues cada quince días o veinte, yo me lavo una noche, ¿sabes? Toda. Y **él**$_i$ me
> 　　　　　　　da por la espalda porque ((laughs)) no llego.
> 　　　　C: Que sí, que no llegas, es / / natural.
> 　　　　L: Las cosas / / como son.
> 　　　　C: Que sí por dios, en el matrimonio es / / así.
> 　　　　L: Y otra noche se lava **él**$_i$, y yo **le**$_i$ doy también…
> 　　　　(Blackwell 1998: 616)

In this case, the knowledge which contributes to the resolution of the pronouns *él* and *le* does not

stem from information stated explicitly in this particular conversation, but rather from the more

general knowledge – shared by the speaker and interlocutor – that L is married. Specifically,

"C's mention of *el matrimonio* confirms that she recognizes the intended referent" of the

pronoun *le* and both occurrences of the pronoun *él* (Blackwell 1998: 616).

Besides alluding to the importance of shared knowledge in the resolution of anaphoric

expressions, Ariel's (1991) definition of antecedent salience also emphasizes the significance of

what Dahl refers to as "egophoric reference," defined as "reference to speech act participants and

generic reference" (2000: 37). Ariel maintains that references to one (or both) of the speech act

participants "are more easily retrievable than other representations" (1991: 445). This is because

the physical presence of the discourse participants within the discourse situation makes them highly accessible as referents (cf. Dahl 2000; Gundel et al. 1993; Ranson 1991).[6] Thus, all else being equal, first- and second-person referents should be more accessible than third-person referents. Consequently, the form of a referring expression which evokes one of the discourse participants as its referent would be more likely to take the form of a null pronoun.

Because Cameron's (1997) data was taken from a somewhat naturally-occurring discourse situation (where context was determined by a relatively large body of text), it is possible that both shared knowledge (textual or extra-textual) and egophoric reference may have affected the interpretation and/or production of anaphora. While the difference between egophoric (first and second-person) references and non-egophoric (third-person) references was not reflected in the results (94% of egophoric references were null, while 92% of non-egophoric references were null)[7], it is my opinion that an analysis which rules out the possibility of such interference would be more methodologically sound.

2.4   Justification of the Current Experiment

This brings us to the link between the previously discussed analyses of plural anaphora and the justification of the methodology for the current production experiment. In summary, Cameron's (1997) study and Moxey et al.'s production experiment (2004) differ in two key ways: the language in which the study was conducted, and the context in which anaphoric reference was analyzed. Since English and Spanish make use of different forms of referring

---

[6] For example, in any given conversation or discourse, I need not specify the referent for the pronoun *I* in order to license its use (e.g., ***Buck Pennington**$_i$ likes ice cream.*I$_i$ also like cake.*) However, in order to license the use of a third-person pronoun, its referent must be specified by 1) a previous, more specific reference (cf. Gundel et al. 1993), 2) contextual knowledge (such as in (8)), or 3) general and/or world-knowledge (such as in (9)).

[7] For the entire corpus, approximately 87% (925/1060) of all first- and third-person references were null (Cameron 1997: 48).

expressions, it is difficult to compare the results of a study whose analysis is based on the form

of referring expressions in English to the results of a study whose analysis is based on the form

of referring expressions in Spanish. Similarly, it is difficult to compare data drawn from a

naturally-occurring discourse situation (where context is determined by a relatively large body of

text) to data drawn from responses to a single, "out-of-the-blue" test sentence (where context is

determined almost entirely by the content of that sentence).

Considering the effect of context on anaphoric reference, the question arises as to how

one should go about conducting an analysis of plural anaphora. The answer to this question

depends ultimately on the goal of the analysis. As we have seen, the production and/or

interpretation of plural anaphora is governed by various constraints (most of which are the same

for singular anaphora): the extent to which the components of a plural antecedent share syntactic

and/or thematic roles (Moxey et al. 2004); the salience of a given antecedent (Ariel 1991);

distance between the antecedent and a corresponding referring expression (Ariel 1991);

competition (Ariel 1991); egophoricity (Dahl 2000); textual knowledge (Ranson 1991); and

extra-textual knowledge (e.g., general and/or world knowledge) (Blackwell 1998; Huang 1993;

Ranson 1991).

One of the primary goals of the current study is to offer an analysis of plural anaphora

which fits within Centering Theory, so as to make preliminary predictions about how the

interpretation and/or production of plural anaphora affects discourse coherence. Centering

Theory, as previously discussed, is concerned with the relationship between the relative salience

of a discourse entity and the form of a corresponding referring expression. Furthermore, within

the Centering framework, relative salience is determined primarily by syntactic role (Walker et

al. 1998: 7).[8] However, since the current analysis focuses on plural anaphora, it is also necessary

to consider the results of Moxey et al.'s (2004) study, which suggest the significance of the

thematic roles of the constituents of a split antecedent in determining whether the perceived

theme of a given utterance is a singular or plural discourse entity. It follows, then, that the

current study is primarily concerned with the effect of the syntactic and thematic roles of the

components of a plural antecedent.

In order to analyze the effects of the syntactic and thematic roles of the constituents of a

plural referent on subsequent anaphoric reference, it is important to ensure that other influential

factors – such as distance, competition, egophoricity, extra-textual knowledge, etc. – are either

excluded or severely limited. Therefore, the current study is an adaptation of the Moxey et al.

(2004) production study, which controls (as much as possible) the constraints on anaphoric

reference by limiting the discursive context to a single, "out-of-the-blue" utterance.

The most important difference between the Moxey et al. (2004) production experiment

and the current production experiment is that the current experiment is conducted in Spanish.

Participants will be less limited in the types of referring expressions they can choose because

Spanish allows for the use of null subject pronouns. Consequently, the results of the current

study are more easily compared to the results of Cameron's (1997) study. More importantly,

analyzing anaphoric expression in Spanish should reflect more accurately the relationship

---

[8] It should be noted that at this point in time, Centering Theory is unable to account for certain constraints on anaphoric reference which are purely pragmatic: namely, general and/or world-knowledge constraints. While this is a significant limitation – especially considering the analyses of Blackwell (1998) and Huang (1993) – quantifying the effect of context on anaphoric reference would be exceedingly difficult. As Cote states, a salience ranking for the antecedents within a discourse segment should "[concern] itself with linguistic features, rather than with features of the more general and largely unmanageable realm of knowledge representation" (1998: 59). As such, an incorporation of general or world-knowledge constraint on anaphoric reference is beyond the scope of the current study.

between the accessibility of an antecedent and the form of a corresponding referring expression. Instead of the binary distinction available to participants in the Moxey et al. study (i.e., the use of an explicit plural pronoun or the re-introduction of singular discourse entities via name), participants in the current study will be able to use null plural pronouns, explicit plural pronouns, and names.

Other than the difference in language, the same limitations which apply to the Moxey et al. production study also apply to the current production study. First, the study does not compare the accessibility of split vs. non-split antecedents, but only different types of split antecedents. As previously explained, however, this difference should be of minimal theoretical importance. Second, if the form of a referring expression was not a subtle enough indicator of accessibility to reflect differences between split and non-split antecedents in Cameron's study, then it is possible that the difference will not be reflected in the current study, either. The benefit of utilizing Moxey et al.'s methodology, however, is that because of the limited context, antecedent accessibility in the current study should be determined almost entirely by syntactic and/or thematic roles. As a result, a lack of correspondence between a particular antecedent and the predicted form of a corresponding referring expression cannot be readily attributed to factors such as competition.

2.5   Methodology

21 native Spanish-speakers from Cochabamba, Bolivia were asked to provide continuations to 3 different types of test sentence, in which split antecedents were introduced with various syntactic and semantic relationships, represented by (10-12), below.

10)   **Juan y María** limpiaron las casas.

11)   **Juan** limpió la casa **con María.**

12)   **Juan** limpió la casa **para[9] María**.

Each test sentence consisted of one inanimate entity and two singular, animate entities which were introduced via proper name. If the inanimate object served as the direct object of a transitive verb, it was introduced as a plural NP (as in (10)). This was done so as not to further facilitate the formation of a plural referent by having the two animate entities acting on the same object, such as in *Juan y María limpiaron la casa.* In the case of *Juan y María limpiaron las casas*, it is at least feasible to construct a mental representation in which *Juan* and *María* realize separate verbal events, possibly even at separate times. This interpretation would be very unlikely for cases such as *Juan y María limpiaron la casa.*

Each respondent received a packet of 33 typed sentences: 21 test sentences with each different construction (modeled above) appearing 7 times, and 12 filler sentences, such as the following:

(13)   Paró la lluvia y salió el sol.

Respondents were asked to provide a hand-written continuation consisting of at least one complete sentence for each prompt sentence in the packet. Responses which did not consist of at

---

[9] The structure of this test sentence is meant to parallel that of the benefactive (*for*) condition investigated by Moxey et al. (2004: 348) (see example (5)). It should be noted that while several native speakers of Spanish have endorsed the use of the Spanish preposition *para* to represent this condition, others have noted that the use of the preposition *por* might be a more accurate translation. This applies not only to the PARA condition used in the production experiment discussed in this chapter, but also to the PARA condition used in the processing experiment discussed in Chapter 3.

least one independent clause, such as (14), and responses which did not include at least one anaphoric reference, such as (15), were excluded from the analysis.

(14)   Antonio y Sofía entregaron los paquetes. *A todos los empleados.*

(15)   Jesús hizo las maletas para Rosario. *Sale el vuelo hacia España.*

Continuations which consisted of at least one independent clause were analyzed according to the way in which reference was made to the entities presented in the test sentence. The form of each anaphoric reference (name, explicit pronoun, or null pronoun), the syntactic role fulfilled by the referring expression (subject, object, or other), and the syntactic role of the evoked referent (*Juan*, *María*, *Juan+María*, inanimate entity, or event) were recorded. The following examples serve to demonstrate how the analysis was carried out.

(16)   Manuel$_i$ y Carmen$_j$ hicieron las camas. *Luego Ø$_{i+j}$ se fueron al trabajo.*

(17)   Juan$_i$ limpió las casas con María$_j$. **Ellos$_{i+j}$** *trabajan siempre juntos.*

(18)   Arturo$_i$ compró los regalos para Ana$_j$. *Ø$_i$ Quiere conquistar**la**$_j$.*

In (16), the anaphoric reference was recorded as a null subject pronoun which evoked the plural referent *Manuel+Carmen*. In (17), the reference *ellos* was recorded as an explicit subject pronoun which evoked the plural referent *Juan+María*. In (18), two anaphoric references were recorded: a null subject which evoked *Juan* (the subject of the previous utterance), and a direct object pronoun (a clitic) which evoked *Ana* (the prepositional object, or oblique entity, of the previous utterance.

Anaphoric references made via possessive pronouns were excluded from the analysis because in the majority of cases the referent (either *Juan, María, Juan+María,* or the inanimate entity) was unidentifiable. In (19), for example, it is unclear whether the pronoun *su* refers to *Eduardo, Catalina,* or the (perhaps unlikely) sum of *Eduardo+Catalina*.

(19)   Eduardo preparó los pasteles con Catalina. *Porque es **su** cumpleaños.*

For responses where multiple references were used to evoke the same referent, only the first reference was recorded for analysis. This was done in order to avoid analyzing references for which the antecedent was a previous reference made by the participant. In (20), for example, only *los dos* was recorded as a reference to *Emilio+Pamela*.

(20)   **Emilio$_i$** cocinó las papas con **Pamela$_j$**. *A **los dos$_{i+j}$ les$_{i+j}$** encantan las papas fritas, por eso siempre $\emptyset_{i+j}$ cocinan juntos.*

2.6   Predictions

In accordance with Cameron's (1997) study, the results of the current experiment should demonstrate that it is possible to for split antecedents to form highly accessible complex reference objects, expressed anaphorically by a null plural pronoun. That is, respondents should be able to make null plural reference to the two animate entities presented in the prompt sentence, regardless of their syntactic and semantic relationship.

If the form of a referring expression is an accurate indicator of accessibility, then the highest ratio of null plural references to explicit plural references (i.e. *ellos, ambos, los dos,* and

etc.) should be produced in response to test sentences with conjoined antecedents, as in (10). Prompts in which the animate antecedents share a thematic role (co-agent) but differ syntactically, as in (11), should yield a lower ratio of null to explicit plural reference. Prompts in which the animate antecedents share neither syntactic nor thematic roles should yield the lowest ratio of null to explicit plural reference.

If the form of a referring expression is not an accurate indicator of accessibility, then the ratio of null plural reference to explicit plural reference should not differ significantly across prompt type. If this latter possibility proves to be true, then, like the Moxey et al. production experiment, the current production experiment will be somewhat limited in what it reveals about the accessibility of different types of split antecedents. At the very least, however, the data should reveal that the extent to which antecedents share syntactic and thematic roles correlates directly with the extent to which respondents interpret the theme of the test sentence as being about a plural entity. For sentences with conjoined antecedents (such as in (10)), the frequency of plural reference should be the highest; for sentences with syntactically and thematically disjoint antecedents (such as in (12)), the frequency of plural reference should be the lowest.

2.7   Results

As a starting point, it is worthwhile to note whether the animate entities introduced in the test sentences were referenced as singular or plural entities. For the purpose of uniformity, I will refer to the animate entities as *Juan* and *María*, as they appear in examples (10-12). Thus, *Juan* will represent reference to the first element of the conjoined subjects for the Y (conjoined) condition, and reference to the subject for the CON (different syntactic roles, shared thematic role) and PARA (different syntactic and thematic roles) conditions. *María* will represent

reference to the second element of the conjoined subjects in the Y condition and reference to a prepositional object in the CON and PARA conditions.

Figure 2.1 compares plural vs. singular references to animate entities (*Juan* and/or *María*) across conditions. Continuations to the Y condition contained the highest frequency of plural references: 96% (72/75) of all anaphoric references to animate entities in the Y condition evoked the plural referent *Juan+María* (as opposed to singling out either *Juan* or *María*). The frequency of plural reference dropped to about 63% (57/90) for the CON condition, and was negligible for the PARA condition (3.7% (128/133)).[10]
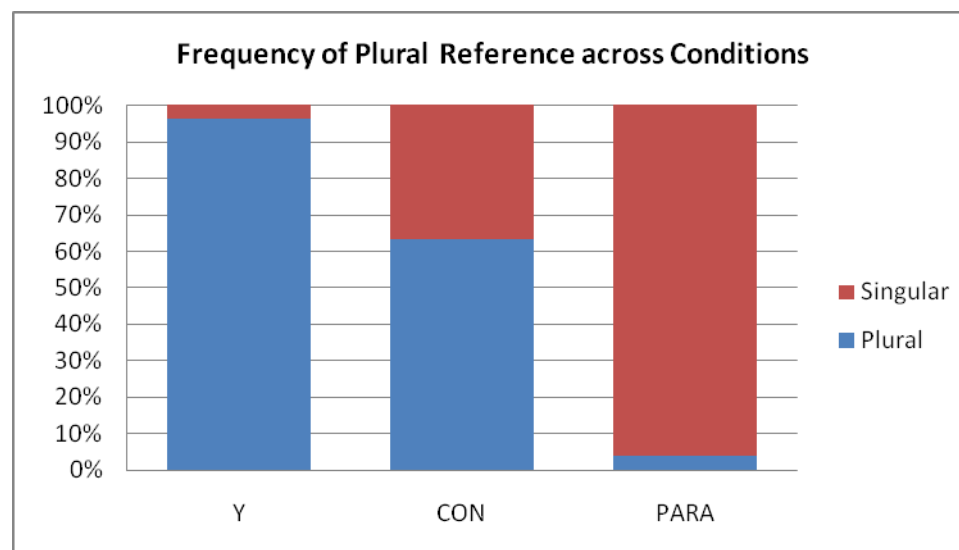

Figure 2.1. Frequency of Plural Reference across Conditions

Figure 2.2 compares the frequency of the forms of referring expression (null subject pronouns vs. explicit subject pronouns) used to reference the plural entity *Juan+María* across conditions. In order to consistently compare null expressions to explicit expressions across

---

[10] These differences are significant at the .0001 level.

conditions, I have only considered references which served as the grammatical subject of the continuation. Other grammatical functions (such as verbal complements and prepositional objects) typically do not allow for null reference in Spanish. As we can see, the ratio of null references to explicit references was roughly the same for both the Y and the CON conditions: 73% (47/64) of subjects which evoked *Juan+María* were null for the Y condition, and 76% (37/49) were null for the CON condition. Curiously, although references to *Juan+María* which appeared in the subject position were relatively infrequent for the PARA condition (only 5 instances total), all of them were null.
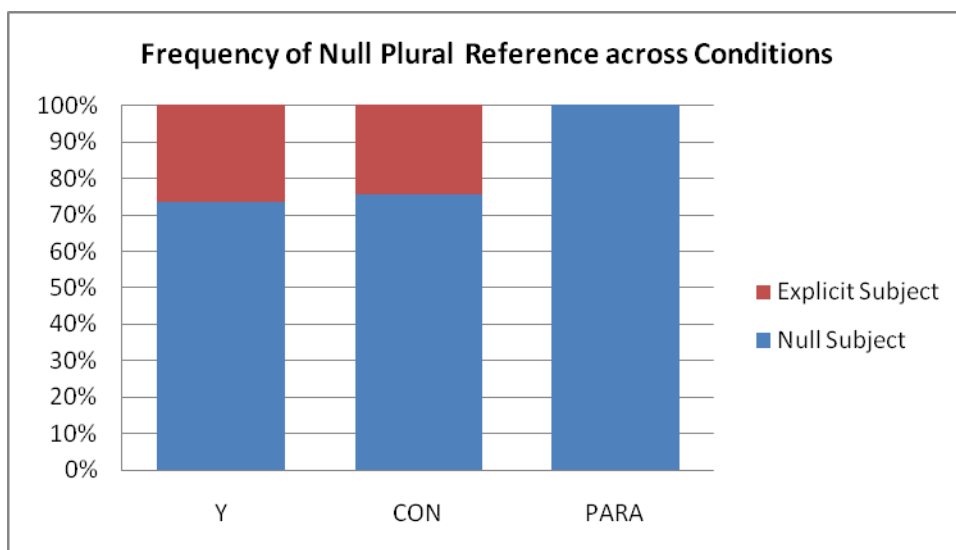

Figure 2.2. Frequency of Null Plural Reference across Conditions

2.8   Discussion and Conclusion

As expected, the data suggest that plural reference is greatly facilitated when split antecedents share syntactic and/or thematic roles (cf. Figure 2.1). This corroborates the results of the Moxey et al. (2004) production study. The data also suggest an inverse tendency: that singular reference is greatly facilitated when discourse entities do not share syntactic and/or

thematic roles. This is evidenced by the fact that the frequency of singular reference increased significantly as the sharing of roles decreased (cf. Figure 2.1).

Additionally, because the study was conducted in Spanish, the data have allowed us to investigate whether the sharing of syntactic and/or thematic roles also affects speakers' choice of plural referring expression (either null or explicit). Given the results displayed in Figure 2.2, it seems that this is not the case. Speakers were no more likely to employ null expressions when split antecedents shared both syntactic and thematic roles, and no less likely to do so when split antecedents did not share roles. If the data from the current experiment were able to accurately reflect the relative accessibility of split antecedents, then we would expect to see significant differences not only in the ratio of singular to plural references across condition types, but also in the ratio of null plural to explicit plural references (in the subject position) across condition types.

This is perhaps further evidence that the higher frequency of plural reference to conjoined entities (and the almost negligible frequency of plural reference to syntactically disjoint entities) from the Moxey et al. (2004) production study is a reflection of the perceived theme of the test sentence, and not necessarily a reflection of the difference in accessibility between various types of split antecedents.

This leads us to our first possible conclusion: that the extent to which antecedents share syntactic and thematic roles, while affecting the perceived theme of an utterance, does not affect the accessibility of those antecedents as a plural referent. However, this conclusion seems difficult to justify considering the close relationship between the theme of a given utterance and the most salient syntactic entity of that utterance – usually the subject/agent[11] (cf. Blackwell

---

[11] The theme of an utterance does not always correspond with the grammatical subject. Evidence of this comes not only from Japanese, where the empathy locus is significant in determining theme (Di Eugenio 1998: 116), but also

1998; Bolinger 1979; Brennan et al. 1987; Gundel et al. 1993; Reinhart 1981). If both the theme and the most salient entity of a given utterance are determined largely by syntactic function, then it seems unlikely that continuations would reflect the effect of syntactic role on the perceived theme of a test sentence (as the data suggest), while not reflecting the effect of syntactic role on the accessibility of split antecedents.

While it is possible, as previously discussed, that the form of referring expression in Spanish is simply not a precise enough measurement to reflect the accessibility of different types of split antecedents, I would like to suggest another possible explanation. In the discussion of the contrasting methodologies of the Cameron (1997) study and the Moxey et al. (2004) study, I pointed out that differences in the context in which anaphoric reference takes place could have affected antecedent accessibility. Specifically, if respondents in the Moxey et al. study (or in the current one) chose to make plural reference, there was only one possible antecedent for that reference: the Complex Reference Object formed by the sum of the two animate entities presented in the test sentence (i.e., *Juan+María*). This means that all plural, animate referents in the study were entirely free of competition – there were no other plural antecedents available within the discourse (i.e., the test sentence). According to Ariel, this lack of competition increases the accessibility of a given antecedent (1991: 445).

Another factor which contributes to antecedent accessibility, again according to Ariel (1991: 445), is the syntactic "prominence" (or salience) of that antecedent within a given discourse segment. It is important to note here that the salience of one entity is ranked according

from western languages such as English and Spanish, in cases of "non-agentive psychological verbs such as *interest*, *seem;* perception verbs such as *feel*, *appear*; and in general, expressions that refer to a character's point of view such as *The thought crossed her mind*. With such expressions, it is the experiencer, which is often in object position, rather than the grammatical subject," which serves as the theme of the utterance (Di Eugenio 1998: 116). However, since none of the test sentences used in the current experiment introduced these types of verbs, empathy could not have affected the perceived theme.

to the salience of other entities presented in the same discourse segment. In other words, salience is relative. Presumably, this is the reason why frameworks such as Centering Theory rely on hierarchical salience rankings (i.e., subject>object>other), rather than assigning some absolute "salience value" to a particular syntactic function. This type of relative ranking for singular entities is somewhat stable because an object or an oblique discourse entity does not typically exist without a subject. However, if we could somehow fabricate an utterance which consisted solely of a direct object (for example), then that object would, by default, occupy the most salient position within the utterance.

I argue that this is precisely the sort of situation created by the test sentences both in the Moxey et al. (2004) production study and in the current experiment. Although in the CON and PARA conditions speakers are presented with 2 singular, animate entities which differ in terms of syntactic prominence, respondents are presented with only one antecedent choice if they choose to make plural reference. By default, then, the referent formed by the sum of *Juan+María*, is the most syntactically prominent plural entity available, regardless of the syntactic roles of its individual components.

Thus, according to the criteria proposed by Ariel (1991: 445), for all three conditions in the experiment, the referent formed by the sum of *Juan+María* represents the most highly accessible type of referent possible: First, it is entirely free of competition; second, it occupies the most salient syntactic position within the discourse segment; and third, in terms of distance, it is minimally separated from corresponding anaphoric references. This explains why the ratio of null plural reference to explicit plural reference did not decrease from the Y condition to the CON and PARA conditions.

This conclusion also suggests that the notions of "prominence" and "competition" may need to be slightly adjusted in order to accommodate an analysis of split antecedents. Based on the results of the current experiment, the syntactic prominence of a split antecedent should be determined as relative to the syntactic prominence of other plural referents within the same discourse segment, rather than according to the syntactic prominence of a split antecedent's individual components. Likewise, the data seem to suggest that competition for a plural reference arises only when more than one plural referent is available; singular referents do not compete with plural referents for plural references.

These conclusions will be investigated further in the following chapter, where the results of the current production experiment will be supplemented by an on-line processing experiment.

CHAPTER 3

THE ACCESSIBILITY OF PLURAL ANAPHORA: A PROCESSING EXPERIMENT

3.1   Introduction

In the previous chapter, the results of the production study demonstrated that frequency of plural reference increased according to the extent to which split antecedents shared syntactic and/or thematic roles. Curiously, however, the extent to which split antecedents shared syntactic and/or thematic roles did not seem to affect respondents' choice of form of referring expression. The ratio of null plural reference to explicit plural reference was roughly the same for the Y and CON conditions; and while only 5 plural subjects were recorded in responses to the PARA condition, all were null.

This leaves us with two possible conclusions concerning the accessibility of split antecedents in Spanish. First, it is possible that the form of a referring expression is not a precise enough measurement by which to gauge the accessibility of different types of split antecedents. This conclusion would also explain the results of Cameron's (1997) analysis without requiring that we categorize all types of plural antecedent as equally accessible. Second, it is possible that different types of split antecedents (or different types of plural antecedents) do not differ in terms of accessibility. This conclusion seems unlikely given that the type of split antecedent presented in a given utterance significantly affects the perceived theme of that utterance. If the sharing of roles facilitates plural reference to two discourse entities, then it should also increase the accessibility of those entities as a plural referent. However, if the latter conclusion proves to be

true, then, as Cameron states, "a structural conception of antecedent inferiority, as distinguishes

split antecedents from non-split antecedents, does not hold" (1997: 48).

      In order to evaluate these preliminary conclusions, further experimentation is needed.

Moxey et al. (2004) supplemented the results of their production study with an on-line

processing study which was designed to measure the processing difficulty of plural references to

different types of split antecedents. As in the production experiment, participants read various

sentences which presented three different types of split antecedents: conjoined antecedents (the

and condition), antecedents which shared a common agentive role (the *with* condition), and

antecedents which shared neither syntactic nor thematic roles (the *for* condition), demonstrated

in (1a-c).

          (1) (a) **Mary$_i$** and **John$_j$** cleared up the garden.
             (b) **Mary$_i$** cleared up the garden with **John$_j$**.
             (c) **Mary$_i$** cleared up the garden for **John$_j$**.
                (cited in Sanford et al. 2004: 158)

          (2)   **She$_i$/They$_{i+j}$** enthusiastically made a bonfire.
                (cited in Sanford et al. 2004: 158)

After reading the first sentence, participants were presented with a second sentence which made

either singular reference to *Mary* or plural reference to *Mary+John*, demonstrated in (2). The

processing difficulty of the referring expression in the second sentence was measured using an

eye-tracking device.[1]

---

[1] The eye-tracker was used to record both the amount of time spent reading a particular anaphoric expression and the amount of time spent re-reading previous discourse once an anaphoric expression was introduced.

As discussed in Chapter 1, the results of the processing study demonstrated that plural anaphora which resolved to syntactically and thematically disjoint antecedents were significantly more difficult to process than plural anaphora which resolved to antecedents which shared syntactic and/or thematic roles (Sanford et al. 2004: 158). Specifically, the plural pronoun *they* was more difficult to process when it followed sentences like (1c) than when it followed sentences like (1a).

However, because the study was conducted in English, only one form of plural reference could be analyzed: the explicit plural pronoun *they*. Consequently, the results of the Moxey et al. processing experiment provide only limited insight into the relationship between the accessibility of a split antecedent and the form of a corresponding referring expression. Nevertheless, the fact that differences in processing difficulty were recorded for the same plural referring expression suggests that differences in the accessibility of split antecedents do not necessarily correspond with differences in the form of corresponding plural referring expressions (at least in English). In other words, determining the accessibility of a split antecedent in English calls for an analysis of relative processing difficulty, rather than an analysis of the form of referring expressions.

One of the ways in which processing difficulty has been measured successfully in order to make claims about the accessibility of split antecedents is through a self-paced reading experiment. For example, Gordon et al. (1999) analyzed mean reading times in order to investigate the "prominence" of discourse entities introduced by a conjoined NP (1999: 363). Participants were asked to read various passages consisting of three sentences, modeled in Table 2.1, below.

Table 3.1. Model Test Sentences for Gordon et al. 1999

Sentence 1:	John and Mary went to the store.
Sentence 2:	**They/He/John** wanted to buy candy.
Sentence 3:	The store was closed.

In the first sentence, two discourse entities were introduced in the form of a conjoined NP. In the second sentence, reference was made either to one of the entities in the conjoined NP (via proper name or subject pronoun) or to the plural entity represented by the sum of the NP's individual components. The sentences in each passage were presented one at a time on a computer screen. Participants indicated that they had read a sentence and initiated the appearance of the next sentence by pressing the space bar on the keyboard of a personal computer. In order to measure the processing difficulty of the anaphoric reference to the discourse entities introduced by the conjoined NP in the first sentence, reading times were recorded for the second sentence (i.e., the time between the appearance of the second sentence and the appearance of the third sentence, as initiated by the participants) in milliseconds. To ensure that the participants read the passages as quickly as possible and paid attention to the content, participants were asked to answer a true/false question appeared at the end of each passage. Feedback was given in the case of an incorrect answer.

The results of the study indicated that mean reading times were greater for sentences which made singular reference – via either proper name (1836ms) or subject pronoun (1864ms) – than for sentences which made plural reference to the conjunct as a whole (1678ms). If reading times are viewed as a reflection of processing difficulty, these results suggests that "the collective entity realized in subject position by a conjoined NP is more prominent and accessible than its component entities" (Gordon et al. 1999: 366).

The processing experiment discussed in the following sections makes use of the methodology employed by Gordon et al. (1999) in order to investigate the relationship between the accessibility of different types of plural antecedents and the form of corresponding referring expressions in Spanish

3.2   Methodology

The current experiment differs from the Moxey et al. (2004) experiment in three main ways. First, as already stated, the current experiment is conducted in Spanish, rather than English. This means that instead of comparing the processing difficulty of one form of plural reference (i.e., the explicit pronoun *they*) to one form of singular reference (i.e., the explicit pronoun *she* or *he*), the methodology allows for a direct comparison of two different forms of plural referring expressions: null (Ø) and explicit (i.e., *ellos* or *ellas*). Second, instead of measuring only the processing difficulty of references to various types of split antecedents, the current experiment is designed to measure the processing difficulty of plural references to both split and non-split antecedents. Third, while processing difficulty in the Moxey et al. (2004) experiment was measured with an eye-tracking device, the current experiment measures processing difficulty with self-paced reading times. Again, this particular methodology is adapted from a processing experiment conducted in English by Gordon et al. (1999) (but see also Gordon et al. 1993; Gordon and Hendrick 1997; Kennison and Gordon 1997; Koh and Clifton 2001; Koh et al. 2005).

22 Native Spanish-speakers from 9 different countries completed a self-paced reading experiment which was conducted using E-Prime Psychology Software, version 1.2. Participants were asked to read various passages consisting of three sentences which appeared one at a time

on a computer screen. The first sentence in each passage introduced either a non-split antecedent

or some type of split antecedent, as demonstrated by in the Table 3.1.

Table 3.2. Test Sentences for the Current Study

Sentence 1:	(a) **Los chicos$_i$** limpiaron las casas.
	(b) **Juan$_i$** y **María$_j$** limpiaron las casas.
	(c) **Juan$_i$** limpió las casas con **María$_j$**.
	(d) **Juan$_i$** limpió las casas para **María$_j$**.
Sentence 2:	**Ø$_{i/i+j}$/Ellos$_{i/i+j}$** tenían mucho que hacer.
Sentence 3:	Las casas estaban muy sucias.
True/False:	Juan y María limpiaron los coches.

The second sentence contained either a null plural reference or an explicit plural

reference which evoked the animate antecedents presented in the first sentence. The third

sentence in the series did not contain any meaningful variables, but was included so that the

stimulus sentence (i.e., the second sentence) did not appear at the end of the series. At the end of

each series of three sentences, a true/false question about the content of those sentences was

included in order to ensure that participants actually read the passages.

Different combinations of the sentences modeled in the first position, above, with those

modeled in the second position yield a total of eight experimental conditions, depending on the

type of plural antecedent and the form of plural reference to that antecedent:

1. NON-SPLIT → null/explicit
2. Y → null/explicit
3. CON → null/explicit
4. PARA → null/explicit

80 test passages, each containing a combination of the types of sentences modeled in Table 3.1, were combined with 85 filler passages for a total of 165 passages. The passages were grouped into 4 test blocks of 40 passages each (20 test passages and 20 filler passages) and an initial practice block of 5 filler passages. Within the 4 test blocks, the 8 experimental conditions were divided so that each block contained either 2 or 3 instances of each condition. Each participant completed each of the 4 test blocks, which took an average of approximately 40 minutes.

Participants completed the experiment on a desktop computer. Each sentence in a series appeared one at a time in the center of the computer screen. In order to move from one sentence to another, participants pressed the space bar. After each group of three sentences, participants pressed the space bar in order to move on to the true/false question for that passage. Participants answered the question by pressing appropriately labeled keys. A feedback screen appeared as soon as participants responded to the question, indicating both response time and whether the response was correct or incorrect. Response time feedback was provided in order to encourage participants to move through the experiment as quickly as possible; accuracy feedback was provided in order to encourage participants to pay attention to the content of the test passages.

In order to measure the processing difficulty of the referring expressions presented in the second sentence, participant response time (i.e., the time between the appearance of the second sentence and the appearance of the third sentence) was recorded in milliseconds.

3.3   Predictions

If different types of plural antecedents differ significantly in terms of accessibility, then the processing difficulty for corresponding plural referring expressions should also differ.

Specifically, if the accessibility of a plural antecedent increases according the extent to which the components of that antecedent share syntactic and thematic roles, then the processing difficulty for a plural reference should decrease according to the extent to which the entities evoked by that reference share syntactic and semantic roles. Thus, for example, plural references to antecedents in the NON-SPLIT condition (modeled in Table 3.1, Sentence (1a)) should be more easily processed than plural references to antecedents in the PARA condition (such as in Table 3.1, Sentence (1d)).

Furthermore, if differences in accessibility are at all reflected in the form of referring expression, as Ariel (1991) maintains, then we may also expect processing difficulty to be affected by form of referring expression. These predictions are based on Ariel's "hierarchy of Accessibility markers," which states that lower accessibility is typically signaled by a more specific form of referring expression (1991: 449). Related to this notion are the recipient's corollaries of Levinson's I- and M- principles. The I-principle states that interlocutors are to "prefer co-referential readings of reduced NP's [such as null expressions]" (cited in Blackwell 1998: 612), while the M-principle states that the use of more explicit (or 'marked') expressions should be interpreted as non-co-referential (cited in Blackwell 1998: 613). If marked expressions (in this case, the explicit plural pronoun *ellos*) are expected to indicate non-co-reference or low accessibility, then readers should have difficulty resolving those expressions to antecedents which are more obviously co-referential or highly accessible within the co-text (i.e., the first sentence of each passage). Likewise, if null expressions are expected to indicate co-reference or high accessibility, then readers should have difficulty resolving those expressions to antecedents which are less obviously co-referential or relatively inaccessible within the co-text.

In summary, it should be more acceptable to resolve explicit reference to antecedents which are relatively inaccessible (i.e., the PARA condition), and less acceptable to resolve explicit reference to antecedents which are relatively accessible (i.e., the NON-SPLIT condition). Likewise, it should be more acceptable to make null reference to antecedents which are relatively accessible, and less acceptable to make null reference to antecedents which are relatively inaccessible. Thus, if antecedent accessibility decreases from the NON-SPLIT condition to the PARA condition, then the processing difficulty for a null plural reference should increase proportionally from the NON-SPLIT condition to the PARA condition. Conversely, the processing difficulty for an explicit plural reference should decrease from the NON-SPLIT condition to the PARA condition.

If there is no detectable difference in accessibility among various types of plural antecedents, then there should be no significant difference in processing difficulty for plural references across conditions. If this is the case, then we are left to decide between the two remaining possible conclusions: We must either classify the types of plural antecedents analyzed in this experiment as equally accessible (which means that the sharing of syntactic and/or thematic roles does not affect the accessibility of plural antecedents), or modify our understanding of the factors which contribute to the accessibility of plural antecedents.

3.4   Results

While it was expected that plural reference in general would be more easily processed when the components of a plural antecedent shared syntactic and thematic roles, an analysis of variance demonstrated that the interaction was not significant ($P = .15$). Disregarding form of referring expression, sentences which contained plural reference to antecedents in the NON-

SPLIT and Y conditions were read only slightly faster (an average of 1771 ms and 1850 ms, respectively) than sentences which contained plural reference to antecedents in the CON and PARA conditions (an average of 1907 ms and 1914 ms, respectively), as evidenced by Figure 3.1.
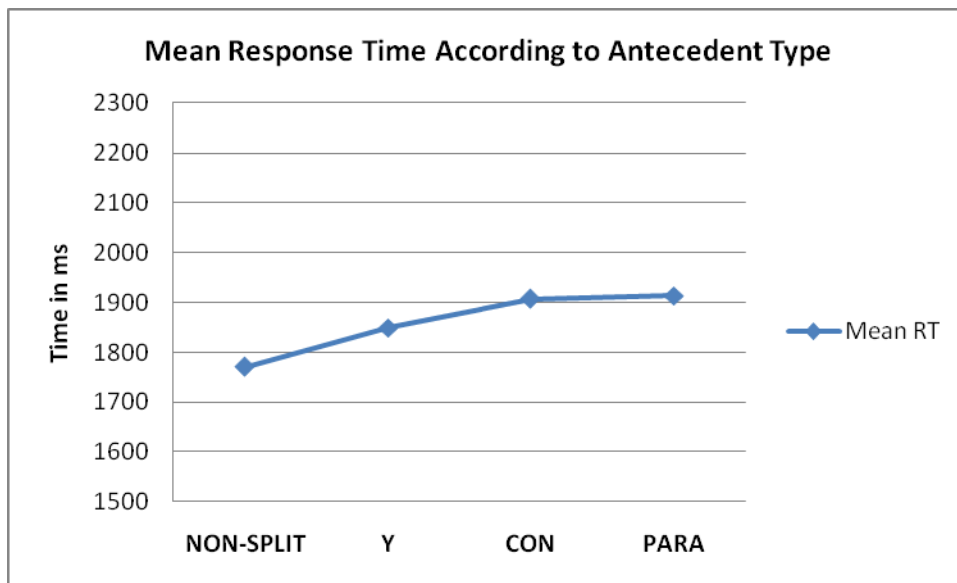

Figure 3.1. Mean Response Time According to Antecedent Type

Even when the results are parsed according to form of referring expression, the data do not reveal a significant interaction between antecedent type and processing difficulty ($P = .29$), as evidenced by Figure 3.2. From the NON-SPLIT condition to the PARA condition, the average reading time for sentences with null plural reference increases by less than 80 ms (from 1788 ms for NON-SPLIT antecedents to 1867 ms for antecedents in the PARA condiiton), indicating that antecedent type had little effect on the processing of null references. For sentences with explicit plural reference, the average reading time increases by roughly 200 ms (from 1753 ms to 1960 ms, respectively), inidicating that antecedent type had a slightly larger (though still insignificant) effect on the processing of explicit references.
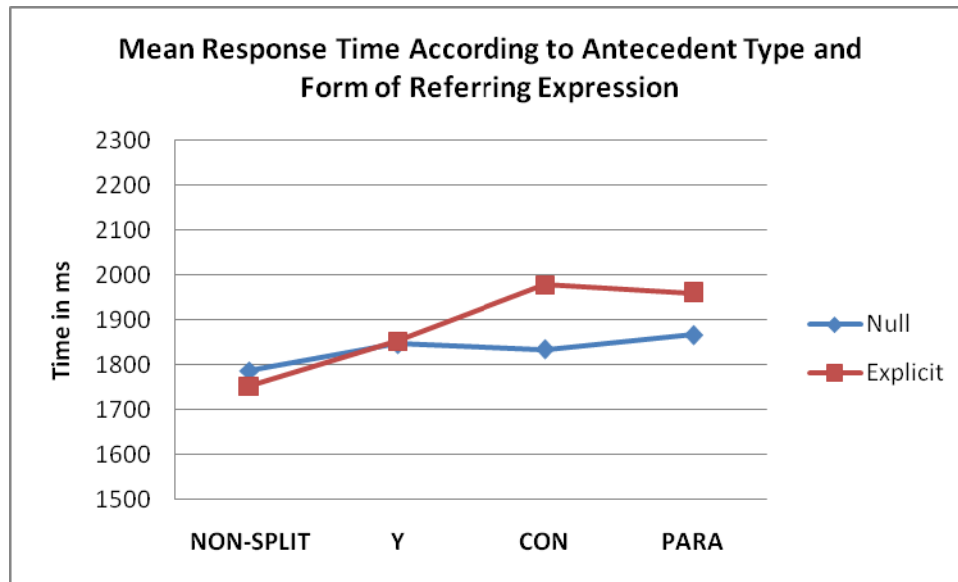
Figure 3.2. Mean Response Time According to Antecedent Type and Form of Referring Expression

Finally, if we assume that the antecedents presented in all 4 conditions are equally accessible (as the data suggest), then we may disregard the conditions and focus our analysis on the effect of form of referring expression. As a result, we see that the difference between the average processing time for null plural references (1834 ms) and explicit plural references (1886 ms) was not significant ($P = .29$), as demonstrated by Figure 3.3. Overall, participants did not process explicit plural reference much differently than null plural reference. Disregarding antecedent type, we may conclude from the data that form of referring expression (or at the choice between a null plural subject pronoun and an explicit plural subject pronoun) does not significantly affect the processing (or resolution) of plural anaphora.
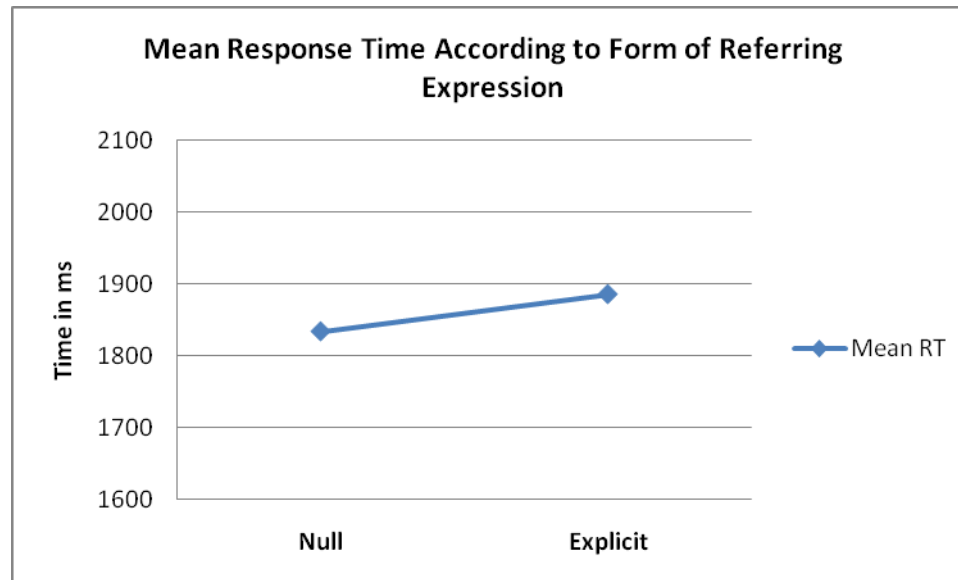
Figure 3.3. Mean Response Time According to Form of Referring Expression

Table 3.1, below, summarizes the results of reported in Figures 3.2 and 3.3 with the corresponding *P* values.

Table 3.3. Mean Response Times according to Antecedent Type and Form of Expression

	NP	Y	CON	PARA	TOTAL
Null	1788	1848	1835	1867	1834
Explicit	1754	1853	1979	1960	1887
	P = .71	*P* = .95	*P* = .18	*P* = .39	*P* = .29

3.5   Discussion and Conclusion

The results of the current processing experiment, while indicating that the resolution of plural reference in general is perhaps slighlty facilitated when the components of a plural antecedent share syntactic and thematic roles, do not indicate a significant interaction between plural antecedent type and the plural reference processing. Furthermore, as the data displayed in

Figures 3.2 and 3.3 suggest, the way in which a plural discourse referent is constructed (i.e., via introduction as a plural NP, a conjoined NP, or separate syntactic entities) does not seem to be reflected in the choice of form of plural referring expression (i.e., in the choice between a null subject pronoun and an explicit subject pronoun).

Here it is helpful to consider an obvious difference between the two experiments conducted for this work: the first experiment investigates the production of anaphora according to antecedent type, while the second investigates the processing of plural anaphora according to antecedent type. Nevertheless, we may still draw important parallels between the results of the two experiments. The production experiment demonstrates that the way in which a plural referent is represented syntactically and/or thematically in the discourse does not significantly affect the form of subsequent reference to that plural referent. The processing experiment demonstrates that the way in which a plural referent is represented in the discourse does not significantly affect the processing of subsequent reference to that referent, regardless of whether reference is made via null or explicit subject pronoun. In this respect, both sets of results lead us to a common conclusion: there is little correspondence between the extent to which the components of a plural referent share syntactic and/or thematic roles and the form of a corresponding plural referring expression (i.e., a null or an explicit subject pronoun) in Spanish. Consequently, according to Ariel's "hierarchy of Accessibility markers" (1991: 449), this should mean that there is likewise little correspondence between the extent to which the components of a plural referent share roles and the accessibility of that referent.

However, there are respects in which the results of the two experiments lead us to slightly different conclusions. Although the results of the production experiment did not demonstrate a significant interaction between the sharing of roles and the form of plural referring expression,

they did reveal a significant interaction between the sharing of roles and the frequency of plural reference in general. Specifically, while the sharing of roles did not increase the ratio of null to explicit plural references, it did significantly increase the ratio of plural references to singular references. If these results were to be paralleled by the processing experiment, we might expect for the processing difficulty of plural reference, disregarding the form of that reference, to increase as the sharing of roles decreased (i.e., from the NON-SPLIT to the PARA condition). While the data summarized in Figure 3.1 do seem to reflect this trend to a certain extent, we see that the interaction is not significant ($P = .15$).

This discrepancy is perhaps simply a reflection of the difference between the processeses of producing and resolving anaphora. While these processes do overlap a great deal, there is at least one key difference between them, exemplified by the different tasks required of participants in the two experiments. In the production experiment, partcipants had to decide whether plural reference was to be made at all, and then mark that reference with an appropriate Accessibility marker (i.e., a null or an explicit subject pronoun). In the processing experiment, the decision of whether or not to make plural reference was already made, given that all references to antecedents presented in the test sentences were plural. Participants were only required to resolve that plural reference to the appropriate referent (which, in every case, was the only plural referent available within the discourse segment).

In other words, in the production experiment, participants had to first form what Moxey et al. refer to as a "Complex Reference Object" (CRO) (2004: 346) from the information provided in the discourse before making plural reference. Here, it is important to note that the process of forming a CRO, as evidenced by the results of the experiment, is significantly affected

by the extent to which discourse referents share syntactic and/or thematic roles. I will come back to this point shortly.

In the processing experiment, plural references were already provided, signaling to the participants that a CRO had already been formed. The one remaining next step, then, was to resolve a given plural reference to the only available CRO within the discourse segment. All of the information required in order to successfully complete this task was provided by either the plural morphology of the verb (in the case of null references) or the plural morphology of the subject pronoun (in the case of explicit references). In other words, participants did not have to rely on syntactic or thematic information in order to successfully resolve plural reference.

In summary, then, we see that the formation of a CRO (or plural referent) in the production experiment depended on the extent to which discourse referents shared syntactic and thematic roles; the resolution of plural reference in the processing experiment did not. Why might this be the case? Returning to my discussion in the conclusion of Chapter 2 of Ariel's criteria for assessing Accessibility (1991: 445), I would like to suggest that the difference between the process of forming a CRO and the process of resolving plural anaphora (within the contexts of the respective experiments) depends largely on the presence or absence of competing referents. In the production experiment, since no reference had yet been made, participants were free to decide between making singular or plural reference to the discourse entities presented in a given test sentence. In this particular scenario, then, there was competition between singular discourse referents and plural discourse referents for subsequent anaphoric reference. As such, participants relied on the syntactic and thematic information provided about the discourse entities in order to "weed out the competition." As a result, related discourse entities (either syntactically or thematically) were more often referred to as a plural entity, while unrelated entities were more

often referred to as singular entities. As soon as a plural reference was chosen over a singular reference, however, competition for that reference disappeared because there was only one plural referent available within the discourse (i.e., the sum of *Juan+María*). In such cases, respondents did not consider the syntactic and thematic information provided about the discourse entities because there was no competition to "weed out." As a result, the form of plural referring expression did not vary significantly according to the extent to which discourse entities shared syntactic and thematic roles.

If we extend this logic to the results of the processing experiment, we understand why there was not a significant interaction between the sharing of roles and the processing of plural reference in general. Once plural reference was made (in the second sentence of each passage), there ceased to be competing referents for that reference (whether null or explicit). As such, participants had no need to rely on syntactic and thematic information in order to disambiguate between potential referents. As a result, there was not a significant interaction between the extent to which discourse referents shared roles and the processing of plural references in general.

In summary, syntactic and thematic factors affected the formation of CRO's because of the presence of competition (between singular referents and plural referents). This explains why antecedent type had a significant effect on the frequency of plural reference in the production experiment. However, syntactic and thematic factors did not affect the processing of plural references because of the absence of competing plural referents. This explains why processing difficulty for plural references (regardless of form) did not increase significantly from the NON-SPLIT condition to the PARA condition in the processing experiment.

This leads us to a more general conclusion about the processing of anaphora, suggested implicitly in frameworks such as Centering Theory (cf. Walker et al. 1998): syntactic and

thematic factors affect the production and/or resolution of anaphora only in the presence of competing referents. In other words, when interlocutors are forced to choose between two or more possible referents for a given reference, they use the syntactic (and perhaps thematic) roles of those referents in order to make the most appropriate selection. However, when there is no competition, interlocutors have no need to rely on syntactic and thematic cues to select an appropriate referent.

Finally, it is important to note that competition was eliminated in both experiments via the introduction of a plural reference. As soon as plural reference was made (whether produced by the participant, in the case of the production experiment, or provided in a test sentence, in the case of the processing experiment), the number of possible referents for that reference was reduced to exactly one. In turn, the necessity of relying on syntactic and thematic information in order to resolve that reference was eliminated. This suggests, perhaps not surprisingly, that within the hierarchy of constraints which govern anaphoric resolution, there is a general agreement constraint which comes into play prior to the effect of syntactic and/or thematic constraints (cf. Beaver 2004: 14). More generally, where number and/or gender agreement is not sufficient to successfully resolve a given anaphoric expression, interlocutors may then rely on other information – syntax, world-knowledge, and etc. – in order to resolve the expression.

This idea will be developed further in the following chapter.

CHAPTER 4

THE ACCESSIBILITY OF PLURAL ANAPHORA: AN APPLICATION

4.1   Introduction

In the production study discussed in Chapter 2, we saw that the frequency of plural

reference increased in direct proportion to the extent to which discourse entities shared syntactic

and thematic roles, confirming the results of the Moxey et al. (2004) production experiment

conducted in English. However, the ratio of null plural reference to explicit plural reference did

not increase according to the extent to which discourse entities shared roles. This was attributed

to the fact that any plural, anaphoric reference made within the context provided by the

experiment necessarily evoked the only – and consequently, the most syntactically salient –

plural referent available. Thus, all plural references evoked a highly accessible antecedent, and

therefore tended to be made via a form which reflected that high accessibility: a null subject

pronoun.

In the processing study discussed in Chapter 3, we saw that the processing difficulty for

plural references – whether null or explicit – did not decrease significantly according to the

extent to which discourse entities shared syntactic and thematic roles. This was attributed to the

fact that, regardless of the way in which plural referents were constructed from the discourse, the

plural markings of any subsequent reference required that that reference be resolved to the only

plural referent available in the context provided by the test passages.

Taken together, the results of the production and processing experiments suggest that in the absence of competition, there is no need to rely on syntactic structure[1] in order to resolve an anaphoric expression to its appropriate referent. In the experiments, all of the information needed to successfully resolve a particular plural reference was provided by the plural markings of that reference. In other words, the resolution of anaphora in both experiments was handled exclusively by a general agreement constraint, which requires that "anaphoric expressions agree with their antecedents in number and gender" (Beaver 2004: 14).

Thus, based on the results of the experiments, we must conclude, like Cameron (1997), that "[Ariel's] structural conception of antecedent inferiority, as distinguishes split antecedents from non-split antecedents, does not hold" (1997: 48). In other words, the accessibility of a particular plural referent does not depend on the structural representation of that referent (whether as a plural NP, a conjoined NP, or two singular and syntactically distinct NP's) within a given discourse segment. However, I argue that this conclusion needs to be qualified. It is not that the structural representation of a particular plural discourse referent (i.e., the extent to which discourse entities share syntactic and/or thematic roles) is irrelevant in the resolution of anaphora. Rather, as I will discuss in further detail in the following section, it is simply that structural representation aids in the resolution of anaphoric expression only when higher-ranking constraints have not specified the appropriate referent.

---

[1] From this point forward, when I refer to "syntactic structure" or "syntactic constraints" without further qualification, I refer specifically to the influence of the sharing of syntactic (and/or thematic) roles on subsequent anaphoric reference. I do not refer to government and binding constraints, which, in the process of anaphoric resolution, would rank higher than influences such as the sharing of syntactic (and/or thematic) roles.

4.2   Cameron 1997 vs. the Current Analysis

It is important to note that while the results of the experiments conducted for the current analysis lead us to roughly the same conclusion as Cameron's (1997) study, the way in which the experiments were conducted require us to qualify that conclusion. What I mean by this is that the carefully controlled context provided by the experiments does not necessarily allow us to make the general claim that the structural representation of a plural antecedent (or, the way in which a plural antecedent's components are presented syntactically within the discourse) has no effect on that antecedent's accessibility. Instead, we are required to make the more specific claim that the structural representation of a plural antecedent has no effect on that antecedent's accessibility *in the absence of competition for a given plural anaphoric reference*.

Given that the context from which Cameron's (1997) data were taken should have allowed for the influence of competition (i.e., the possibility that two or more plural referents competed for a given plural reference), we might question the validity of the more specific claim required by the current methodology. In other words, if the influence competition was present in the Cameron (1997) study, and if the results of the analysis demonstrated that the structural representation a given antecedent did not affect the accessibility of that antecedent, then it should be unnecessary for us to qualify our claim by saying that structural representation is not influential *in the absence of competition*.

There are two problems with this logic, however. First, split antecedents of any type are somewhat rare in natural discourse – Cameron found only 111 examples (out of 1060 antecedents analyzed) in the speech from 20 recorded interviews (1997: 48-49). Thus, it seems likely that split antecedent constructions which could provide competition for a given plural reference would be even rarer, especially when the context of analysis is limited to five clauses,

as it is in Cameron's study (1997: 39).[2] Consider the following excerpt from Cameron's data, which was provided as an example of a split antecedent.

(1) Diego: Él fue instructor.
               Se hizo instructor de equitación.
    Richard: ¿Ah sí?
    Cecilia: Sí, de niños porque **él** le andaba siempre en – mucho con **Diego**
               y **ellos** salieron mucho los dos juntos.
               (Cameron 1997: 37)

From the excerpt we see why the occurrence of split-antecedents might be rather rare in natural discourse. First, there must be at least two (human) discourse entities mentioned within five clauses of one another (if we adopt Cameron's envelope of analysis). Second, there must be plural reference made to these discourse entities, again within the same five clauses. In this case, the reference *ellos* refers to the only possible plural referent available within the discourse segment. Thus, based on Ariel's criteria for determining antecedent accessibility (1991: 445) – distance, salience, and competition – we also see why split antecedents in this type of context tended to yield null reference instead of explicit reference. Much like the context provided in the experiments conducted for the current study, the context of this excerpt presents a plural antecedent that is without competition, and therefore the most syntactically salient plural antecedent available. Furthermore, *ellos* is minimally separated from its antecedent, occurring in the very next clause.

In order for there to have been competition for *ellos* in this case, minimally, a third human referent would have had to be introduced somewhere in the four clauses preceding the

---

[2] Cameron analyzed "only those plural subjects in which antecedents [were] present within five preceding clauses" (1997: 39).

reference. While this would have increased the possible referents for *ellos* from just [*él+Diego*] to [*él+Diego*+3rd referent], [*él*+3rd referent], [*Diego*+3rd referent], and [*él+Diego*], given the requirements for such a scenario to exist, it seems that the influence of competition would have been extremely rare and difficult quantify. In summary, just because the context analyzed in Cameron's study theoretically allowed for the influence of competition does not mean that this influence was prevalent or even existent.

A second reason why we might not uphold Cameron's more general claim about the influence of structural representation on the accessibility of plural antecedents is that Cameron did not include the factor of competition as an independent variable in his analysis. Thus, even if competition were present for many of the references analyzed in the study, we are not aware of its effect on the form of those references (i.e., a null or an explicit subject pronoun).

Considering this, we must uphold our more specific claim that the structural representation of a plural antecedent has no effect on that antecedent's accessibility *in the absence of competition for a given plural anaphoric reference*. In the following section, I will discuss one of the implications of this claim: namely, that the structural representation of a plural antecedent could have an effect on that antecedent's accessibility in the presence of competition. The following discussion suggests a preliminary answer to the question posed by Cameron at the close of his study, of "whether inferior[3] antecedents actually exist for personal plural subjects in Spanish" (1997: 48).

---

[3] The term inferior refers to a sum of singular discourse entities which might be more suitable for plural reference than others (e.g., the sum of the constituents of a conjoined NP vs. the sum of a grammatical subject and an oblique argument). Specifically, Cameron's (1997) study investigates Ariel's suggestion that null reference to the sum of a grammatical subject and an oblique argument, for example, should be impossible (1991: 457).

4.3   Koh and Clifton 2001: Resolving Plural References in the Presence of Competition

As a transition between the conclusion taken from the results of the current experiments and a discussion of how an analysis of plural anaphora might fit within Centering Theory, it is helpful to consider a study conducted by Koh and Clifton (2001) which investigated some of the factors affecting the resolution of plural anaphora in contexts where multiple plural referents competed for a single plural reference.

In a questionnaire conducted for the study, participants were asked to provide continuations for two different types of test sentence, modeled below.

(2) Tom sang with Jim and Tony at the school. They _____

(3) Tom recognized Jim and Tony at the school. They _____
       (Koh and Clifton 2001: 20)

In example (2), *Tom* is related to *Jim* and *Tony* through what the authors refer to as "predicate symmetry" (2001: 20). In other words, all three animate entities presented in example (2) fulfill the same thematic role, namely that of co-agent. In (3), however, the predicate is "nonsymmetric" (2001: 20), meaning that *Tom* fulfills a different thematic role (agent) than *Jim* and *Tony* (recipients).

Notice that these two experimental conditions are somewhat parallel the CON and PARA conditions used for the current experiments.

(4) Juan limpió las casas con María.

(5) Juan limpió las casas para María.

In the CON condition modeled in (4), the predicate is symmetric; both *Juan* and *María* share the thematic role of co-agent. In the PARA condition modeled in (5), the predicate is nonsymmetric; *Juan* acts as the agent while *María* is the beneficiary. The main difference between the conditions tested by the current experiment and those tested by Koh and Clifton (2001) is that there is only one possible plural referent presented in the CON and PARA conditions. In the conditions modeled by (2) and (3), however, there are a total of four possible plural referents: [*Tom+Jim+Tony*], [*Jim+Tony*], [*Tom+Jim*] and [*Tom+Tony*].[4]

In addition to providing a continuation for the sentences modeled in (2) and (3), participants in the Koh and Clifton (2001) study also indicated the intended antecedent of the reference *they*. The results of the questionnaire indicated that for both conditions – the symmetric predicate and the nonsymmetric predicate – the intended antecedent for *they* was most often the plural referent formed by the sum of all three animate entities presented in the test sentence: [*Tom+Jim+Tony*]. Furthermore, "participants circled all three entities as the intended antecedent of *they* more frequently for [test sentences] with symmetric predicates [88%] than with nonsymmetric predicates [61%]" (Koh and Clifton 2001: 21).

These results lead us to two conclusions about the resolution of plural anaphora in cases where multiple plural referents compete for a given plural reference. First, it is clear that the structural representation of split antecedents affects the resolution or interpretation of subsequent plural reference. When discourse entities are related via predicate symmetry (i.e., when they share thematic role), they are more likely to be joined together in order to form a plural referent.

Second, it seems that the preferred referent for an ambiguous reference such as *they* consists of what Howe refers to as the "maximal subset" of the set of entities within a given

---

[4] It should be noted that these last two possibilities are highly unlikely given that it is relatively difficult to break apart members of a conjoined NP (cf. Albrecht and Clifton 1998; Gordon et al. 1999). This difficulty is referred to by Albrecht and Clifton (1998) as the Conjunction Cost.

utterance (2003: 17). In other words, if an utterance contains a list of entities *i...n*, then the preferred antecedent for a plural anaphor evoking any combination of those entities will be that which includes every entity presented in the utterance – i.e., the powerset {i,...n} (Howe 2003: 17). It should be noted that this set of entities is sensitive to the ruling out of entities which do not satisfy general agreement or animacy constraints (cf. Howe 2003; Beaver 2004). For example, the entity *the school* in examples (2) and (3) could not be combined with the other animate entities in the utterance in to serve as a viable antecedent for *they*. Thus, even in the absence of predicate symmetry (such as in (3)), Koh and Clifton explain the preference for all three entities as a function of "other contextually salient sources of equivalence among all three entities apart from the equivalence conferred by the symmetric predicate (e.g., they are all humans referred to by proper names)" (2001: 21). In other words, the grouping together of discourse entities to form plural referents is governed by various kinds of equivalency constraints (Koh and Clifton 2001: 10).[5] Equivalence between two or more discourse entities may arise as a function of syntactic structure (as in (2)), or as a function of belonging to what Koh and Clifton refer to as the same "ontological category" (e.g., [+human], [+animate], and etc.) (2001: 12).

Before moving on to a discussion of how the current analysis of plural anaphora in Spanish might fit within Centering Theory, there is another result from the Koh and Clifton (2001) study which merits discussion. I mentioned in the conclusion of section 4.2 that the results of Cameron's (1997) study led him to pose the question of "whether inferior antecedents actually exist for personal plural subjects in Spanish" (1997: 48). In the Koh and Clifton (2001)

---

[5] In addition to testing the influence of equivalence produced by predicate symmetry, Koh and Clifton also explicitly tested the influence of equivalence produced by "ontological homogeneity" (2001: 12), or the extent to which discourse entities occupy same existential category (e.g., human, animate, inanimate, and etc.). They found, perhaps not surprisingly, that discourse entities were more often joined together to form a plural referent when they were "ontologically homogeneous" (2001: 14).

experimental conditions, and especially in the case of the symmetric predicate condition, a superior (or preferable) antecedent is clearly created: [*Tom+Jim+Tony*]. Thus, the three remaining possible antecedents, [*Jim+Tony*], [*Tom+Jim*] and [*Tom+Tony*], are relatively inferior.

However, because in each case respondents were forced to make plural reference via the explicit pronominal expression *they*, we cannot know how reference to one of the inferior plural antecedents would have been reflected in the form of referring expression (such as in Spanish, where a plural subject pronoun may be either null or explicit). We only know that the form *they* – the form which signifies the highest level of accessibility for a plural referent in English – most often resolved to the maximal subset of discourse entities.

In summary, it appears (at least in English) that the presence of more than two discourse entities in a given discourse segment gives rise to competing plural referents which may be ranked in terms of relative accessibility. Accessibility, in turn, is determined in part by the structural representation of discourse entities within the utterance. Thus, if we assume that the same is true for Spanish, we may answer Cameron's question of "whether inferior antecedents actually exist" (1997: 48) affirmatively. However, what we do not know is whether the difference between superior and inferior plural antecedents corresponds with a difference between forms of referring expression, namely null and explicit plural subjects in Spanish.

4.4 Plural Anaphora within Centering Theory

One of the questions posed by Koh and Clifton is whether the "centering phenomena that seem to be so important in the comprehension of singular pronouns play a role in plural pronoun processing" (2001: 28). Here, I will consider this question specifically as it pertains to plural anaphora in Spanish.

As discussed in Chapter 1, Centering Theory bases its claims about discourse coherence on the existence of different types of transitions between utterances. These transition types, in turn, are based on whether the center of a given utterance is maintained from that utterance to a subsequent one. Crucially, the center of a given utterance is determined largely by syntactic role.[6] Taboada gives the following ranking for discourse entities in Spanish:

(6) Experiencer > Subject > Animate Indirect Object > Direct Object > Other >
      Impersonal/Arbitrary Pronouns (Taboada 2005: 14)

In Spanish, the Cp (or preferred center) of a given utterance in Spanish will be the experiencer; if no empathetic/animate entity is present, then the center defaults to the grammatical subject, and so on. If the Cp of an utterance is also a Cb (backward-looking center), and if that Cb is refers to the same entity as the Cp in the previous utterance, then the transition between those utterances is defined as CONTINUE (the most coherent transition type), as evidenced by (7).

(7) (a) **Juan$_i$** compró un regalo$_k$ para **Pedro$_j$**.
        Cf: [JUAN, PEDRO, REGALO]
        Cp: [JUAN]
        Cb: [?]
    (b$_1$)Ø$_{i/?j/*k}$ Es un buen amigo.
        Cf: [JUAN, DINERO]
        Cp: [JUAN]
        Cb: [JUAN]

Because there is no empathetic entity in (7a), the grammatical subject *Juan* serves as the Cp. *Pedro*, the only other human entity in the utterance, is ranked lower than *Juan* because it is an oblique argument. The null subject in (7b), because it is the Cb, gives rise to a CONTINUE

---

[6] In English, the center of a given utterance is determined entirely by syntactic role (Walker et al. 1998: 7).

transition only if it evokes the Cp of the previous utterance: [JUAN]. We see that the preference for *Juan* as the antecedent for the null subject in (7b) comes about as a result of the salience hierarchy (or Cf list) generated from the syntactic roles of the discourse entities introduced in (7a). In other words, relative to the other discourse entities presented in (7a) which could serve as an antecedent for Ø in (7b), *Juan* is the entity which contributes most to discourse coherence. As such, *Juan* is the preferred antecedent for Ø. Conversely, *Pedro* is not preferred because it is not the Cp of (7a). In other words, relative to the other discourse entities in (7a) which could serve as an antecedent for Ø in (7b), *Pedro* is an entity which would contribute less to discourse coherence.

Note that *Juan* and *Pedro* are not the only discourse entities introduced in (7a). A third entity, *regalo*, is also introduced. While an entity in this position might normally be ranked higher than *Pedro* given that Direct Object > Other (Taboada 2005), *regalo* is in this case an unacceptable option because the null subject pronoun in (7b) strongly prefers at least an animate (if not a human) referent. In other words, the option of *regalo* as an antecedent for Ø in (7b) is eliminated, not because of its syntactic role (or, its position within the structural representation of the entities in (7a)), but because of a general semantic constraint which restricts *un buen amigo* to animate referents.

Applying this same logic to an analysis of plural anaphora within Centering introduces several problems, as evidenced by examples (8) and (9), below.

(8) (a) Los chicos$_i$ fueron a la tienda$_j$.
Cf: [CHICOS, TIENDA]
Cp: [CHICOS]
Cb: [?]

(b) Ø$_i$ Compraron una botella$_k$ de vino.
Cf: [CHICOS, BOTELLA]
Cp: [CHICOS]
Cb: [CHICOS]

(9) (a) Juan$_i$ fue a la tienda$_k$ con María$_j$.
Cf: [JUAN, MARÍA, TIENDA]
Cp: [JUAN]
Cb: [?]

(b) Ø$_{i+j}$ Compraron una botella$_l$ de vino.
Cf: [JUAN+MARÍA, BOTELLA]
Cp: [JUAN+MARÍA]
Cb: [JUAN+MARÍA]

As expected, the preferred antecedent for the null subject pronoun in (8b) is *los chicos* because it is the most syntactically salient entity (i.e., the Cp) in (8a). Thus, when the null subject in (8b) resolves to *los chicos,* the transition from (8a) to (8b) is defined as CONTINUE. Example (9), however, is more problematic. First, based on a Cf list generated solely from the syntactic role of the entities presented in (9a) – [JUAN, MARÍA, TIENDA] – we might expect a subsequent null reference to evoke the referent [JUAN]. Second, given that the entities in the Cf list of (9a) are all singular, we might expect subsequent anaphoric reference to be singular. However, not only does the anaphoric reference in (9b) not evoke [JUAN], but it does not evoke a singular referent at all. In fact, it is clear that the anaphor in (9b) does not resolve to any of the referents of the Cf list for (9a).

The first problem to be addressed is how anaphoric reference can be made to entities which are never entered into the Cf list for a given utterance – i.e., split referents. As an answer to this question, Howe proposes "updating the process of building a Cf list" to include both individuals and sums of individuals which could serve as split referents (2003: 10). After such an

update, the Cf list for (9a) includes both individual entities and the sums of those individuals: <i,j, i+j, k, i+k, j+k, i+j+k>. The second problem to be addressed is how these entities are ranked within the Cf list. Here, it is important to make a distinction between the analysis developed for the current study and the goals of Centering Theory.

That the null subject in (9b) resolves to *Juan+María* is easily explained in terms of the analysis that has been developed for plural anaphora within Accessibility Theory. Based on general agreement and semantic constraints, the null reference in (9b) must resolve to a plural, human referent. As such, we are left with only one option for the antecedent for that null reference: *Juan+María*. In terms of Accessibility Theory, since there is only one antecedent available for plural reference, this antecedent is also the most syntactically salient antecedent available. Furthermore, since *Juan+María* is also minimally separated from corresponding anaphoric reference, it is highly accessible, licensing the use of the null anaphor. Thus, we see that general agreement constraints, coupled with Ariel's Accessibility criteria (1991: 445), justify the resolution of the plural null subject pronoun in (9b) to *Juan+María*, whereas a Cf ranking based solely on the syntactic role of the entities presented in (9a) does not.

However, the goal of Centering Theory is not to provide a model for the resolution of anaphora in discourse, but rather to explain how speakers' choices related to anaphoric production affect discourse coherence. In other words, Centering is not necessarily concerned with correctly resolving a particular reference to its referent, but rather with how reference (as a Cb) to one referent instead of another affects discourse coherence. For example, consider a dual application of both Centering Theory and the analysis of plural anaphora developed for the current study in (10).

(10) (a)   Juan$_i$ fue a la tienda$_k$ con María$_j$.
                        Cf: <i,j, i+j, k, i+k, j+k, i+j+k>
       (b)   Ø$_i$ Compró una botella de vino.
       (b')   Ø$_{i+j}$ Compraron una botella de vino.
       (b'')   Ø$_k$ Estaba cerrada.

If anaphoric reference to the entities in (10a) is singular, as in (10b) and (10b'), then the Cf list for (10a) is reduced to singular entities which may be ranked in the order <i,j,k>. If subsequent anaphoric reference is plural, as in (10b'), then the Cf list for (10a) is reduced to the sums of those singular entities. Applying a semantic constraint which prevents *Juan* or *María* from being joined with *tienda* to form a plural referent, we are left with only one possible referent for the anaphor: *Juan+María*. Again, once we are left without competing referents for *Juan+María,* it is the most highly accessible referent available, and null form of reference is therefore justified. However, as I will explain, a justification of the classification of *Juan+María* as highly accessible does not necessarily mean that the transition from (10a) to (10b') can be defined as CONTINUE. In more general terms, just because a particular referent is the most highly accessible referent for a given reference does not mean that subsequent reference to that entity by a Cb results in optimal discourse coherence.

As an example, we see in (10b'') that semantic and/or agreement constraints would limit the possible referents for the null anaphor to *tienda*. Nevertheless, within Centering framework, reference to *tienda* in the subject position would still result in less than optimal discourse coherence, because Cb(10a) (i.e., *Juan*) ≠ Cb(10b'') (i.e., *tienda*). Transition from (10a) to (10b''), even though the anaphoric reference in (10b'') is completely unambiguous, is defined as a SMOOTH-SHIFT.

So it seems that while the analysis of plural anaphora developed in this study helps to explain why split antecedents may be highly accessible (depending on the presence or absence of

competition) this finding does not necessarily help us to rank split antecedents within the

Centering framework. As such, several questions remain: First, how should split antecedents be

ranked within a Cf list? Second, how can this ranking be applied in order to make predictions

about discourse coherence? And third, how might the Cf ranking for split antecedents affect the

form of referring expression when those antecedents are referenced in discourse? While

providing definitive answers to these questions is beyond the scope of the present work, I will

offer some preliminary predictions before closing. Consider the following examples:

(11)   Juan$_i$ comió con Pedro$_j$ y Marcos$_k$ en el restaurante$_l$.
Cf: <i, j, k, l>
Cp: [JUAN]
Cb: [?]
Cf$_{CRO}$: {i+j, i+k, i+l, j+k, j+l, k+l, i+j+k, i+j+l, j+k+l, i+k+l, i+j+k+l}
Cf$_{CRO}$+Filter: {i+j, i+k, j+k, i+j+k}

(12)   Juan$_i$ vió a Pedro$_j$ y Marcos$_k$ en el restaurante$_l$.
Cf: <i, j, k, l>
Cp: [JUAN]
Cb: [?]
Cf$_{CRO}$: {i+j, i+k, i+l, j+k, j+l, k+l, i+j+k, i+j+l, j+k+l, i+k+l, i+j+k+l}
Cf$_{CRO}$+Filter: {i+j, i+k, j+k, i+j+k}

For examples (11-13), I have provided one Cf list based only on the singular entities

presented in the example (Cf), and another based on the Complex Reference Objects (CRO's)

(i.e., plural referents) formed from the sum of those entities (Cf$_{CRO}$). A choice between Cf or

Cf$_{CRO}$ would depend on the number markings of a subsequent anaphoric reference. Applying a

general semantic equivalence constraint to Cf$_{CRO}$ would also exclude combinations which

included an inanimate referent, such as {i+l}. For now we will focus on the Cf list generated by a

plural reference: Cf$_{CRO}$+Filter (henceforward, the Cf$_{CRO}$).

Based on the results of Koh and Clifton's (2001) experiment, we might rank the maximal subset (for animate entities) <i+j+k> as the most salient member of both (11) and (12), regardless of the fact that the entities in (11) are related via predicate symmetry while the entities in (12) are not. Also, based on observations and experimental data which suggest the difficulty of separating the members of a conjoined NP into single constituents (i.e., the Conjunction Cost) (Albrecht and Clifton 1998; Gordon et al. 1999), we might rank referents which split a conjoined NP at the bottom of the Cf list. Thus, for (11) and (12), our Cf_{CRO} would be ranked as follows: <i+j+k, j+k, i+j, i+k>. This ranking is supported by the results of Koh and Clifton (2001), which suggest that after analogous constructions in English, the preferred interpretation the subject pronoun *they* is the maximal subset. The second most preferred interpretation, according to the data, is the plural referent produced by the NP: <j+k>. In terms of discourse coherence, then, plural reference to the maximal subset in (11) and (12) (repeated here as (13) and (14)), such as in (13b) and (14b), should give rise to a CONTINUE transition.

(13) (a) $Juan_i$ comió con $Pedro_j$ y $Marcos_k$ en el $restaurante_l$.
         Cf_{CRO}: {i+j+k, j+k, i+j, i+k}
         Cp_{CRO}: [JUAN+PEDRO+MARCOS]
         Cb:[?]

     (b) $Ø_{i+j+k}$ Conversaron mucho.
         Cf_{CRO}+Filter: {i+j+k}
         Cp_{CRO}: [JUAN+PEDRO+MARCOS]
         Cb: [JUAN+PEDRO+MARCOS]

(14) (a) $Juan_i$ vió a $Pedro_j$ y $Marcos_k$ en el $restaurante_l$.
         Cf_{CRO}: {i+j+k, j+k, i+j, i+k}
         Cp_{CRO}: [JUAN+PEDRO+MARCOS]
         Cb:[?]

     (b) $Ø_{i+j+k}$ Conversaron mucho.
         Cf_{CRO}: {i+j+k}
         Cp_{CRO}: [JUAN+PEDRO+MARCOS]
         Cb: [JUAN+PEDRO+MARCOS]

We must be careful about positing a general rule which states that the maximal subset is always preferred in a Cf$_{CRO}$ list. In both (11) and (12), plural reference which includes the most syntactically salient individual entity (i.e., the grammatical subject) must also include the entities introduced but the conjoined NP. Again, this is because of the Conjunction Cost associated with splitting the members of that NP (Albrecht and Clifton 1998). While making reference solely to <j+k> would not split the constituents of the conjoined NP, it would exclude the most salient individual entity from the reference. It seems strange to assume that the most salient split reference would exclude the most salient individual entity of the utterance if there were no other constraints which prevented its inclusion. In short, the structure of Koh and Clifton's test sentences may have contributed to the preference for the maximal subset for a given plural reference.

Preference for the maximal subset might not exist for situations like those modeled in (15) and (16).

(15)   Juan$_i$ fue con Pedro$_j$ a la casa$_k$ de Marcos$_l$.
        Cf: <i, j, k, l>
        Cp: [JUAN]
        Cb: [?]
        Cf$_{CRO}$: {i+j, i+k, j+k, i+j+k}

(16)   Juan$_i$ y Pedro$_j$ fueron a visitar a Marcos$_k$ y Miguel$_l$.
        Cf: <i, j, k, l>
        Cp: [JUAN, PEDRO]
        Cb: [?]
        Cf$_{CRO}$: {i+j, i+k, i+l, j+k, j+l, k+l, i+j+k, i+j+l, j+k+l, i+k+l, i+j+k+l}

Here we see that the inclusion of another entity outside of a conjoined NP in the oblique position allows for a plural reference to include the grammatical subject without having to reference the maximal subset. In these cases, the grammatical subject is accompanied by another

individual entity which is equivalent either thematically (as in (15)) or syntactically and

thematically (as in (16)). Based simply on a relative salience hierarchy, we might suppose that

the overall salience of a plural entity is a function of the extent to which the more salient entities

are included less salient entities are excluded. As such, in the case of (15) and (16) (repeated

below as (17) and (18)), the plural referent composed of the grammatical subject and the entities

most equivalent to it (syntactically or thematically) would serve as the Cp_{CRO}: $<i+j>$. If this is

true, then we must also suppose that plural reference to $<i+j>$ would be more coherent than, for

example, reference to $<i+j+l>$. Reference to $<i+j>$, then, would result in a continue transition.

But is this the default interpretation of the null reference in (17b) and (18b), below?

(17) (a) $Juan_i$ fue con $Pedro_j$ a la $casa_k$ de $Marcos_l$.
            Cf_{CRO}: {i+j, i+j+l, i+l, j+l}
            Cp_{CRO}: [JUAN+PEDRO]
            Cb: [?]
    (b) $?Ø_{i+j}$Vieron una película.
    (b') $Ø_{i+j}$ No lo_l encontraron.

(18) (a) $Juan_i$ y $Pedro_j$ fueron a visitar a $Marcos_k$ y $Miguel_l$.
            Cf_{CRO}: {i+j, i+j+k+l, k+l, etc.}
            Cp_{CRO}: [JUAN+PEDRO]
            Cb: [?]
    (b) $?Ø_{i+j}$Vieron una película.
    (b') $Ø_{i+j}$ No los_{k+l} encontraron.

Pragmatically speaking, it would be strange to imagine (without additional context) that

Juan and Pedro watched a movie at Marco's house without Marco, or that Juan and Pedro visited

Marcos and Miguel, but watched a movie without them (modeled in (17b) and (18b),

respectively). In other words, pragmatic constraints would force the plural reference in (17b) and

(18b) to include the maximal subset. If we slightly alter the context, however, such as in (17b')

and (18b'), we find that the maximal subset is excluded. In summary, considering their

respective contexts, we might assume that the null references in (17b) and (18b) resolve to

<i+j+l> and <i+j+k+l>, while the null references in (17b') and (18b') resolve to <i+j>. So which

continuation – (b) or (b') – is more coherent?

A more general question which stems from this is whether the members of a Cf list which

includes split referents (i.e., our Cf_{CRO}) should be ranked before or after considering the effect of

semantic and or pragmatic constraints on subsequent reference to members of that list. If we rank

the members which exist on the list before considering the influence of semantic or pragmatic

filters, we would have to exclude split referents from the list. This is because it is the number

markings of a plural reference (which reflect the semantic idea of plurality) which cause us

include split referents to begin with. For example, without adjusting the Cf list for (9a),

introducing split referents according because of our agreement filter, we are unable to resolve the

plural reference in (9b) to any entity in (9a).

If we rank members after considering the influence of semantic and pragmatic filters,

then, given an appropriate context, any Cf list might be reduced to exactly one member. This

seems to be the case in both (10b'') and (18b'), where the only possible antecedents for the null

expressions are *la tienda* and *Juan y Pedro*, respectively. The problem with such a method is that

it prevents us – both in the case of singular and split reference – from making general predictions

about discourse coherence resulting from reference to one entity or another. For example, if we

apply semantic and pragmatic filters to the null reference in (10b''), reducing the Cf list to only

one plausible member (i.e., *la tienda*), then subsequent reference to *la tienda* in the subject

position would result in a CONTINUE transition. By the same logic, the null reference in (10b),

although it evokes a completely different entity, would also be defined as CONTINUE.

However, the original rules and constraints of Centering define (10b) as a CONTINUE (because

Cb(10a) = [?] and Cb(10b) = Cp(10b)) and (10b'') as a SMOOTH-SHIFT (because Cb(10a) =

[?] and Cb(10b) ≠ Cp(10b)). Thus, the question of how to rank split referents within a Cf list

remains unresolved.

One final issue which merits consideration before closing is the relationship between split

referents and form of referring expression in Spanish as it pertains to Centering Theory. Consider

the following example, for which I have constructed a tentative Cf_{CRO} list for explanatory

purposes.

(18) (a) $Juan_i$ fue con $Pedro_j$ a la $casa_k$ de $Marcos_l$.
            Cf_{CRO}: {i+j+l, i+j, i+l, j+l}
            Cp_{CRO}: [JUAN+PEDRO]
            Cb: [?]
     (b) $Ø_{i+j+k}$Vieron una película.
     (b')   ? $Ellos_{i+j}$ vieron una película.

Based jointly on Ariel's treatment of form of referring expression within Accessibility

Theory and Levinson's I- and M-implicatures, we may suppose that null reference would

typically be assigned to the most salient or "default" referent (determined according to Ariel's

three critieria (1991: 445)) within a particular discourse segment. Conversely, according to the

recipient's corollary of Levinson's M-implicature (cited in Blackwell 1998: 613), we may

suppose that the use of a more marked expression (e.g., an explicit subject pronoun) would

typically signal some other referent – necessarily, one that is less salient. In other words, if a

speaker uses an explicit pronoun, he means something different than he would have had he used

a null pronoun (adapted from Blackwell 1998: 613).

Given the context, we might assume that <i+j+l> is the most appropriate referent for the null reference (18b). Again, it would be strange to assume without additional information that Juan and Pedro went to Marco's house to watch a movie without Marcos. So if <i+j+l> is the default interpretation for the null reference in (18b), then how do we resolve the more marked expression in (18b')? According to the recipient's corollary of Levinson's M-implicature, *ellos* in (18b') would have to be assigned to a different referent than Ø in (18b). It might be logical to assume that the explicit pronoun, because it signals a level of accessibility one unit lower than that signalled by a null pronoun (according to Ariel's hierarchy of Accessibility marker's (1991: 449)), would evoke the second most accessible referent within the discourse segment. Using the tentative ranking created for (18a), this would be <i+j>. However, this intuition has yet to be investigated.

REFERENCES

Albrecht, Jason E. and Charles Clifton Jr. 1998. Accessing Singular Antecedents in Conjoined Phrases. Memory & Cognition 26.599-610.

Ariel, Mira. 1991. The Function of Accessibility in a Theory of Grammar. Journal of Pragmatics 16.443-463.

Barker, Chris. 1992. Group Terms in English: Representing Groups as Atoms. Journal of Semantics 9.69-93.

Beaver, David I. 2004. The Optimization of Discourse Anaphora. Linguistics and Philosophy 27.3-56.

Blackwell, Sarah E. 1998. Constraints on Spanish NP Anaphora: The Syntactic Versus the Pragmatic Domain. Hispania 81.606-618.

Blackwell, Sarah E. 2000. Anaphora Interpretations in Spanish Utterances and the Neo-Gricean Pragmatic Theory. Journal of Pragmatics 32.389-424.

Blackwell, Sarah E. 2001. Testing the Neo-Gricean Pragmatic Theory of Anaphora: The Influence of Consistency Constraints on Interpretations of Coreference in Spanish. Journal of Pragmatics 33.901-941.

Brennan, Susan E., Marilyn Walker Friedman, and Carl J. Pollard. 1987. A Centering Approach to Pronouns. Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics, Stanford.

Bolinger, Dwight. 1979. Pronouns in Discourse. Syntax and Semantics: Vol. 12: Discourse and Syntax, ed. by T. Givón, 289-310. New York: Academic Press.

Cameron, Richard. 1997. Accessibility Theory in a Variable Syntax of Spanish. Journal of Pragmatics 28.29-67.

Cote, Sharon. 1998. Ranking Forward-Looking Centers. Centering Theory in Discourse, ed. by Marilyn A. Walker, Aravind K. Joshi and Ellen F. Prince, 55-69. Oxford: Clarendon Press.

Culicover, Peter W. 1997. Principles and Parameters. New York: Oxford University Press.

Dahl, Osten. 2000. Egophoricity in Discourse and Syntax. Functions of Language 7.37-77.

Di Eugenio, Barbara. 1998. Centering in Italian. Centering Theory in Discourse, ed. by Marilyn A. Walker, Aravind K. Joshi and Ellen F. Prince, 55-69. Oxford: Clarendon Press.

Eschenbach, Carola, Christopher Habel, Michael Herweg and Klaus Rehkamper. 1989. Remarks on Plural Anaphora. Proceedings of 4th Conference of European Chapter of the Associations for Computational Linguistics, Manchester.

Gordon, Peter C., Barbara J. Grosz and Laura A. Gilliom. 1993. Pronouns, Names, and the Centering of Attention in Discourse. Cognitive Science 17.311-347.

Gordon, Peter C. and Randall Hendrick. 1998. Intuitive Knowledge of Linguistic Co-reference. Cognition 62.325-370.

Gordon, Peter C., Randall Hendrick, Kerry Ledoux and Chin Lung Yang. 1999. Processing of Reference and the Structure of Language: An Analysis of Complex Noun Phrases. Language and Cognitive Processes 14.353-379.

Grosz, Barbara J., Aravind K. Joshi and Scott Weinstein. 1995. Centering: A Framework for Modeling the Local Coherence of Discourse. Computiational Linguistics 21.203-225.

Gundel, Jeanette K., Nancy Hedberg and Ron Zacharski. 1993. Cognitive Status and the Form of Referring Expressions in Discourse. Language 69.274-307.

Kamp, Hans and Uwe Reyle. 1993. From Discourse to Logic. Dordrecht: Kluwer.

Kennison, Shelia M. and Peter C. Gordon. 1997. Comprehending Referential Expressions during Reading: Evidence from Eye-Tracking. Discourse Processes 24.229-252.

Koh, Sungryong and Charles Clifton Jr. 2002. Resolution of the Antecedent of a Plural Pronoun: Ontological Categories and Predicate Symmetry. Journal of Memory and Language 46.830-844.

Koh, Sungryong, Anthony J. Sanford, Charles Clifton Jr. and Eugene J. Dawydiak. 2005. Good-enough Representation in Plural and Singular Pronominal Reference: Modulating the Conjunction Cost. Interdisciplinary perspectives on reference processing, ed. by Jeanette K. Gundel and Nancy Hedberg. Oxford: OxfordUniversity Press.

Morales, Amparo. 1997. La hipótesis funcional y la aparición de sujeto no nominal: el español de Puerto Rico. Hispania 80.153-165.

Moxey, Linda M., Anthony J. Sanford, Patrick Surt and Lorna I. Morrow. 2004. Constraints on the Formation of Plural Reference Objects: The Influence of Role, Conjunction, and Type Description. Journal of Memory and Language 51.346-364.

Ranson, Diana. 1991. Person Marking in the Wake of /s/ Deletion in Andalusian Spanish. Language Variation and Change 3.133-152.

Reinhart, Tanya. 1981. Pragmatics and Linguistics: An Analysis of Sentence Topics. Philosophica 27.59-94.

Sanford, Anthony and Simon Garrod. 1981. Understanding Written language: Explorations in Comprehension beyond the Sentence. Chichester, England: John Wiley & Sons.

Sanford, Anthony J., Patrick Sturt, Linda Moxey, Lorna Morrow and Cathy Emmott. 2004. Production and Comprehension Measures in Assessing Plural Object Formation. The On-Line Study of Sentence Comprehension, ed. by Manuel Carreiras and Charles Clifton Jr., 150-166. New York: Psychology Press.

Silva, Vera Lúcia Paredes. 1993. Subject Omission and Functional Compensation: Evidence from Written Brazilian Portuguese. Language Variation and Change 5.35-49

Taboada, Maite. 2002. Centering and Pronominal Reference: In Dialogue, In Spanish. Paper presented at Proceedings of the 6th Workshop on the Semantics and Pragmatics of Dialogue (EDILOG), Edinburgh.

Taboada, Maite. 2005. Reference, Centers and Transitions in Spoken Spanish. Simon Fraser University manuscript.

Walker, Marilyn, Aravind Joshi and Ellen Prince. 1998. Centering in Naturally Occurring Discourse: An Overview. Centering Theory in Discourse, ed. by Marilyn A. Walker, Aravind K. Joshi and Ellen F. Prince, 1-28. Oxford: Clarendon Press.

Zabala, Loreley Hadic and Maite Taboada. 2003. Centering Theory in Spanish: Coding Manual. Simon Fraser University manuscript.