

8-31-2011

NPS IN CONVERSATION AND NARRATIVES: THE EFFECTS OF GENRE, INFORMATION FLOW AND INTERACTION

Jenny Dumont

Follow this and additional works at: https://digitalrepository.unm.edu/span_etds

Recommended Citation

Dumont, Jenny. "NPS IN CONVERSATION AND NARRATIVES: THE EFFECTS OF GENRE, INFORMATION FLOW AND INTERACTION." (2011). https://digitalrepository.unm.edu/span_etds/14

This Dissertation is brought to you for free and open access by the Electronic Theses and Dissertations at UNM Digital Repository. It has been accepted for inclusion in Spanish and Portuguese ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.

Jenny Dumont

(Signature)

Spanish and Portuguese

(Signature)

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

(Signature)

Chairperson

(Signature)

(Signature)

(Signature)

**NPS IN CONVERSATION AND NARRATIVES:
THE EFFECTS OF GENRE, INFORMATION FLOW
AND INTERACTION**

by

JENNY DUMONT

B.A., Spanish and Anthropology, University of Cincinnati, 2001
M.A., Spanish, University of New Mexico, 2004

DISSERTATION

**Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy
Spanish and Portuguese**

**The University of New Mexico
Albuquerque, New Mexico**

July 2011

I dedicate this work to my son, Alex

Acknowledgments

I consider myself very fortunate to have been able to work with so many accomplished scholars throughout my graduate studies at the University of New Mexico. I am extremely grateful to Catherine Travis who has been an inspirational teacher, mentor, advisor, colleague and friend to me for the last nine years. I must also express my gratitude to Rena Torres Cacoullas, who has helped me in too many ways to count. I thank her for her kindness, her wisdom, her encouragement, and her generous spirit. I am also grateful to Bill Croft and Sandy Thompson for their time, for their enlightening comments and for their extremely helpful conversations. I do not think that I would be the scholar that I am today were it not for the many other wonderful professors whose paths momentarily crossed mine at UNM, especially Melissa Axelrod, Alejandra Balestra and Jill Morford.

My dissertation research would not have been possible without the help of many Ecuadorians, and I thank them for all that they have shared with me. My life is much fuller due to one extended Ecuadorian family in particular. Pilar Bastidas welcomed me into her home in 2006, enabled me to collect the data for this dissertation, and taught me many things that cannot be learned from reading books. Her daughter, Andreíta Rosero, was a sister to me. I must also recognize José Luis and Paola Santana and their beautiful daughter Christa Amalia, my Ecuadorian family in Albuquerque, for their tremendous support, generosity, and love over the past five years. *Estoy muy afortunada tenerles en mi vida; que Dios les pague.* I am also indebted to my three Ecuadorian research assistants, Diego Bastidas, Fabricio Rosero and his wife Anabel Lanzáduri and Rossy Santana, and their families and friends who helped with the recordings.

Many friends and family members also gave me the support needed to get through life while completing my dissertation. Highlights of the past five years include many memorable moments spent with Sonia Balasch, Damián Vergara Wilson, Agripino Silveira, Pedro Torres, Carson Wassermann, Karina Peña, Kiley Guyton, Erin Amason Montero, Jaspall Singh and Vanessa de Veritch. My friendship with Mary Copple, which has blossomed over time and distance, has proven invaluable to me. My family in Ohio and my son Alex have grounded me in reality and have helped me enjoy life outside of academia.

Lastly, Kate Merrill, Martha Hurd and Rosario Johnson have been an indispensable part of my journey. I truly believe that without these three ingenious women I would not have accomplished all that I have at UNM.

Abstract

This dissertation investigates the grammar of noun phrases as produced in two genres (spontaneous conversations and Pear Film narratives) of spoken Ecuadorian Spanish. As speakers converse and narrate, they make choices about what form (full NP, pronoun, unexpressed) to use for a noun in a particular context, how to best modify that noun (choice of article and other determiners) and what syntactic role to use; this is a study of the different factors that motivate those decisions.

The patterns of use described here show that noun phrase grammar is shaped by the need to track referents and also by a variety of other dialogic interactional needs. On the one hand, speakers must manage the flow of information in a way that communicates what information is new or old, whether or not referents are presumed to be shared or not, and which NPs represent entities (referential NPs) and which perform other discourse functions (non-referential NPs). The grammar of a NP shows the routinized coding of these factors. On the other hand, during the course of a conversation, speakers are also navigating the interactional dimensions of communication, such as turn-taking, expressing stance and attitude, agreeing and disagreeing. The grammar of an NP is also influenced by these interactional factors. It is through the comparison of two genres that a fuller understanding of these different and sometimes competing motivations and their grammatical consequences is achieved.

Table of contents

List of Tables.....	ix
1 Introduction	1
1.1 Discourse Analysis.....	1
1.2 The present study	4
2 Research Questions.....	14
2.1 Research questions.....	14
2.2 Grammar and competing motivations.....	14
2.3 Information flow	17
2.3.1 Activation status.....	17
2.3.2 Discourse Referentiality.....	24
2.3.3 Identifiability.....	28
2.3.4 Specificity	32
2.4 Genre (information flow pressures)	33
2.5 Summary of research questions pertinent to information flow.....	36
2.6 Other factors influencing NP form	37
2.7 Summary	40
3 Data.....	41
3.1 Data collection	41
3.2 Pear Film Narratives	42
3.3 Spontaneous conversations	43
3.4 Transcription method	47
4 Coding	49
4.1 Syntactic and semantic factors.....	53
4.1.1 NP realization.....	53
4.1.2 Modification.....	54
4.1.3 Syntactic role	55
4.1.4 Semantic class of verb	58
4.1.5 Animacy and grammatical number.....	59
4.1.6 Summary	60
4.2 Factors related to information flow.....	60
4.2.1 Activation State.....	60
4.2.2 Specificity	61
4.2.3 Distance from previous mention.....	62
4.2.4 Realization of previous mention of the referent	63
4.2.5 Discourse referentiality.....	64
4.2.6 Summary of factors related to information flow.....	66
4.3 Additional factors.....	67
4.3.1 Increments.....	67
4.3.2 Repetition.....	68
4.3.3 Co-construction.....	69
4.3.4 Lists.....	70
4.4 Summary	71
5 Discourse referentiality.....	72
5.1 Factors related to discourse referentiality	72

5.1.1	Persistence.....	72
5.1.2	NP Realization	75
5.1.3	Modification.....	77
5.1.4	Specificity	79
5.1.5	Grammatical role	81
5.2	Summary of grammatical manifestations of discourse referentiality	86
5.3	Recoding referentiality.....	87
6	Discourse Referential NP Results.....	89
6.1	Overview of NPs.....	89
6.2	Formal characteristics of discourse referential NPs.....	95
6.2.1	Discourse referential NPs and realization.....	95
6.2.2	Discourse referential NPs and modification	105
6.2.3	Discourse referential NPs and syntactic role	119
6.2.4	Summary of new NPs	122
6.2.5	Summary of given referents.....	123
6.3	Additional considerations for referential NPs.....	123
6.3.1	New referents and semantic class of verb.....	124
6.3.2	Other types of 'given' referents	127
6.4	Preferred Argument Structure.....	133
6.5	Summary.....	137
6.5.1	Given vs. new referents.....	137
6.5.2	Genre-specific differences	139
7	Discourse Non-referential NPs.....	143
7.1	Discourse non-referential NPs in core roles	144
7.2	Objects of prepositions	150
7.3	Predicate nominals	154
7.4	Adverbials.....	155
7.5	Vocatives.....	156
7.6	Summary of non-referential NPs	158
7.7	Referential NPs, non-referential NPs, and transitivity.....	160
8	Additional Interactional Factors Results	163
8.1	Increments.....	163
8.1.1	Forms of NPs in increments in Ecuadorian Spanish.....	172
8.1.2	Summary of NPs in increments	175
8.2	Repetitions	176
8.3	Co-constructions	179
8.4	Lists.....	182
8.5	NPs in 'Other' syntactic role.....	185
8.5.1	Referential 'Other' NPs	186
8.5.2	Non-referential 'Other' NPs	188
8.5.3	Summary of 'Other NPs'	192
8.6	Given Full NPs.....	193
8.7	Summary.....	202
9	Conclusions	204
9.1	Summary of major findings	204
9.1.1	Grammar and competing motivations.....	204

9.1.2	Referentiality.....	206
9.1.3	Genre.....	207
9.2	Avenues of future research	210
List of Appendices.....		211
Appendix 1		212
References		212

List of Tables

Table 1: Pear Stories	43
Table 2: Summary of conversational data	46
Table 3: Verb classes	59
Table 4: Discourse Referentiality and Persistence in both genres (N=4405)	74
Table 5: Discourse Referentiality and Expression type in Javanese (Ewing 2005)	75
Table 6: Discourse Referentiality and Expression type in both genres (N=4405)	77
Table 7: Discourse Referentiality and Determiner presence for Lexical mentions in both genres (N=2665)	78
Table 8: Discourse Referentiality and Specificity in both genres (among personal pronouns and unexpressed mentions; N=1347).....	80
Table 9: Discourse Referentiality and Specificity in both genres (lexical NPs only; N=2665).....	81
Table 10: Discourse Referentiality and grammatical role in Spanish Narratives (lexical mentions only; N=693)	82
Table 11: Discourse Referentiality and grammatical role in Spanish Conversation (lexical mentions only; N=1972).....	82
Table 12: Specificity, Grammatical role and Discourse Referentiality of full NPs in Spanish Conversation (N=1972).....	83
Table 13: Specificity, Grammatical role and Discourse Referentiality of full NPs in Spanish Narratives (N=693)	86
Table 14: Coding for discourse referentiality	88
Table 15: Overview of NPs by discourse referentiality (N=4418).....	90
Table 16: Overall distribution of core vs. non-core NPs (N=4418)	90
Table 17: Overview of discourse referential NPs by activation status (N=1861)	91
Table 18: Persistence of new referents (N = 453)	92
Table 19: New referents and realization (N = 453)	95
Table 20: Given referents and realization (N = 1408).....	99
Table 21: Distance in clauses from previous mention for given full NPs (N=419)	100
Table 22: Realization of Given NP, Distance in Clause from Previous Mention and realization of previous mention in conversations (N=734)	103
Table 23: Realization of Given NP, Distance in Clause from Previous Mention and realization of previous mention in narratives (N=655).....	104
Table 24: New NPs and definiteness (lexical NPs only; N = 414).....	108
Table 25: New NPs and type of determiner (lexical NPs only; N = 414)	108
Table 26: Proportion of new referents introduced with definite article (excluding proper names) (N = 372)	110
Table 27: Overall proportion of referential NPs (lexical NPs and pronouns) modified by relative clauses (N=1093)	115
Table 28: New referents and relativization (N=440).....	116
Table 29: Given referents and type of modification (N=423)	117
Table 30: Given referents modified by relative clause (lexical NPs and pronouns; N=784).....	118
Table 31: New referents and syntactic role (N = 423).....	119
Table 32: Given referents and syntactic role (N=1408)	122

Table 33: New referents (in A, S, DO and IO roles) and semantic class of verb (N=378)	125
Table 34: Distribution of verb type for Given and New referents in narratives (N=754)	126
Table 35: Distribution of verb type for Given and New referents in conversations (N=912)	126
Table 36: Other types of given referents (N=133)	127
Table 37: Realization of 'members of a set' (N=79)	128
Table 38: Modification of 'members of a set' (N=54)	129
Table 39: Syntactic role and activation status in Narratives (N=790)	134
Table 40: Syntactic role and activation status in Conversations (N=1071)	134
Table 41: Form of expression for subjects of transitive verbs (N=312)	135
Table 42: Core vs. Non-core non-referential NPs (N=2424)	144
Table 43: Non-referential core NPs by syntactic role (N=777)	145
Table 44: Subject realization of non-referential NPs	148
Table 45: Realization of object non-referential NPs	149
Table 46: Modification of full non-referential NPs in core syntactic roles	149
Table 47: Non-core non-referential NPs and syntactic role (N = 1647)	151
Table 48: Frequency of prepositions in Narratives and Conversations (N=929)	151
Table 49: Realization of object of preposition NPs	153
Table 50: Modification of full NPs that are objects of prepositions	153
Table 51: Realization of predicate nominals	154
Table 52: Modification of predicate nominals	155
Table 53: Realization of adverbial NPs	156
Table 54: Modification of full adverbial NPs	156
Table 55: Modification of vocatives	157
Table 56: Modification of proper names	158
Table 57: Non-referential NPs and syntactic role (N = 1647)	159
Table 58: Third person referential and non-referential subjects and verb transitivity in narratives and conversations (N = 1596)	161
Table 59: Proportion of direct objects by genre (N=4418)	161
Table 60: Referentiality of NPs in transitive clauses in narratives (N=923)	162
Table 61: Referentiality of NPs in transitive clauses in conversations (N=1668)	162
Table 62: Increments by syntactic role (N=109)	173
Table 63: Increments by realization (N=109)	174
Table 64: Determiners and full NPs in increments (N=96)	175
Table 65: NPs in repetitions by syntactic role (N=47)	179
Table 66: NPs in repetitions by realization (N=47)	179
Table 67: Proportion of NPs fulfilling interactional functions in conversations (N=3190)	185
Table 68: Distribution of NPs in syntactic role 'Other' by referentiality (N=280)	186
Table 69: Activation status of referential NPs in 'Other' syntactic role (N=128)	186

1 Introduction

1.1 Discourse Analysis

The study of language, or the discipline of linguistics and related fields, has been revolutionalized over the course of the last five decades with the advent of voice recorders and other technology that have allowed researchers to capture spontaneous language at the time of production and preserve it in a way in that it can be analyzed at a later date. As such, the field of Discourse Analysis has grown exponentially as researchers have access to phenomena that were once fleeting. Consequently, our understanding of language structure and grammar has been radically altered.

At first glance, to the untrained eye, a transcript of conversational language appears messy and disorganized. Language in its raw and unedited form appears chaotic—full of unfinished sentences, the haphazard stringing together of clauses or fragments of speech with no punctuation, and the frequent occurrence of nuisances or interruptions such as pauses, hesitations, truncated words, laughter, and other speakers. It bears little resemblance to the more polished written language. However, when our expectations are adjusted and we dismiss the notion that speech is but a poor representation of an underlying grammar, structure and order become visible. Organization is viewed in the highly sophisticated turn-taking system, which allows for speakers to practically seamlessly transition from one speaker to another with remarkably few problems (cf. Sacks, Schegloff & Jefferson 1974). Structure is revealed in the architecture of the clause, in which the same patterns are repeated over and over, building conversations and narratives (cf. Chafe 1994, Du Bois 1987). Punctuation is found in the

intonation patterns that regulate the flow of information and contribute to the turn-taking system (cf. Chafe 1994, Ford & Thompson 1996).

What, then, are the factors that contribute to the organization and structure of naturally occurring language? A great deal of work has concentrated on the cognitive factors that shape the grammar of spoken language. The work of scholars such as Wallace Chafe (1980, 1987, 1994, *inter alia*), Talmy Givón (1981, 1983, 1995, *inter alia*) and John Du Bois (1980, 1987, 2003) have shown how cognitive concerns contribute to patterns of language use. Through these and similar studies, we have learned a great deal about how speakers manage the translation of thoughts and memories into coherent speech. In the process of verbalization, speakers are faced with limitations regarding how much information can be in the focus of consciousness at one time (cf. Chafe 1994), as well as how to keep track of multiple referents in a conversation at a time without creating confusion. Grammar reflects speakers' need to distinguish between what information is presumed to be in the consciousness and what is not.

Discourse analysis has also shed light on categorization. Hopper and Thompson (1984) show that while speakers universally orient toward *things*, which are represented by nouns, and *actions*, represented by verbs, it is only within actual discourse that categories are imposed on the forms. In a similar vein, Hopper and Thompson (1980) and Thompson and Hopper (2001) show that intransitive, transitive and ditransitive are not discrete grammatical categories, but that transitivity is scalar and depends upon several markers of transitivity within the clause. The valency of a verb is determined by its use in discourse, rather than the other way around. The significance of these and other studies is that language is a dynamic system which is grounded in use. Without discourse, there are

no nouns or verbs, nor transitive or intransitive verbs. Without discourse, our assumptions about linguistic structure are often misguided. Where the traditional view of language or grammar may be somewhat unidimensional, the analysis of discourse reveals that there are a multitude of interrelated dimensions (cognitive, pragmatic and interactional, among others) which bear relevance on linguistic form.

This study of language in its most natural form allows us a privileged glimpse into how social and interactional practices, which are absent from many other genres, influence grammar. In addition to the cognitive pressures associated with creating language on-line with no time to edit, speakers are also simultaneously faced with additional interactional concerns, ranging from taking turns to ways of expressing stance and attitude. These interactional pressures have grammatical consequences that are manifested in a number of ways.

Numerous studies have examined the turn-taking system (Ford & Thompson 1996; Ford, Fox & Thompson 2002; Goodwin 1981; Lerner 1991; Sacks et al. 1974; Schegloff 1996, *inter alia*). This research indicates that turn-taking is a highly organized and predictable system within conversation. Participants recognize the appropriate times in a conversation in which to take a new turn at talk, have also developed strategies for dealing with problems that arise when turn-taking does not follow as anticipated, and can even collaboratively construct sentences across two speaker turns.

One of the fundamental principals of interactional linguistics is the idea that grammar emerges as a set of patterns that arise in response to repeated actions (cf. Bybee 2007, 2010). For example, Thompson's (2002) study of complement clauses challenges the traditional view of these clauses as subordinate to complement-taking predicates

(e.g., *think, know, realize, wonder, etc.*) by showing that complement-taking predicates are more accurately described as epistemic/evidential/evaluative fragments denoting a speakers' stance toward the clause. Here, our understanding of these structures as main clause + subordinate clause must be abandoned as we see the rich interactional functions that they perform in conversation.

It must be noted that discourse analysis is not limited to just the study of conversation, but includes other spoken and written genres of language. It should be emphasized that one genre does not take precedence over another in the field of discourse analysis, rather the focus is on naturally-occurring data and the study of language beyond isolated sentences. This study, building upon previous studies of the cognitive and interactional dimensions of language use, examines two genres of spoken language—spontaneous conversations and monologic narratives—and finds that genre plays an important role in dictating how patterns emerge.

1.2 The present study

This study is a case study of just one part of speech—the noun phrase—as it is used in two genres of spoken Ecuadorian Spanish. Why was the NP chosen? The sheer abundance of NPs in any given corpus immediately makes it an attractive candidate for a quantitative study. It is through the quantitative study of language that we come to understand how routinized linguistic structures are the grammatical realization of recurrent social and communicative needs. In addition, NPs are a veritable gold mine of information about the underlying cognitive processes involved in language production—it is through the NP that we trace the information status of referents, study anaphora, and measure the information flow pressures. The frequency of the NP also guarantees that it

occur in a wide variety of interactional contexts, allowing for an understanding of how interactional factors help shape grammar.

Only third person NPs are examined here. One reason for this is purely practical—1st and 2nd person NPs are nearly absent from one of the genres under consideration (narratives of the Pear Film). It is also because third person has been less widely studied than first and second person in Spanish (e.g., Cameron 1994; Flores-Ferrán 2002; Travis 2005, 2007, *inter alia*) and because there are different factors that affect realization of first and second person NPs that do not apply to third person. Lastly, there is simply more variation in the information flow parameters (discourse referentiality, information status, specificity, identifiability) of third person NPs than first or second person, which are central to the investigation of the cognitive processes involved in language production examined here.

How do we investigate the role of noun phrases in the cognitive and interactional processes of language production? Or conversely, how do we investigate the role of cognitive and interactional pressures in shaping the grammar of NPs in spoken language? The best approach toward answering these questions is to first sort out the cognitive factors from the interactional pressures. For that reason, two genres of spoken language—monologic narratives and dialogic conversations—were chosen for this study. These genres differ from each other in two important ways—the information flow pressures and the level of interaction. In terms of cognitive processes, the narratives have higher information flow pressures (cf. Du Bois 1987), which allows us to understand how speakers maximize the available grammatical resources of a language to keep track of a high number of referents. The low information pressures of the conversations allow for a

comparison of the grammar of NPs under different conditions. As far as interactional concerns, the conversational data teems with interaction as speakers take turns, compete for the floor, finish either other's sentences, question what others have said, and jointly build narratives. Interaction is not absent from the narratives (cf. Schegloff 1982 and Goodwin 2007), but we can expect interactional concerns to be less pervasive in narratives than in the conversations and there may also be fewer types of interactional concerns (i.e., turn-taking is less important in a largely monologic narrative). Compare examples

(1) and (2), for example. In the first example, two speakers are jointly constructing a narrative, taking turns with each other, competing for the floor (overlapping speech is seen in square brackets, see the Appendix for a full list of transcription conventions), agreeing and disagreeing with each other. In the second example, the speaker is recalling the story on her own—she has the floor to herself, there is no one to help her remember details or to challenge her memory, nor to agree with her.

(1) Jointly constructed narrative in conversational data

A: ... (H) vos te fuiste una ves trotando con nosotros?
R: .. hasta el --
claro,
pues,
hasta= --
y después [Ø cogimos bus].
A: [que fue] --
no,
y esa camioneta,
que nos llevó hasta la Mitad del Mundo,
R: ah,
claro,
para llegar hasta la Mitad del Mundo.
A: .. (H) y que luego nos Ø querían llevar <@ más allá @>.
.. y Ø estaban borrachos,
Ø creo.
R: claro,
se iban para Calacalí,
creo.
A: .. ah,

a la costa,
a dónde también Ø se irían?

R: .. pero Ø estaban --
bien borrachos,
nosotros golpeábamos el [vidrio],

A: [hm],

R: .. para que Ø nos pare,
A: y Ø no nos pa- --
Ø nos paró más allá del redondel?

R: ... <@ claro @>.

A: una media [cuadra].

R: [y de ahí] Ø queríamos subir trotando de nuevo.

A: ... no,
Ø cogimos bus hasta .. <@ Pomasqui @>.

R: Ø cogimos bus y Ø nos bajamos en --
claro,
[en Pomasqui].

A: [Pomasqui].
y de ahí [Ø seguimos trotando].

A: ... (H) did you go jogging with us one time?
R: .. to the --
of course,
well,
to= --
and then [we caught a bus].
A: [that went] --
no,
and that pickup truck,
that took us to La Mitad del Mundo,
R: ah,
of course,
to get to La Mitad del Mundo.
A: .. (H) and then later (they) wanted to take us <@ farther @>.
.. and (they) were drunk,
(I) think.
R: of course,
they were going to Calacalí,
I think.
A: .. ah,
to the coast,
where would (they) go?
R: .. but (they) were --
really drunk,
we hit the [window],
A: [hm],
R: .. so that (they) would stop,
A: and (they) didn't st- --
(he) let us out past the roundabout?
R: ... <@ of course @>.
A: a half a [block].
R: [and from there] we all wanted to go jogging again.
A: ... no,
(we) took a bus to .. <@ Pomasqui @>.
R: (we) took a bus and (we) got off in --
right,
[in Pomasqui].
A: [Pomasqui].
and from there [(we) kept jogging].

(Fumar: 155-194)

(2) Excerpt from monologic narrative

Ø está yendo en la bicicleta,
eh,
.. en el que misma --
en el mismo camino,
pero en el sentido contrario,
pasa una niña,
en otra bicicleta.
.. (H) a lo que Ø están pasando juntos,
eh,
.. el sombrero del niño vuela,
(H) y el niño,
por regresar a ver el sombrero,
no ve una piedra grande y se choca.
... (H) eh,
Ø se choca y se caen las manzanas,
se cae el canasto,
se riegan todas las peras,
(H) la niña sigue su camino,
(H) y el niño a lo que= --
.. a lo que Ø esté en el piso,
se levanta su pantalón,
se baja sus medias,
se X su pierna lastimada,
y se da cuenta que a lado de él están parados tres niños.

(he) is going on the bicycle,
eh,
.. in the same --
on the same road,
but on the other side,
comes a girl,
on another bicycle.
.. (H) when (they) pass by each other,
eh,
.. the boy's hat flies off,
(H) and the boy,
upon looking back at the hat,
doesn't see a big rock and crashes.
... (H) eh,
(he) crashes and the apples fall,
the basket falls,
the pears all spill out,
(H) the girl goes on her way,
(H) and the boy when --
.. when (he) is on the ground,
raises his pants,
lowers his socks,
X his hurt leg,
and realizes that standing next to him are three boys.

(PS 100:131-154)

The corpus used for this study is described in detail in Chapter 3. The data are from a larger, multi-genre corpus of Ecuadorian Spanish. Fifteen narratives of the Pear Story were selected, totaling 6430 words, from which all 3rd person NPs were extracted (exclusions are discussed in Chapter 4). An equal number of speakers were chosen for the conversational data. A sample of third person NPs from these speakers were extracted from transcripts of ten conversations, totaling 74,673 words.

The findings presented in this work shed light on a number of important issues. The comparison of two genres allows for an understanding of what recurrent grammatical patterns are linked to the different actions of narrative and conversational data. For example, the distribution of the NPs in the narrative data reflect the higher information flow pressures of the genre in a number of ways—higher rate of transitive constructions, proportion of referential to non-referential NPs, and a stricter adherence to Preferred Argument Structure (cf. Du Bois 1987). Other grammatical patterns are seen at much higher rates in the conversations, such as the introduction of new referents with definite markers, which reflect genre-specific discourse needs.

Categoriality is explored, especially with respect to the distinction between referential and non-referential NPs, and the typology of the different groups of non-referential NPs. The terms referential and non-referential are used here in the sense of Du Bois (1980) and Hopper and Thompson's 1984 'discourse manipulable' sense. Referential NPs are those that track a referent whereas the different kinds of non-referential NPs perform different discourse functions, including predicating, classifying and characterizing. Compare the noun *sol* ('sun') in (3) and (4). It is the same lexical form, yet the discourse function, and consequently the grammar, is very different

between the two examples. In the first example, *sol* is referential, that is, it exhibits the prototypical functions of a noun—it refers to an entity that has continuity of identity within the discourse. In addition, it shows nominal morphology (i.e., the definite article), whereas non-referential NPs tend to exhibit less morphology characteristic of NPs (e.g., bare NPs, as in (4)). The difference between referential and non-referential NPS will be discussed in depth in Chapter 5, where the operationalization of referentiality is explained, but perhaps one of the easiest ways to understand the difference between the two is the often used analogy of computer files. When a referential NP is used for the first time in a conversation, the listener creates a new file or accesses an existing cognitive file for that NP. The identity of this NP is stable, information about this particular NP is stored in the mind, and new information can be added to the file. Non-referential NPs, on the other hand, such as *sol* in (4), have no associated cognitive file. The noun *sun* in this example is not used to talk about the entity that is yellow, is a star, and is located approximately 149.6 million kilometers from the Earth. The discourse function of *sol* in this second example is part of a verbal predicate. As we will see throughout this volume, there is ample evidence in the discourse that the grammar of NPs reflects the differences between referential and non-referential NPs and that speakers use them for quite distinct discourse functions.

(3) Referential NP

F: amarillo es **el sol**,

F: **the sun** is yellow,

(Birthday: 286)

(4) Non-referential NP

A: pero está haciendo **bastante** **so**l.

A: but it's really sunny (lit. '(it) is making a lot of sun')

(Clases: 7)

The previously understudied roles of free NPs (see *gastroenteritis* and *hepatitis* in (5)) are thoroughly explored in this dissertation. We see that speakers use these NPs for a variety of functions related to both information flow and interactional concerns.

(5) Free NP

E: qué problema es el que da,
cuando está --
.. cuando hacen cosas sucias?
...(2.0) cómo se llama?

L: .. **gastroenteritis**.

E: .. **gastroenteritis**.
o la otra,
que te pones amarillo?

L: .. **hepatitis**.

E: .. **hepatitis**.

E: what problem is it,
When it -
.. when they do dirty things?
...(2.0) what's it called?

L: .. **gastroenteritis**.

E: .. **gastroenteritis**.
or the other,
when you turn yellow?

L: .. **hepatitis**.

E: .. **hepatitis**.

(Food, 6:13-622)

As we trace the form and distribution of NPs throughout discourse, we see how the varying forms and roles reflect local cognitive and interactional demands. At the local level, we can situate a NP within the immediately surrounding discourse and understand why it is that a speaker has chosen a particular way of encoding the NP. On a more global scale, we see which patterns are borne out over and over again, attesting to the robustness of particular communicative needs. The findings here attest to the view that spoken

language is in fact orderly and structured and that speakers are highly capable of managing both cognitive and interactional demands in the on-line production of language.

2 Research Questions

2.1 Research questions

This chapter presents an overview of previous research that has explored factors related to NP realization and distribution and draws attention to the limits and shortcomings of this previous research. A great deal of research has focused on information flow, and there has recently been an interest in how interactional factors may also contribute to NP form and distribution. What is lacking is a methodologically refined study of information flow, particularly with respect to Spanish, and a holistic approach to the form and distribution of NPs which takes into account factors beyond information flow. The current study will enhance our understanding of NP form and distribution by 1) addressing methodological concerns, and 2) broadening our understanding of the effects of both cognitive and interactional factors.

The discussion is presented in three sections. First, there is a discussion of what ‘grammar’ means as used here. Secondly, there is discussion and definitions of information flow properties and how they have been shown to influence NP form and distribution, and an overview of how the present study will contribute to the body of literature dedicated to understanding information flow. The third section reviews the literature on other factors known or hypothesized to influence NP realization.

2.2 Grammar and competing motivations

When engaging in conversation or story-telling, speakers need to use language in a way that effectively, accurately and efficiently reflects the message(s) they are trying to

convey. From a functional perspective, grammar is a resource that speakers employ to negotiate the different needs presented in the discourse context. Grammar arises from repeated patterns of use (cf. Hopper 1987, Bybee & Hopper 2001) and is said to ‘code best what speakers do most’ (Du Bois 1985:363).

To the non-linguist and to many linguists (e.g., Chomsky 1965), grammar is often understood to be a rigid set of rules that speakers conform to. To the functional or socio-linguist, grammar is adaptable, flexible, and speakers are presented with choices in constructing grammatical units. However, these choices are not entirely up to the whim of the speaker, rather they are constrained by a number of different factors. The two types of factors of interest here are the cognitive factors that influence the flow of information in a discourse and the social and interactional factors that shape conversation.

Du Bois (1985) points out that discourse is shaped by competing motivations. Unmarked grammar, he says ‘is distributed on the basis of “unmarked” (statistically most frequent) discourse’ (Du Bois 1985:362). With this in mind, we expect to see broad patterns of use emerging, repeated across speakers and among genres. This is the case with information flow—speakers are always managing the flow of information, and as a result, the broad and statistically frequent patterns of grammar reflect this discourse need. For example, Preferred Argument Structure (cf. Du Bois 1987) has been observed in a number of typologically unrelated languages. The two key findings are that there tends to be only one new argument per clause and that new arguments are typically found as subjects of intransitive verbs or objects. Clauses that are consistent with the pragmatic and morphosyntactic tendencies of PAS can be considered unmarked. Marked grammar, on the other hand, may be evidenced in situations in which another factor, for example

interactional needs, outweighs the cognitive motivation behind the expected pattern, for example, a deviation from PAS. There may be instances in which there is no clash between the competing motivations, that is, the grammar of the language allows for one grammatical coding that accommodates both factors. On the other hand, there may be times when the grammar of the language does not allow for both factors to be accommodated, therefore, speakers must ‘mark’ the grammar according to the strength of each external motivation.

Regarding the grammar of noun phrases in Spanish, speakers are faced with a number of choices. For third person, they have to choose among three types of realization (full NP, pronoun and unexpressed), and the choice is commonly thought to be governed by information flow, pragmatic and discourse-sensitive considerations (Blackwell 2003, Silva-Corvalán 1994)¹. In addition, they must choose how to modify the noun (i.e., choice of determiner), and how to fit the noun into the rest of the discourse (i.e., syntactic role). Information flow is one of the most influential forces in helping speakers to make these choices, and the effects of information flow have been widely observed in a variety of languages, with speakers roughly adhering to universal tendencies such as presenting only one new piece of information at a time. Yet speakers are also often dealing with social and interactional pressures at the same time which can also effect how they choose to code and employ NPs in a discourse. This purpose of this study is to examine the patterns of Spanish NP grammar and to look for evidence of the different types of pressures that contribute to the choice of a particular NP form in spoken discourse.

¹ One of the basic assumptions here is that discourse shapes grammar. For alternate (formal) approaches to subject expression in Spanish, see Toribio 1993, 1996 and 2000.

2.3 Information flow

Information flow involves the relationship between the activation status, identifiability, specificity and referentiality of a referent and the bearing that this has upon the linguistic coding of the referent. When participating in a conversation or other speech event, speakers and interlocutors are faced with the task of keeping track of which referents are being talked about and which NPs serve other discourse functions (e.g., orienting, predicating) and require a different type of cognitive effort. When coding NPs, speakers use a combination of linguistic devices to express to the hearer whether or not the information is new, whether or not they expect the hearer to be able to identify the referent, and whether or not the referent is ‘important’, or in other words, whether or not it has continuity of identity within the discourse. The devices that speakers use are a combination of morphological markers (e.g., definite versus indefinite articles), NP realization (e.g., full NP or pronoun), and syntactic role.

The different information flow properties of NPs affect different areas of NP grammar—specifically the realization of the NP, definiteness and syntactic role. In the sections that follow I provide explanations of the different dimensions of information flow, show how they are interrelated, and how they have been found to affect NP distribution and form. In addition, I draw attention to research questions that have not yet been addressed.

2.3.1 *Activation status*

Many researchers have highlighted the role of activation status in NP realization and distribution (cf. Ariel 1990; Du Bois 1987; Givón 1983; Gundel, Hedberg & Zacharski 1993; *inter alia*), and as this research has emerged, advances have been made toward

defining activation state and operationalizing methods of data analysis according to activation state. As Chafe (1976) explains it, activation status has to do with where a speaker assumes a referent is in the consciousness of a hearer. A concept can be considered active (given) if it is 'knowledge which the speaker assumes to be in the consciousness of the addressee at the time of the utterance' (Chafe 1976:30). Conversely, it is inactive if it is not in the consciousness of the addressee (or, more specifically, not assumed so by the speaker). For a concept to be inactive, it does not mean that the addressee has no prior information about the referent, but that the speaker assumes that the referent is not in the addressee's focus of consciousness at that particular moment in time. However, Chafe also points out that this characterization is based upon introspection, and it is irresponsible for the linguist to assume that he or she has access to what the speaker is thinking.

Instead, the linguist must rely on cues from the discourse, including other mentions of the referent and the forms which the speaker uses to code the referent to know whether or not they expect the speaker to recognize them as already active. Many researchers (see, for example, Ariel 1990, Chafe 1987, Givón 1983, Prince 1981, 1992) agree that if a referent has been mentioned in the recent discourse, then it is logical that the speaker and hearer will regard it as active or given, but is debatable how long a referent stays active after it has been activated. Chafe (1987) suggests that active concepts are found in short-term memory, and given that short-term memory can only hold seven plus or minus two items at a time (Miller 1956) the number of active concepts at a time must be very limited. Givón (1983) first proposed a measure of twenty clauses for referential distance, and later modified the number to ten (1995). As Fox (1987)

observes, however, this linear view of discourse and anaphora does not always represent what speakers are *doing* in the discourse, and she shows in her research how the notion of conversational sequences and other discourse structures are relevant to studies of anaphora. Tomlin (1987) also finds that discourse structure is more relevant than elapsed time (as measured in clauses in Givón's work); specifically he reports that episode boundaries are crucial in the realization of referents in narrative data.

While the precise amount of time that a referent remains active in the mind and the role of other factors is unknown, most would agree that activated referents can and do change status after a relatively short period of inactivity. This begs the question of what happens to active referents after a certain amount of time. Prince's 1981 model and Chafe's 1987 model with three states are two of the most often cited models. Prince's model is more detailed than Chafe's model, with seven different categories of assumed familiarity to his 3 activation statuses, but both models share several core traits. In both models there are 'new' and 'given' referents, with intermediate states.

Chafe (1987) argues that active referents are moved to an intermediate activation state, semi-active ('accessible'). Semi-active referents are not limited to those which have been previously active, and Chafe (1987) considers those referents that are part of a greater schema to also be semi-active, i.e., 'they are more accessible to recall than they would have been as inactive concepts' (Chafe 1987:29). Note that Givón's (1983) model does not include a semi-active or accessible state.

The third activation status distinguished by Chafe is inactive, or new. This status applies to referents that have not been active or semi-active via previous mention, nor can they be considered semi-active due to association with a schema. According to Chafe

(1987:31), the change in state from inactive to active carries the heaviest cognitive burden; the cost of activating these referents is much higher than other changes in activation status (i.e., activating a semi-active concept, the change from active to semi-active). As a result, it is expected that this higher cost will be reflected linguistically.

Prince's 'brand-new' (both unanchored and anchored) corresponds most closely to Chafe's 'inactive' or 'new'. New-unused², noncontaining inferrable and containing inferrable are all similar to Chafe's 'semi-active' or 'accessible', and her 'textually and situationally evoked' referents could correspond to Chafe's semi-active or active. Prince's later (1992) work reflects the important distinction between *hearer* new/old and *discourse* new/old.

Example (6) illustrates several changes in activation state. In line 2, the boy is first introduced, representing a change from inactive to active. The cognitive burden is represented in its expression as a full NP. The referent then continues as an active concept, seen in lines 5, 7 and 8, which show lesser coding (pronoun in line 5, unexpressed in lines 7 and 8, and again pronominally in 10, 11 and 12). With the introduction of the girl in line 15 the boy becomes semi-active, and when becomes active again he is referred to pronominally (line 16), which is more coding than the previous unexpressed mentions in lines 7 and 8, but less than the first mention and activation (line 2).

(6) Changes in activation state

² Other researchers may have different interpretations of how Prince's categories align with Chafe's. O'Dowd (1990), for example, codes Prince's Unused referents as New, and only inferrable and containing inferrable as Accessible.

1 el momento que él estaba distraído recogiendo,
2 pasó **un niño**.
3 ... (H) con --
4 en una bicicleta.
5 .. **é=l** se acer- --
6 se acercó,
7 (H) y **ø** quiso coger --
8 .. **ø** quiso coger una,
9 pero=,
10 la tentación **le** --
11 .. **le** --
12 .. (H) **le** venció y **ø** cogió todo un canasto en su³ bicicleta.
13 (H) al seguir por el camino,
14 eh,
15 pasó una niña,
16 y **ø** **le** botó el sombrero.
17 (H) al momento de botarle el sombrero **él** se --
18 **é=l** .. se descuidó y cayó

1 the moment that he was distracted collecting (pears),
2 along came **a boy**.
3 ... (H) with --
4 on a bicycle.
5 .. **he** appr- --
6 approached,
7 (H) and **(he)** wanted to take --
8 .. **(he)** wanted to take one,
9 bu=t,
10 the temptation to **him** --
11 .. to **him** --
12 .. (H) conquered **him** and **(he)** took a whole basket on his bike..
13 (H) upon following the path,
14 eh,
15 along came a girl,
16 and knocked off **his** hat (lit. knocked off the hat to **him**).
17 (H) with the hat being knocked off **he** --
18 **he=** .. got distracted and fell.

³ Although the boy is not the referent of this NP, he is referred to by the use of the possessive adjective (i.e., the possessor) and this also contributes to the maintenance of the activation state.

It should be noted that Chafe (1987:25) applies different information states to *concepts*, which includes ‘the ideas of objects, events, and properties’, which are expressed as verbs, adjectives, etc. In practice, however, most researchers find it practical to only code for the information status of NPs. As Du Bois (1987:816-17) points out, ‘it is easier to decide whether a nominal reference to an entity is new or not, or whether there are one or two entities within a single clause, than it is to decide whether a particular sequence of verb + adverb + complement constitutes three two, or just one (complex) piece(s) of new information’. Furthermore, he states that ‘it may be that speakers attend to new nominal information on a separate cognitive track, as it were, from new verbal information’ (1987:817).

While there is a consensus that given-new distinction is not simply binary, but rather scalar or multi-dimensional, and that different factors (previous mention, shared knowledge, participation in a schema, episodic structure, etc.) can influence the information status of a referent, in practice most researchers find it sufficient to distinguish between new and non-new (given), although they arrive at these distinctions in different ways. Although it is possible to examine a text or transcript and find clear examples of each of Chafe and Prince’s categories, when the activation status of each and every NP in a dataset must be accounted for, researchers are faced with the task of operationalizing these notions. A variety of different practices for operationalizing this variable have been used, with both Chafe and Prince’s models often cited. Du Bois (1987) codes his data according Chafe’s three states of activation: given referents are those that have been previously mentioned or are notably present in the context of

situation (e.g., speaker and addressee), accessible are those that are part of an already evoked schema or where a previous mention occurred at a distance greater than 20 Intonation Units (IUs), and new mentions are those which have not been previously mentioned. However, he finds that given and accessible referents pattern similarly, therefore he combines these two and contrasts new and non-new states in his analysis. Ashby and Bentivoglio (1997) follow the same criteria as Du Bois (1987), and Bentivoglio (1993) similarly distinguishes between new and non-new in her study of Full NPs in spoken Spanish. Ashby and Bentivoglio (1993) cite Prince's model in their study of Preferred Argument Structure in French and Spanish, and consider those that are new to correspond to Prince's 'brand new', and all others to be non-new. Ewing (2005) uses previous mention as the only criterion for distinguishing between given (previously mentioned) and new (not previously mentioned) referents. Given that this is the only criterion that can be neatly operationalized for 3rd person referents, it is this measure that will also be used in this study. However, alongside this I consider other dimensions, including distance from previous mention.

This study seeks to add to the body of literature that explores how speakers code NPs according to information status. Specifically, this study will examine the preferred form and distribution of both given and new referents with special attention to preferred presentational structures across genres. It also considers other dimensions of information flow in conjunction with information status and fills a gap in previous research by showing how different dimensions of information flow are intertwined with each other and have bearing on the grammar of NPs.

2.3.2 *Discourse Referentiality*

It is clear that not all NPs have the same discourse function. Prince (1981:235) states this very succinctly: ‘all discourse entities in a discourse-model are represented by NPs in a text, though not all NPs in a text represent discourse entities’. Hopper and Thompson (1984:710) explain that the prototypical function of nouns is to ‘introduce participants and “props” and to deploy them’. These NPs are referential, in the sense that they ‘have continuity of identity’ (Du Bois 1980:208), and ‘are important in the subsequent discourse’ (Hopper & Thompson 1984:711). In (7), the referent ‘Mishell’s cousin’ (and subsequent mentions) is clearly referential. In the first mention, a participant (*la prima de la Mishell*) is introduced, and is later seen to have both continuity of identity and to be important in the discourse.

(7) Referential NP

A: .. (H) y te acuerdas que una vez **la prima de la Mishell**,
<@ se pegó una matada @>?
.. (H) jugando a las escondidas?
... en la casa mismo de la Mishell,
afuera que [había tierra]?
R: [cuál prima]?
A: **una prima de la Mishell**,
que tenía.
R: **la Verónica**?
A: ah=.
.. ø iba corriendo y shu=,
ø se resbala,
y ø se pega una matada,
y ø se raspó la rodilla,
te acuerdas?

A: .. (H) and do your remember one time that Mishell's cousin,
 <@ took a fall @>?
 .. (H) playing hide-and-go-seek?
 ... at Mishell's house,
 outside where there [was dirt]?
 R: [which cousin]?
 A: one of Mishell's cousins,
 that (she) had.
 R: Verónica?
 A: yeah=.
 .. **(she)** was running and shu=,
(she) slips,
 and **(she)** takes a fall,
 and **(she)** scrapes her knee,
 remember?

(Fumar:218-232)

Du Bois (1980:210) noted that ‘the form of a nonreferential mention is not responsive to the presence or absence of a prior mention’, and as this study will show, it is important to exclude non-referential NPs from studies of how grammar is shaped by information status. For example, the NP *escalera* (‘ladder’) in (8) below is non-referential⁴ (a detailed discussion of the different categories of discourse referentiality and their grammatical manifestations is presented in Chapter 5), and as such, is not responsive to the effects of information status, as seen in the coding as a full NP in three mentions over four clauses (and is more of a ‘prop’ than a central figure to the action), as opposed to the subject (clearly an ‘actor’), which has been established and receives no overt linguistic coding here.

(8) Non-referential NP

∅ se vuelve a subir **la escalera**,
 ∅ se --
 ∅ se limpia su ropa,
 y se ∅ vuelve a subir a **la escalera**.
 .. (H) ∅ se sube a **la escalera**,
 (he) goes back up **the ladder**,

⁴ Note here the distinction between referential and referring. Most linguists would agree that the NP *escalera* in example (8) is a referring NP, but the important distinction in this work is that it is non-referential in Du Bois' (1980) sense.

(he) --
(he) cleans his clothes,
and (he) goes back up **the ladder**.
.. (H) he goes up **the ladder**,

(PS100: 59-63)

In addition, several scholars have noted that non-referential NPs exhibit different syntactic behavior than referential NPs. Hopper and Thompson (1984) show that referential NPs take different forms than non-referential NPs in that they show the morphological marking of NPs, whereas non-referential NPs lack some or all of the morphological marking associated with NPs. Thompson (1997) finds that core arguments tend to be tracking (i.e., referential), whereas obliques tend to be non-tracking (i.e., non-referential). Torres Cacoullos and Aaron (2003) find that a significant amount of bare, lone English origin nouns in New Mexican Spanish tend to be non-referential, but they do not show that there is a tendency for non-referential nouns in general to be bare. Du Bois (1980:212) and Ono and Thompson (1994) also note the correlation between non-referentiality and lack of determiner in English.

While the difference between referential and non-referential NPs has been well-noted in the literature (cf. Du Bois 1987; Givón 1981; Hopper & Thompson 1984; Prince 1981), its importance is often overlooked and it is not apparent that researchers distinguish between the two types of NPs when conducting quantitative research on NPs, particularly where information flow is concerned.

A survey of ten studies that examine the relationship between information status, NP form and distribution (Ashby 1995; Ashby & Bentivoglio 1993, 1997; Bentivoglio 1993; Du Bois 1987; Durie 2003; Ewing 2005, Kärkkäinen 1996; Matsumoto 2003; Thompson 1997) reveals that only one researcher (Durie 2003) states that they excluded non-referential NPs from their data. While some take referentiality into account to some

degree, many appear to disregard it. Ashby & Bentovoglio (1997) only include full, specific NPs in the S and O syntactic roles in their study; as will be seen in Chapter 5, these criteria correspond with discourse referentiality for Spanish. Thompson (1997) excludes Predicating NPs, for example *ornaments* in the phrase *put up ornaments*, which are non-referential. She does, however, include obliques and codes those for information status. It should also be noted that although Durie (2003) does mention that non-referring nominals were excluded from his study of Acehnese, and he discusses the differing discourse roles of non-referring NPs⁵, he does not clearly explain how he distinguished between referring and non-referring NPs in selecting his data, and says (2003:193n10) that ‘the typology of non-referring expressions is too complex to explore here. It remains an important topic for research at the grammar-discourse interface’.

In attempting to understand discourse referentiality, the analogy of a referent to a computer file is often invoked (cf. Du Bois 1980:220-23, 1987:817). Referential NPs can be understood as having cognitive files—when a new referent is introduced, a new file is opened, and as that NP is tracked throughout the discourse, new information is added to that file. Non-referential NPs do not have an associated file, and as they do not have continuity of identity, they are not tracked in subsequent discourse. As Du Bois (1987) and Croft (1991) have claimed, and as I will argue here, the evidence suggests that speakers attend to new, referential NPs on a separate cognitive track than new (in the sense of not previously mentioned), non-referential information.

Importantly, unlike activation status (e.g., given and new), where researchers are often very clear in defining how they code these variables, a replicable way of

⁵ The terms referential, tracking, and referring have all been used in the literature in different ways. It is my understanding that the NPs that Durie calls non-referring are the same as the NPs that are labeled non-

operationalizing the coding for referential versus non-referential NPs has not been proposed. Even where researchers do distinguish between the two types of NPs (e.g., Durie 2003), they are often vague in describing how they code for referentiality. What is lacking, then, in the research is 1) a replicable and operationalized way to code for the difference between referential and non-referential NPs, and 2) research which directly compares the different variables that influence the form and distribution of referential and non-referential NPs. This investigation offers a comprehensive operationalization of referentiality based on morphosyntactic properties identified through empirical analysis.

2.3.3 Identifiability

Identifiability, or whether or not the speaker believes that the hearer will be able to identify the referent, is often linked to information flow. On the one hand, it seems very simple: given referents within a discourse should all be identifiable, and among new referents, some will be identifiable and others will be unidentifiable (cf. Thompson 1997:68). However, it is not always the case that once a referent is introduced into the discourse it becomes identifiable to the speaker. In example (9), we can see how identifiability is negotiated between the speakers. After the first mention of *tu pelada* ('your girlfriend') by speaker R, speaker J asks for clarification. At the same time, speaker R has realized that speaker J is not making the connection, so he gives more information to help J identify the referent, he says *la recepcionista* ('the receptionist'). It is at this point that the identity of this referent, which was first introduced as definite, is established. And J goes on to explain that this person is not, in fact, his girlfriend, she is his employee, and therefore the first mention of this person cannot truly refer to her. As was discussed above, the notion of 'givenness' is complicated at best, and as this example

referential in this work.

shows, identifiability is even more difficult to define in a manner which the linguist can use to accurately identify and code NPs for identifiability.

(9) Negotiating identifiability

R: y tu pelada?
... [la recepcionista]?
J: [cuál pelada]?
R: <@ no,
mentira @>.
J: Ø no es mi pelada,
Ø es mi --
Ø es mi [empleada].
R: [Ø es mi] recepcionista.
no,
mentira.
J: Ø es la que me pasa las llamadas.
R: @@@@
J: @@@@
(H) no=,
Ø no es mi pelada.
y además=s,
yo no creo en el amor.

R: and your girlfriend?
... [the receptionist]?
J: [what girlfriend]?
R: <@ no,
just kidding @>.
J: (she)'s not my girlfriend,
(she)'s my --
(she)'s my [employee].
R: [(she)'s my] receptionist.
no,
just kidding.
J: she's the one that transfers my calls.
R: @@@@
J: @@@@
(H) no=,
(she)'s not my girlfriend.
and anyway,
I don't believe in love.

(Jugo:1089:1106)

Identifiability is often linked to the use of articles. For languages like English and Spanish which have definite and indefinite articles, the indefinite article is said to be used for unidentifiable mentions, and the definite article (and by extension, other definite markers such as possessive adjectives and relative clauses (Du Bois 1980:273)) are

thought to mark identifiability. However, Du Bois (1980) shows that this is not always the case; in his data he finds first mentions with the definite article and second or subsequent mentions with the indefinite article, and either article with non-referential NPs. Similar examples were found in the data used in this study. Example (10) illustrates a definite first mention, which is unidentifiable, and (11) shows a second mention of the boy occurring with the indefinite article.

(10) Definite first mention⁶

primero **el= señor** recogía sus manzanas,
first **the= man** was collecting his apples,

(PS36:17)

(11) Indefinite identifiable NP⁷

... primero pasa ... un niño,
no entendí mucho,
pero pasó **un niño** con un= --
... con un becerro,

... first along comes a boy,
(I) didn't understand much,
but along came **a boy** with a= --
... with a calf,

(PS12:20-23)

Many researchers report coding NPs for identifiability (see, for example, Bentivoglio 1993; Ewing 2005; Thompson 1997), yet a standard method for coding for identifiability is not widely employed. Bentivoglio (1993), for example, does not state how she codes for identifiability. Thompson (1997) mentions the link between the definite article and

⁶ The way in which the Pear Stories were collected also draws into question the coding for identifiability as the speakers were asked to relate the story to the investigators as if they were telling the story to someone who had not viewed the film, when in actuality, the researchers had seen the film.

⁷ Some may argue that this is a repetition of the first mention and therefore not a true subsequent mention. It should be noted that there are no prosodic clues (i.e., truncation) that this the second mention is a repair intended to replace the first mention, and in addition, it is not an exact repetition, as the tense changes from present in the first mention to past in the second mention. Du Boi (1980:260) refers to these types of

identifiability, but does not state if she uses this as a criterion for determining identifiability or not. Ewing (2005) provides a very comprehensive account of how he codes for identifiability of a NP, via seven different identifiability pathways. A system such as the one that Ewing employs for Cirebon Javanese may not easily be replicated across researchers and languages, however, as notions such as what constitutes ‘world knowledge’ may differ from researcher to researcher, or from speaker to speaker, and may require that the researcher attempt to intuit what constitutes world knowledge for a particular speaker.

Coding for identifiability would seem to require some insight into a speaker’s thought process, or at least knowledge of what a speaker considers their interlocutor capable of identifying. As linguists do not have the privilege of accessing speaker’s thoughts, nor can they predict with any great certainty what two referents will be mutually identifiable to two speakers, it is very difficult to objectively code for identifiability. In addition, the method of data collection used in collecting the narratives for this study complicates the ability to determine what a speaker might judge to be identifiable.⁸ While the notion of identifiability is very useful to understanding the differential treatment of new referents, for this study, identifiability was not coded independently, instead it is evidenced through the speakers’ use of definite and indefinite articles with new referents (see Chapter 6).

repetitions as ‘degression-and-repeats’ and notes that they are responsible for ‘late indefinites’ in English language Pear Film narrations as well.

⁸ Participants narrated the Pear Film to the investigators, whom they knew had seen the Pear Film. However, they were instructed to narrate the film as if they were telling the story to someone who had not viewed the film. This differs from the method of collection described in Chafe (1980:XV), in which narrators were told that they were narrating the film to someone who had *not* previously viewed the film.

2.3.4 Specificity

Another NP property linked to information flow is specificity. NPs can fall into two types of classes: general and specific. General NPs refer to a class of entities, or a member of a class which is interchangeable with others, as in (12)⁹. Specific NPs are NPs of unique reference, they are not interchangeable with other members of the class, such as (13).

(12) General NPs

B: inclusive en **los aviones** hay --
hay un --
un gran cuidado de las --
(H) de **las azafatas**,

B: including in **(the) airplanes** there is --
there is a --
great care from the --
(H) from **the flight attendants**,

(Hermanos:358-361)

(13) Specific NP

R: no ves ø que está súper descargado ese teléfono?
R: don't (you) see that **that phone** is super uncharged?

(Jugo:819)

Specificity has also been linked to anaphoric coding of an NP. Ewing (2005) shows that specific NPs are more likely to be referred to with lesser coding (i.e., pronouns, unexpressed mentions) than general NPs. This is likely due to the relationship between specificity and discourse referentiality. As most general NPs are also non-referential, their linguistic coding is not subject to influence by information status, and therefore tend not to be referred to with pronouns and unexpressed mentions. Conversely, specific NPs tend to be referential, and therefore may be coded as pronouns and unexpressed mentions when given.

Ashby and Bentivoglio (1993) incorporate specificity into their analysis of PAS in French and Spanish and find that it has significant correlations with two phenomena—newness and syntactic role. They find that new participants are more likely to be general than specific (generalizing and particularizing in their terms), and that in Spanish, specific NPs are more likely to be coded in the S (subject of an intransitive verb) role than in the O (object) role. Ashby (1995) reports additional findings for new referents in French, with specific referents more likely to appear in locative, subject of copular verb, S, oblique, *avoir*-cleft and *avoir*-simple syntactic roles, while dispreferred roles for specific NPs include A (subject of a transitive verb), O, and two structures with *y'a*. These last two studies do not exclude non-referential NPs from their data, and it is likely that this has resulted in a skewed look at the influence of specificity on syntactic role of new referents.

In this study, the relationship between discourse referentiality and specificity is explored.

2.4 Genre (information flow pressures)

Variation in distribution of linguistic form between genres is well noted. Biber (1988) dedicates an entire volume to genre effects on a wide range of linguistic phenomena, from tense and aspect to negation. Biber et al. (2006) conduct a Multi-Dimensional analysis¹⁰ of register variation in Spanish and describe six different dimensions of

⁹ A terminological distinction between general and generic should be made here. General NPs are all NPs that are not specific; generic NPs are a type of general NP, but there are other general NPs which are not generic.

¹⁰ Multi-Dimensional analysis involves constructing a large corpus of different genres and then comparing “the co-occurrence patterns of linguistic feature through a factor analysis of frequency counts” (Biber et al. 2006:6), then functionally interpreting the factor analysis. See Biber (1995: Chapter 5) and Conrad and Biber (2001: Chapter 2) for more details on Multi-Dimensional analysis.

variation (that correspond to different registers or genres) that can be identified by the co-occurrence of linguistic features (e.g., the subjunctive mood, progressive aspect, present tense, etc.). Travis (2007) finds genre to be a significant factor in the rate of first person subject expression in spoken Spanish, as well as in the duration of the priming effect.

Most studies of information flow have focused on spoken language, with different genres represented in quantitative analysis. The bulk of the literature has focused on narratives and monologic speech (e.g., Ashby & Bentivoglio 1993, 1997; Du Bois 1987; Durie 2003; Givón 1983, *inter alia*), although some have examined conversation (e.g., Ewing 2005; Kärkkäinen 1996; Kumpf 1992; Matsumoto 2003), and other genres (Kumpf 2003; O’Dowd 1990).

Oral narratives and conversations are both genres of spoken language and are similar in that they represent naturally produced, unplanned speech. There are marked differences between the two genres, however¹¹. Narratives tend to be largely monologic, with a focus on relating one person’s experience, and there is greater continuity of subject in narratives than in conversations (cf. Travis 2007:130). Conversations are multi-party, and speakers are challenged with simultaneously balancing cognitive (i.e., information flow) pressures with social/interactional pressures (i.e., turn-taking).

Another way to think about the differences between the genres is in terms of information flow pressures. Both Du Bois (1987) and Durie (2003) discuss information flow pressure (i.e., the greater the number of referents, the higher the information flow pressures). Du Bois (1987:835) predicts that in texts with higher information flow

¹¹ It should be noted that the terms ‘narrative’ and ‘conversation’ are used here to describe the two genres, but that there are different types of narratives and different types of conversations as well. The narratives of the Pear Film used in this study are one type of narrative, and may not be representative of other

pressures (i.e., narratives), there will be a stricter adherence to Preferred Argument Structure (as evidenced in the alignment of S and O roles for containing lexical mentions), and that the proportion of intransitive predicates (i.e., NPs in the S role) used for introducing new mentions will be higher than in texts with lower information pressures (i.e., conversations). He proposes two measures of information pressures: the ratio of new humans to clauses, and the ratio of new entities of any kind to clauses (or IUs) (Du Bois 1987:834). Durie (2003) finds that, in Acehnese, the greater the information pressure, the greater the tendency for the S role to fragment, i.e., to pattern in some cases like O (object) and sometimes like A (subject of a transitive verb), rather than show a more ergative alignment (pattern like O), as predicted by Du Bois (1987).

A few studies (Kärkkäinen 1996; Kumpf 2003; O'Dowd 1990) do look specifically at genre, and have tested Du Bois' hypotheses. O'Dowd (1990) studied PAS in the high-information flow pressure genre of an orally-delivered paramedic training session. She finds some of the principles of PAS do apply (e.g., avoid lexical A, one lexical argument per clause), but that in this genre S does not align with O in introducing new information, but it does align with O in maintaining topic continuity. Kärkkäinen (2003) looked at PAS in a low information pressure genre, American English conversation. Like O'Dowd, she found that the S and A role pattern similarly (non-lexical NPs, given information) in non-narrative portions of conversation, and that new information is generally introduced in the O or Oblique roles. She concludes that her observations of the patterns of American English are not due to genre differences, rather that there is no evidence of ergative patterning in nominative/accusative languages.

narratives (i.e., personal narratives described in Labov's work (cf. Labov To Appear)), and conversations can include various different genres.

Kumpf (2003) noted conformity with the principles of PAS in teacher's classroom discourse. She also finds in her data that free (unattached) NPs can be used for introducing new information, although the preferred information status of free NPs appears to be given (85%) in her data.

In sum, there is substantial evidence that PAS principles (one lexical argument, avoid lexical A, etc.) are fairly universal, but that the S role is subject to different pressures, including information pressure¹² and, perhaps, genre. This study will contribute to this body of literature by direct comparison of two different genres, Spanish language, with the exclusion of non-referential NPs.

2.5 Summary of research questions pertinent to information flow

To recapitulate, while there has been extensive research on how information flow affects NP form and distribution, gaps in the literature remain. This study examines how the different information flow properties of an NP are linguistically manifested. Specifically, one important contribution of this study is the methodological importance of distinguishing between referential NPs and the different types of non-referential NPs, in order to create an appropriate database of NPs among which to investigate information status, as the inclusion in a study of the differences among given and new NPs of NPs whose form is not affected by information flow results in a skewed picture of the results.

Among those NPs that have forms subject to the influence of information status, this study seeks to see how they are treated differently, what general tendencies are cross-

¹² See also Dutra (1987) for an account of the split S role in Brazilian Portuguese, where she reports that the preverbal S aligns with A (also strongly preverbal) in having a strong tendency to code given information and postverbal S aligns with O in both word order and preferred information status (given).

genre-neutral and which are genre specific, and how identifiability is evidenced in the treatment of these NPs.

This study also investigates the factors that contribute to the form and distribution of NPs which are not affected by information status (i.e., non-referential NPs), and seeks to provide a better understanding of the discourse roles performed by these NPs.

This then allows us to better understand the different dimensions of information flow and how they each affect the grammar of NPs. While this case study of the different issues related to information flow provides a much needed synthesis of information flow factors, it was found that information flow factors alone are insufficient for accounting for the form and distribution of NPs in spoken Spanish. The next section introduces other influential factors that have often been overlooked in the literature—the role of dialogic interaction, as illustrated in (14) below, where information flow fails to account for the realization of the NP *Daisy*.

(14) NP realization not explained by information flow

J: solo está la Daisy,
Ø te das cuenta?
R: .. solo **la Daisy** viene.
J: just Daisy is there,
do (you) realize?
R: .. just **Daisy** is coming.

(Jugo: 516-518)

2.6 Other factors influencing NP form

While much research has been done on the cognitive constraints that influence NP form, less research has focused on other factors that may influence NP form. Fox (1987) points out that when a stretch of speech is examined, not all instances of NP form can be accounted for strictly in terms of information flow or referential distance. In fact, there

are often NPs whose form is markedly different than what would be predicted by information flow, and she suggests how speakers may be using these devices for interactive means.

One of the main difference between Fox's (1987) work on anaphoric coding and traditional (e.g., Givón 1983) perspectives lies in the conceptualization of discourse units. Some researchers (e.g., Chafe 1987, Du Bois 1987, Givón 1983, Croft 1995) take the stance that the basic unit of discourse is the clause or the intonation unit. Fox (1987), following the tradition of Conversation Analysis, embarks from a different point of reference for the discourse unit—the adjacency pair¹³ (cf. Levinson 1983). In her study she finds the following: 'The first mention of a referent in a sequence is done with a full NP. After that, by using a pronoun the speaker displays an understanding that the proceeding sequence has not been closed down' (Fox 1987:18). In addition, full NPs are not found only for the first mention of a referent, but they are also used to signal that a preceding sequence has been closed. Fox also reports non-structural uses of full, non-new NPs, including disagreements, assessments, and referent negotiation.

Another important contribution to the literature on interactional factors that affect NP distribution is the role of increments. Increments are 'constituents that are added to turns that, a just prior point, are interpretable as possibly complete syntactically and prosodically' (Ford, Fox and Thompson (2002:15-16), cf. Schegloff 1996). Building on previous work on turn-taking (cf. Sacks et al. 1974 and Ford & Thompson 1996), Ford et al. 2002 find that speakers resolve speaker change problems by adding increments that syntactically follow from the previous utterance, or in other cases, unattached NPs that

¹³ Adjacency pairs are not found in all genres; they are a product of dialogue and are absent in monologic data.

are syntactically isolated and provide an evaluation of a previous utterance. This is typically done through full NPs. This shows how syntactic role of an NP can be manipulated by speakers to accomplish ‘interactional work’.

Helasvuo (2001) has also contributed to the literature on NP distribution and interactional factors, including the interactive functions of NPs in list construction. Jefferson (1990) first drew attention to the occurrence of list-construction in conversation, noting that lists tend to have three parts, which tend to be structurally similar (e.g. three NPs, three infinitival verbs); Sánchez Ayala (2003) confirmed the similarities in list construction for Spanish. This implies that while lists are under construction, participants in the conversation orient to the three-part structurally similar list and can project, to a certain extent, what will come next. This projectability also allows for the co-construction of by multiple parties in the conversation. It has also been noted that each item in a list tends to occupy its own intonation unit (see Sánchez Ayala 2003 for a prosodic analysis of lists in English and Spanish). Helasvuo (2001) also observes that full NPs in predicate nominal constructions provide an opportune place for participants to co-construct the conversation, however this is not a very common occurrence in her data. Dumont (2006) finds that the same full NP is sometimes repeated in two consecutive clauses in Spanish interview data, which is linked to two interactional factors—repetition and contrast.

Given these findings, an examination of how interactional factors such as speaker orientation to discourse structure, turn-management, referent negotiation, list-construction, co-constructions, and ‘stylistic’ or uses of NPs such as disagreement, repetition, and contrast should not be excluded from a comprehensive analysis of NP

form and distribution. While the broad patterns of NP form and distribution may be attributed to the cognitive constraints, there are deviations from patterns predicted by information flow, and these may be understood by examining the more immediate and local interactional factors that speakers are navigating.

2.7 Summary

The following research questions have been developed after taking into account the research mentioned above.

- What are the different and ‘competing’ factors that influence NP form and distribution?
- Does our understanding of the effects of information flow change when NPs are examined according to referentiality?
- How do information flow constraints show different manifestations under different genres (i.e., different levels of information flow)?
- What is the role of interaction in the grammar of conversation?

Chapter 4 discussing the coding and Chapter 5 examines the methodological question of how to distinguish between referential and non-referential NPs. Chapter 6 examines the form and distribution of referential NPs, with special attention to the affects of information flow and information pressures. Chapter 7 shows how non-referential NPs respond to different discourse pressures than referential NPs, and how this is reflected in their form and distribution. Interactional factors that ‘compete’ with information flow factors in affecting NP form and distribution are discussed in Chapter 8.

3 Data

3.1 Data collection

The data for this investigation come from a large corpus of Ecuadorian Spanish that I collected in Quito between June and August 2006. I worked with an Ecuadorian colleague and three local assistants to collect 35 narrations of the Pear Film (Chafe 1980), 31 spontaneous conversations, and 12 sociolinguistic interviews. All speakers are middle class, high school or college educated, and range in age from 18 to 84.

Two genres are included in this investigation: the elicited monologic narratives and spontaneous conversations. Each genre provides its own advantages to the study of NPs, and the inclusion of both genres allows for comparison between the two. In one genre, the narratives, the information pressures are high due to the large number of new referents that are introduced over a relatively short span of speech (cf. Du Bois 1987), yet the social/interactional dynamic is low, as one speaker has the floor and there is no competition for taking a turn at talk. Therefore, the study of narratives allows for an in-depth analysis of how speakers maximize syntactic resources while introducing and tracking referents throughout a narrative. In the other genre, conversation, the information pressure is low, yet the social/interactional dynamic is high. A comparison between the two genres provides an ideal circumstance in which to examine the presumably differing effects that information flow and interactional concerns have upon NP realization and distribution.

3.2 Pear Film Narratives

The Pear Film is a six minute film that was made at the University of California at Berkeley in 1975. The film begins with a man picking pears in a rural setting. A man and a goat pass by, and then a young boy on a bicycle approaches and takes one of the baskets of the pears that the man has collected. He rides away, passes a young girl on a bike, loses his hat, hits a rock while distracted and topples over, spilling the basket of pears. Three young boys come into the scene to help the boy with the bicycle pick up the pears and load the basket back onto the bike. The boy starts to continue on his way, when one of the boys whistles to him to come retrieve his hat. He collects his hat from the boy and gives him three pears in return. The three boys set off in the direction of the man who was picking pears, who has just noticed that he is missing one of the baskets. The boys pass by the man, eating the pears, at which point the film ends.

The Pear Film was designed as a tool for linguists to elicit narrations of the same experience from different speakers. There is no dialogue in the film, but other sounds were included. It has been used to collect data from many different languages (cf. Chafe 1980) and the narratives have been used to investigate many linguistic phenomena.

Fifteen narrations of the Pear Story were selected for this investigation, from eight male speakers and seven female speakers. Each recording lasts between one minute, thirty seconds and five minutes, eight seconds, with most lasting between two to three minutes. Details of the narrations are presented in Table 1.

Table 1: Pear Stories

PS #	Speaker	Age	Duration (minutes.seconds)	# of words	# of IUs
12	male	23	2.58	552	139
13	male	50	2.05	287	66
14	male	19	1.36	271	77
29	female	25	1.58	338	81
30	male	27	1.54	375	101
33	female	30	1.42	291	58
34	male	23	1.44	239	57
36	female	43	2.24	441	99
38	male	58	2.23	404	105
40	female	33	2.38	539	130
41	male	37	1.30	277	53
42	male	66	3.02	379	82
43	female	26	3.02	542	122
52	female	57	3.26	434	97
100	female	26	5.08	1061	246
Totals			37.30	6430	1513

The participants viewed the Pear Film on a laptop computer with headphones, and after watching the movie were instructed to narrate to one of the investigators what they had seen in detail, as if they were re-telling the story to a friend who had not seen the film. Each digital recording was made in the presence of one of the corpus investigators, but interaction from the investigator was kept to a minimum. In a few cases there are a minimal number of backchannels produced by the investigator during the narration (saying *mhm*, laughing, etc.), but aside from this the narrations are entirely monologic.

3.3 Spontaneous conversations

From the spontaneous conversations, fifteen participants were also selected for this investigation, 7 males and 8 females, and 250 NPs were extracted from each speaker's speech. Three assistants were hired in Ecuador to make digital recordings of routine conversations with their spouses, families, friends and co-workers. A brief

description of the participants and conversations follows; the bolding of a participant's name indicates that their speech from that recording was used in this investigation¹⁴. A summary of the conversational data is given in Table 2.

Clases (12 minutes, 40 secs)

Participants- **Amanda** (26 years old) and **Carolina** (21 years old). In this conversation, these two sisters talk about various topics including Carolina's university classes, digital cameras, laptops, and other computer related items.

Birthday (14 minutes, 46 seconds)

Participants- **Amanda** (26 years old) and **Francisco** (28 years old). This husband and wife talk about their jobs, their nephew's upcoming birthday and raising children in this conversation.

Jugo (25 minutes, 8 seconds)

Participants- **Rodrigo** (19 years old) and **Julio** (19 years old). This conversation takes place in Rodrigo's home, and the friends discuss a soccer game they played earlier in the day, gossip about various friends, and talk about Julio's job.

Café (1 hour, 28 minutes, 52 seconds)

Participants- Rodrigo (19 years old), Rodrigo's parents (50 year old **Gerardo** and 46 year old **Paca**), and Rodrigo's maternal great-grandmother, **Señora Manuelita** (84

¹⁴ Tokens were extracted from one conversation for speakers, except for Amanda, whose sample includes tokens from two conversations because she did not produce enough NPs in *Clases* for the analysis. Of the 250 NPs for this speaker, 217 are from *Clases*, the remaining 33 from *Birthday*.

years old). In this four-party conversation, participants come and go as they catch up on each other's lives, discuss future plans, and talk about other family members.

Chicas (17 minutes, 58 seconds)

Participants- Rodrigo (19 years old), **Adán** (19 years old) and Martín (18 years old). In this three-party conversation, Adán tells his friends about the domestic help his mother has recently hired, and the three discuss the upcoming birth of Martín's child and his relationship with the mother of the child.

Fumar (24 minutes, 2 seconds)

Participants- Rodrigo (19 years old) and **Adrián** (18 years old). In this conversation, the friends discuss Adrian's desire to quit smoking, they gossip about other friends, and they reminisce about growing up together in the same neighborhood.

Viejos Amigos (40 minutes, 3 seconds)

Participants- Rodrigo (19 years old) and **Eva** (19 years old). These friends discuss their friendship, other friends and family members, and Eva's classes in this dialogue.

Hermanos (1 hour, 13 minutes, 45 seconds)

Participants- Gerardo (50 years old) and his 57 year old sister, **Beatriz**. The siblings talk about a wide range of topics, including Beatriz's recent foot injury, the death of their parents, family disputes, and their children.

Deportes (14 minutes, 56 seconds)

Participants- Francisco (28 years old) and **Alicia** (23 years old). These co-workers discuss the weekend, possible plans to travel together, and soccer in this conversation.

Food (24 minutes, 55 seconds)

Participants- **Lucy** (28 years old) and **Eduardo** (30 years old). In this conversation, this married couple discusses the food options at their respective workplaces, Lucy's stomach ailments, and mutual friends.

Table 2: Summary of conversational data¹⁵

Conversation	Participants included in this study	Total N of participants	Duration (hours.minutes.seconds)	# of words	# of IUs
Clases	Amanda, Carolina	2	12.40	2814	754
Birthday	Amanda, Francisco	2	14.46	3006	762
Jugo	Rodrigo, Julio	2	25.08	5225	1425
Café	Gerardo, Paca, Señora Manuelita	4	1.28.52	19,031	4941
Chicas	Adán	3	17.58	4603	1335
Fumar	Adrián	2	24.02	5139	1261
Viejos Amigos	Eva	2	40.03	9069	2459
Hermanos	Beatriz	2	1.13.45	17,535	4684
Deportes	Alicia	2	14.56	3765	933
Food	Lucy, Eduardo	2	24.55	4486	1107
Totals	15 participants		5.57.05	74,673	19,661

¹⁵ The duration, number of words and number of IUs is given for each transcript. In most cases, the sample from which the NPs in this study were taken did not include the entire transcript.

3.4 Transcription method

All data were transcribed according to the conventions of Du Bois, Schuetze-Coburn, Cumming and Paolino (1993). In this method, each intonation unit (IU) occupies its own line in the transcription, and a number of production and interactional features (pauses, laughter, overlap, inhalation, etc.) are captured in the transcription. An IU is identifiable as 'a stretch of speech uttered under a single coherent intonation contour' (Du Bois et al. 1993: 47). The transitional continuity of the IU, that is, 'the marking of the degree of continuity that occurs between one intonation unit and the next' (Du Bois et al. 1993: 53) can be realized as continuing, final, appeal, or truncated, and is noted for each IU. It is important to note that final intonation does not always occur at the end of a clause or the end of a turn, continuing intonation does not necessarily mean that the speaker will continue, appeal intonation is not exclusively reserved for questions, and truncated IUs will not always be followed by a completion or repair of the previously truncated IU. In other words, IUs are defined prosodically, not functionally, although there is some functional correlation.

According to Chafe (1994:69), 'Intonation units are hypothesized to be the linguistic expression of information that is, at first, active in the consciousness of the speaker and then, by the utterance of the intonation unit, in the consciousness of the listener.' IUs can range from small, one or two word units such as backchannels (*mhm*, *oh*), discourse markers (*pues* 'well'), and free NPs (as illustrated in Chapter 1), to clauses and parts of clauses. Croft (1995) finds for English that nearly all intonation units are also grammatical units (clauses and phrases), but that not all grammatical units are intonation units (that is, an IU can be comprised of more than one grammatical unit). Chafe (1994)

notes a correlation between substantive IUs and clauses in English¹⁶, but also points out that substantive IUs can also be comprised of just a referent (especially free NPs). A number of linguists (Chafe, 1994; Croft 1995; Du Bois 1987:812; *inter alia*) would argue that IUs constitute the basic unit of the grammar of spoken speech. Chafe (1994) proposes a constraint of one new idea per IU, linking the IU to memory, cognition and information flow. By studying how speech and grammatical structures are distributed by speakers across different IUs, we come closer to understanding what constitutes a meaningful 'unit of speech' to a speaker at the time that an utterance is constructed. Appendix I provides a summary of the symbols used in the transcription.

The Pear Stories were transcribed by myself, and five were checked for accuracy by a native speaker of Ecuadorian Spanish. The conversations were all roughly transcribed by local (Ecuadorian) assistants, in most cases the project assistants themselves, then checked by myself, sent back to a native speaker (most often a participant in the conversation) for clarification, and finally transcribed in the final format by myself.

¹⁶ See, however, Matsumoto (2003) for a discussion of the correlation between clauses and IUs in spoken Japanese.

4 Coding

All 3rd person NPs, pronominal, nominal and unexpressed mentions¹⁷, in all syntactic roles were extracted from the 15 Pear Story narrations, yielding 1565 NPs, and the first 250 NPs/third person verbs produced by each of the 15 speakers in the spontaneous conversations (3750 tokens) were extracted.

Excluded from the analysis were: truncated NPs and those NPs for which syntactic role cannot be determined due to truncation of the IU, subjects of relative clauses that are coreferential with the head¹⁸, headless relative clauses, NPs embedded within other NPs, infinitival verbs, (partially) unintelligible NPs, 3rd person verbs with impersonal subjects, and fixed clause length phrases such as *Dios le pague* 'thank you' (where neither *Dios* nor *le* were included) and *que el Señor te bendiga* 'may God bless you'). After these exclusions 4418 (1228 from the narratives and 3190 from the conversational data) tokens remained for analysis and were coded for the factors described in the following sections. The excerpts in (15) and (16) illustrate exclusions and the discussion that follows will clarify which tokens were included or not included in this study and why. The NPs in italics were not included, the bolded NPs were¹⁹.

¹⁷ Unexpressed tokens were only extracted for subjects and were identified through the verbal morphology.

¹⁸ These were not included in the study because their expression is not variable in Spanish, they are obligatorily unexpressed.

¹⁹ Remember that NPs embedded within a larger NP were excluded. In the case of (52) below, *un sombrero. que (53) le pertenecía (54) al niño, que -- .. que sustrajo (55) la -- (56) la canasta*, '(52) a hat. that (53) (le) belonged to (54) the boy, that --.. that stole (55) the --(56) the basket', the entire NP *un sombrero... la canasta* (a hat... the basket' is included, but the NPs embedded within the relative clause (*le, el niño, la canasta*) are excluded.

(15) Excerpt 1

J: ... (2.0) (1) ∅ entro en (2) tu [<X cocina X>]?
R: [(3) ∅ estoy] con (4) el Julio.
... (5) qué?
J: y (6) otra vez (7) me toca (8) firmar (9) la hoja?
R: .. no,
ya,
solo (10) una vez.
J: .. ah,
fresco.
.. (11) esa es.
... [qué (12) chévere <X grabadora X>].
R: [siéntate,
(13) Julio].
J: (14) Dios (15) te pague.
R: (16) ∅ quieres (17) juguito?
J: .. bueno.
.. gracias (18) Rodriguito.
(19) se (20) te agradece con (21) toda el alma.
R: ... (2.0) toma (22) ∅.
J: .. gra=cias,
@@
(23) qué más,
(24) Rodriguito?
qué tal --
(25) qué hiciste (26) ∅ hoy?
R: ... (H) hoy --
(27) ∅ estaba pasando (28) las .. conversaciones.
... (2.0) (H) y ya.
(29) tú?
J: cómo es (30) ∅?
como (31) un diálogo?
R: .. sí,
más o menos.
J: .. (32) o sea,
(33) tú pones (34) Julio dice (35) esto,
(36) Roddy dice (37) esto.
.. (38) Junior dice (39) esto.
... (40) Junior es vago.

J: ... (2.0) (1) do I go into (2) your [<x kitchen x>]?
R: [(3) I'm with (4) Julio].
... (5) what?
J: and (6) this time do (7) I (8) have to sign (9) the paper?
R: ..no,
yeah,
just (10) one time.
J: .. ah,
cool.
.. (11) that's it.
... [what (12) a cool <X recorder X>].
R: [sit down,
(13) Julio].
J: (14) God pay (15) you.
R: do (16) you want (17) juice?
J: .. ok.

.. thanks (18) Rodriguito.
 (19) (20) you are appreciated with (21) all my heart.
 R: ... (2.0), (22) \emptyset take.
 J: .. thanks,
 @@
 (23) what else,
 (24) Rodriguito?
 how -
 (25) what did (26) you do today?
 R: ... (H) today -
 (27) I was transcribing (28) the .. conversations.
 ... (2.0) (H) yeah.
 (29) you?
 J: how is (30) it?
 like (31) a dialogue?
 R: .. yes,
 more or less.
 J: .. (32) so,
 (33) you put (34) Julio says (35) this,
 (36) Roddy says (36) this,
 .. (38) Junior says (39) this.
 ... (40) Junior is lazy.

(Jugo:59-96)

(16) Excerpt 2

es (41) lo que hizo (42) el niño,
 (43) \emptyset siguió,
 ... pero,
 ... (44) \emptyset siguió (45) su --
 (46) su camino,
 no?
 (H) (47) los tres niños,
 .. también comenzaron a --
 a --
 (H) a regre- --
 a regresar.
 por donde (48) el ca- --
 por donde (49) el niño había venido.
 (H) y (50) \emptyset encontraron (51) un= --
 (52) un sombrero.
que (53) le pertenecía (54) al niño,
que --
.. que sustrajo (55) la --
(56) la canasta,
 no?
 Entonces,
 (57) los niños,
 al encontrar (58) el sombrero,
 ... (59) lo --
 (60) lo recogieron,
 y (61) le --
 y (62) \emptyset (63) le llamaron (64) la atención (65) al niño,
que ya estaba --

.. ya estaba (66) algunos metros .. lejos de (67) ellos.
Entonces (68) ellos (69) le devolvieron muy amablemente (70) su sombrero,

'that's (41) what (42) the boy did,
(43) (he) went,

... but,
... (44) (he) went on (45) his -
(46) his way,

no?

(H) (47) the three boys,

.. also started to --

to --

(H) to re- --

to return.

from where (48) the ro- --

from where (49) the boy had come.

(H) and (50) (they) found (51) a= --

(52) a hat.

that (53) (le) belonged to (54) the boy,

that --

.. that stole (55) the --

(56) the basket,

no?

So,

(57) the boys,

upon finding (58) the hat,

... (59) (it) --

took (60) it,

and (61) (le) --

and (62) (they) (63) (le) caught (64) the attention of (65) the boy,

who was already --

.. was already (66) a few meters .. away from (67) them.

So (68) they very nicely returned (69) his hat to (70) him,

(PS12:74-103²⁰)

Included in the analysis were:

- full NPs ((4), (6), (9), (10), (13), (17), (18), (21), (24), (28), (31), (34), (36), (38), (40), (46), (47), (49), (52), (57), (58), (64), (65) and (70));
- pronominal NPs (including (60), (63), (68) and (69));
- unexpressed NPs, such as (43), (44), (50) and (62);

²⁰ PS stands for 'Pear Story', the number refers to the speaker, and the number following the colon refers to the line number in the transcription. Examples taken from the conversational portion of the corpus include the title of the conversation and transcript line number(s) for that example.

- interrogative pronouns, see (5), (25);
- demonstrative pronouns, i.e., (11), (35), (37) and (39);
- and other pronouns (such as *algo* ‘something’, *alguien* ‘someone’, etc., not illustrated above).

Excluded from this study were:

- first and second person NPs ((1), (3) (7), (15), (16), (20), (26), (27), (29), and (33) above);
- NPs that were unclear to the transcriber, as in (2) and (12);
- infinitival subjects, such as (8);
- impersonal subjects, for example, (19) and (30);
- NPs occurring in set phrases and discourse markers ((14), (22), (32)),
- headless relative clauses (41);
- truncated NPs (see (45), (48), (51), (55), (59) and (61);
- and NPs embedded within other NPs (42), (53), (54), (56), (66) and (67).

4.1 Syntactic and semantic factors

4.1.1 NP realization

Referent realization is often said to be related to information flow (Ariel 1990; Ashby & Bentivoglio 1993; Du Bois 1987, 2003; Givón 1983; Prince 1981). More specifically, it has been proposed that there is a link between the availability/accessibility of a referent and the realization – more available/accessible referents show a tendency toward being realized pronominally or as unexpressed mentions, and conversely, less available/accessible referents tend towards showing more explicit marking (full NPs, cleft and focus constructions, etc.). In other studies, referent realization has also been

linked to interactional factors (Ford & Fox 1996; Fox 1987; Helasvuo 2001a). Seven different types of NP realization were coded in the data²¹:

- full NP²²
- personal pronoun (subject *él* 'he' and *ella* 'she', direct object *lo*, *la*, *le*, *los*, *las*, *les*²³ and indirect object *le* and *les*)²⁴, demonstrative pronoun (*eso*, *esa*, *ése*, *esos*, etc.), other pronoun (*algunos* 'some', *otros* 'others', *alguien* 'someone', *todos* 'all, everyone', *nada* 'nothing', *una* 'one', etc.), unexpressed mentions, interrogative pronouns (*qué* 'what', *quién* 'who', etc.).

4.1.2 Modification

In order to examine patterns of information flow and referent modification, all full NPs were coded for type of determiner (none (bare), definite article, indefinite article, demonstrative adjective, descriptive adjective, possessive adjective, quantifier, interrogative adjectives), and whether or not the NP was modified by a relative clause.

Definiteness in Spanish (i.e., modification by a definite article, demonstrative adjective, possessive adjective or quantifier) is thought to be associated with identifiability of a referent (cf. Bosque 1989; Ortiz Ciscomani forthcoming), indefiniteness with newness and unidentifiability (Garachana Camarero, forthcoming), while lack of modification is associated with non-referentiality (Torres-Cacoullós &

²¹ Three additional realizations were found but excluded: subjectless phrases (*ø es que* 'it's that', *ø es cierto* 'it's true'), headless relative clauses (*lo que ...*), and truncated NPs.

²² This includes proper names, which, like other full NPs, contrast with pronouns and unexpressed mentions. Other researchers (e.g., Prince (1981) and Ewing (2005)) have pointed out that proper names often differ from other full NPs with respect to information flow properties (accessibility, referentiality, etc.); these properties are captured in other aspects of the coding for this study.

²³ Ecuadorian Spanish is a *léista* variety in which direct object pronouns often take the forms *le* and *les*, which are only used as indirect objects in other varieties.

²⁴ There are different ways of coding for these factors. One way which may be fruitful for future studies would be to separate the object pronouns that occur as part of a construction (particularly when they appear with the full NP) from those which do not.

Aaron 2003). Du Bois (1980) has shown a similar relationship between the use of definite and indefinite markers and identifiability in English. Other uses of modifiers not associated with information flow are explored more thoroughly in Chapter 6.

Fox and Thompson (1990) show that in American English conversation there are distinct patterns in the distribution of grammatical roles of NPs that are heads of relative clauses and the grammatical role of the coreferent within the relative clause²⁵, which they attribute to speakers' attention to information flow. While NPs embedded within relative clauses were not extracted for this study, it was noted which NPs are heads of relative clauses in order to better examine their role in managing information flow. An example of an NP modified by a relative clause follows in (17).

(17) NP modified by relative clause

pero decidió al final,
coger **todo el canasto que estaba lleno.**

'but (he) finally decided,
to take **the whole basket that was full.'**

(PS12:41)

4.1.3 *Syntactic role*

Numerous studies have pointed out the relationship between syntactic role and information flow (Ashby & Bentivoglio 1993; Du Bois 1987, 2003; Du Bois, Kumpf & Ashby 2003). Studies of Preferred Argument Structure (i.e., Ashby & Bentivoglio 1993; Du Bois 1987, 2003) indicate that speakers tend to limit the amount of new information in a clause to just one new argument per clause, and that new information is most often found in the S (subject of intransitive verb) or O (object) roles. It has also been

²⁵ For example, in the sentence 'probably the only thing you'll see is like the table' from Fox and Thompson (1990:302), the NP 'the only thing' is the subject of the main clause, but the object of the relative clause.

consistently found that referents in the A role are almost exclusively given, and therefore generally unexpressed.

On the other hand, there have also been studies demonstrating links between particular syntactic role and interactive tasks not related to information flow (Ford, Fox & Thompson 2001; Helasvuo 2001a). Eleven different syntactic roles were distinguished. Core arguments include subject of transitive verb, subject of intransitive verb²⁶ (which includes single arguments of verbs that are not truly transitive, such as (18) and (19)), direct object, indirect object; non-core arguments include object of preposition, adverbial, predicate nominal, and vocative. A separate code was given to those interrogative pronouns that do not fulfill a clear argument function, but whose role is merely grammatical (see (20))²⁷. Another code was created for those NPs which are an extension of a previous syntactic frame, as in (21), where *una parte (en el aeropuerto)* is interpreted as directly following from the verb *ver* (which has a separately expressed direct object), as opposed to being a syntactically isolated (free) NP. Lastly, there were a number of NPs that did not fit into any other the above roles, including appositives as in (22) and syntactically isolated ('free') NPs as in (23), which were put into a heterogeneous category of 'Other'.

(18) Single argument of *haber*

siempre hay **alguien** que nos va a ayudar.
'there is always **someone** that is going to help us out.'

(PS12:63)

²⁶ Note that all subjects of 'Complement-taking predicates', i.e. *decir* 'say', *pensar* 'think', *ver* 'see' were coded as subjects of intransitive verbs (cf. Thompson 2002).

²⁷ In traditional analyses, many of these instances would be assigned traditional syntactic roles, i.e., direct object for (20). For the purposes of this study, they were given a separate code because the discourse functions interrogative pronouns perform are quite different from discourse functions of direct objects that are directly related to the semantics of the verb.

(19) Single argument of *hacer*

pero está haciendo **basta=nte so=l**.
'but it is really sunny [lit. it makes a lot of sun].'

(Clases:7)

(20) Interrogative/grammatical role

qué dijo tu mami ayer,
'**what** did your mom say yesterday, '

(Jugo:386)

(21) Extension of previous syntactic frame

S: .. vi una parte,
 en mi casa,
F: ya.
S: de ahí **una parte** en el aeropuerto,

S: .. I saw part [of the soccer game],
 at my house,
F: yeah.
S: and then **another part** at the airport,

(Deportes:83- 86)

(22) Appositive

(H) pasaron los tres niños,
**los que le ayudaron al niño a recoger las --
las peras,**

'(H) the three boys passed by,
the ones that helped the boy collect the --
the pears, '

(PS36:56-58)

(23) Syntactically isolated NP

pero a la una,
de una a dos hay supletorio.
.. **supletorio de productos del banco.**

'but at one o'clock,
from one to two there is a make-up exam.
.. **make-up of [the exam on] banking products.** '

(Birthday:158-60)

4.1.4 Semantic class of verb

There are patterns between different types of verbs and different discourse functions. Motion verbs, for example, have been shown to often introduce new referents (Bentivoglio 1993; Dumont 2006). Scheibman (2001) shows that relational predicates perform specific discourse functions. In her study of American English, a large number of third person singular subjects are nonhuman referents that occur with relational predicates (mostly copular clauses, e.g., ‘that is a good deal’), which permit 'speakers to index their opinions and attitudes' (2001:85). Accordingly, in this study all subjects and objects were coded for the semantic class of the verb, which were categorized as follows: copula; motion; cognitive process; perception; speech act; 'other'; *haber* ('there is'); and stative, impersonal and experiencer verbs. Table 3 exemplifies the kinds of verbs found in each class.

Table 3: Verb classes

Class	Examples
Copula	<i>estar</i> ('to be'), <i>parecer</i> ('to seem'), <i>ser</i> ('to be')
Motion	<i>andar</i> ('to go'), <i>bajar(se)</i> ('to go down'), <i>entrar</i> ('to enter'), <i>ir</i> ('to go'), <i>pasar</i> ('to pass'), <i>regresar</i> ('to go back'), <i>subir(se)</i> ('to go up'), <i>venir</i> ('to come'), <i>volver</i> ('to return'), etc.
Cognitive process	<i>aprender</i> ('to learn'), <i>creer</i> ('to believe'), <i>darse cuenta</i> ('to realize'), <i>decidir</i> ('to decide'), <i>entender</i> ('to understand'), <i>enterarse</i> ('to find out'), <i>pensar</i> ('to think'), <i>recordar</i> ('to remember'), <i>saber</i> ('to know'), etc.
Perception	<i>escuchar</i> ('to listen'), <i>mirar</i> ('to watch'), <i>observar</i> ('to observe'), <i>oír</i> ('to hear'), <i>oler</i> ('to smell'), <i>sentir</i> ('to feel'), <i>ver</i> ('to see'), etc.
Speech Act	<i>decir</i> ('to say'), <i>agradecer</i> ('to appreciate'), <i>pedir</i> ('to ask for'), <i>sugerir</i> ('to suggest'), etc.
Other	<i>abrir</i> ('to open'), <i>dar</i> ('to give'), <i>hablar</i> ('to speak'), <i>hacer</i> ('to make/do'), <i>poner</i> ('to put'), <i>recoger</i> ('to pick up'), <i>silbar</i> ('to whistle'), <i>tener</i> ('to have'), <i>terminar</i> ('to finish'), <i>tomar</i> ('to take'), <i>trabajar</i> ('to work'), <i>vivir</i> ('to live'), etc.
Haber	<i>haber</i> (existential verb, 'there is/are')
Other stative, impersonal and experiencer verbs	<i>encantar</i> ('to love'), <i>existir</i> ('to exist'), <i>gustar</i> ('to like'), <i>importar</i> ('to matter'), <i>llamarse</i> ('to be named'), <i>molestar</i> ('to bother'), <i>pertenecer</i> ('to pertain'), <i>quedar</i> ('to stay, remain'), <i>suced</i> ('to happen'), etc.

4.1.5 Animacy and grammatical number

All NPs were coded for animacy (human, non-human animate and inanimate) and grammatical number (singular or plural). Animacy is related to referent realization, as inanimate subjects are rarely expressed pronominally in Spanish, and is also correlated with syntactic role (animate subjects tend to be more agentive than inanimate subjects). In addition, the discourse profiles of animate and inanimate NPs are very different, as observed by Fox (1995) and Scheibman (2001), among others. Grammatical number proves useful in cross-tabulations of the results for identifying trends associated with specific characters in the narratives as well as for patterns associated with non-specific (impersonal) third person verbs in the conversation.

4.1.6 Summary

In sum, based on insights from previous studies of NPs in English, Spanish and other languages, the NPs in these data were coded for 6 syntactic and semantic factors (realization, modification, syntactic role, semantic class of verb, animacy and grammatical number), in order to investigate how the grammar of NPs reflects relationships between syntactic, semantic and other factors (specifically information flow and interactional factors, as described in the following sections).

4.2 Factors related to information flow

The effects of information flow on referent realization and distribution have been well documented (Ariel 1990; Ashby & Bentivoglio 1993; Du Bois 1987, 2003; Du Bois et al. 2003; Givón 1983; Prince 1981), as was discussed in Chapter 2. Many ways of measuring information flow and topic continuity have been proposed; the coding of factors related to information flow used in this study draws from several previous studies and attempts to provide a comprehensive yet objective manner of coding for different parameters of information flow in a way which can be reproduced in future studies.

4.2.1 Activation State

Different models of activation state have been proposed; perhaps the two most influential are those found in Chafe (1987) and Prince (1981)²⁸. In Chafe's (1987) model, recall that there is a three-way distinction between active, semi-active, and inactive referents. In Prince's (1981) paper, she discusses seven different activation states: brand new unanchored, brand new anchored, new-unused, noncontaining inferrable, containing inferrable, textually evoked and situationally evoked. However, because of the

difficulties that arise trying to objectively code for states such as 'accessible' or 'inferred', a binary distinction between new (first mention of a referent in the transcript) and given (NP which has previously been mentioned in the discourse) was used here²⁹. There is a subset of referents (general, non-referential referents, described in detail below) for which activation state does not apply. The following example, (24), taken from the beginning of one of the narrative transcripts, illustrates the first and second mention of a referent (*un campesino* 'a peasant'). The first mention is coded as new, the second (unexpressed) mention as given.

(24) New referent

```
.. Eh=,
estaba un campesino recogiendo= .. unas=,
eh,
manzanas,
de= su árbol.
ya  $\emptyset$  tenía .. (H) dos canastas llenas de manzanas,

.. Eh=,
there was a peasant gathering .. some,
eh,
apples.
from his tree.
(he) already had .. (H) two baskets full of apples,
```

(PS29:3-8)

4.2.2 *Specificity*

The distinction made here is between general and specific referents; general referents do not name a specific entity, as in *guitarra* in the example below, but specific referents do, such as *Tebi*, the subject of (25).

²⁸ But see also Prince (1992) for another model. One major addition to her revised schema is the distinction between *discourse* new and old NPs and *hearer* new and old.

²⁹ Many factors (distance from previous mention, previous knowledge, schema, and episode boundaries among others) have been proposed to influence the accessibility of a referent, and these are difficult to separate from one another and to code independently. See Ewing (2005:122-124) for a detailed discussion of why a binary distinction between given/new referents proves to be a practical and reliable way of coding activation status.

(25) Specific versus general referents

... pero **el Tebi** ya sabe tocar **guitarra**?

'... but **Tebi** already knows how to play **guitar**?

(Café:1699)

4.2.3 Distance from previous mention

It is often thought that the number of clauses between two mentions (Du Bois 1987:816; Givón 1983) can give insight into the different degrees of 'old' information in discourse. Specifically, it is hypothesized that the greater the distance between two (co-referential) NPs in discourse, the more likely a previously introduced referent will receive linguistic coding associated with less continuous/accessible referents.

The number of intervening clauses was noted for those tokens that had a coreferential mention within the preceding 10 clauses; tokens with a coreferential mention at a distance greater than 10 clauses after the previous mention were given the same code. In (26) there are five mentions (four unexpressed) of the same referent (Verónica), with no intervening clauses between them. In (27), there are three intervening clauses (marked by arrows in the margin) between two mentions of the referent 'he/him.'

(26) No intervening clauses

R: **la Verónica**?

A: ah=.

.. \emptyset iba corriendo y shu=,
 \emptyset se resbala,
y \emptyset se pega una matada,
y \emptyset se raspó la rodilla,

R: **Verónica**?

A: ah=.

.. (**she**) was running and shu=,
(**she**) slips,
and (**she**) takes a huge spill,
and (**she**) scratched her knee.

(Fumar:226-231)

(27) Three intervening clauses

- M: ... sus abuelitos también tienen que haberle apoyado mu=cho,
en fin,
→ no es cierto?
G: [cla=ro,
→ así es].
M: [la familia],
la familia unida,
→ que es lo mejor.
G: ... pero,
él,
qué va a seguir ahí?
- M: ... his grandparents also must have supported him a lot,
after all,
→ isn't that right?
G: [of course,
→ that's the way it is].
[the family],
the united family,
→ that's the best.
G: ...[but,
he,
what is (he) going to study there?

(Café:100-110)

4.2.4 Realization of previous mention of the referent

This variable was included in the study in order to investigate the relationship between activation state and referent form; my hypothesis was specifically that newer referents are coded with relatively longer forms, and given referents receive relatively less coding (i.e., a pronoun is more likely to be preceded by a coreferential mention as a full NP than as an unexpressed mention (cf. Givón 1983)).

Seven types of previous realization were found here: personal pronoun, demonstrative pronoun, other pronoun, unexpressed, truncated; a full NP that was the same as the target token (e.g. *el hombre ... el hombre*); and a full NP that was a different lexeme than the target token (e.g. *Sara ... mi hermana*), a kind of re-naming of the same referent.

4.2.5 Discourse referentiality

The importance of discourse referentiality, that is, whether a NP refers to an entity that plays an important role in the subsequent discourse, has been acknowledged by many scholars (Du Bois 1980; Givón 1981; Hopper & Thompson 1984). It is very clear that speakers treat referential NPs in a different manner than non-referential NPs in speech – both in terms of NP form and syntax. It is also clear that the discourse referentiality of an NP intersects with other dimensions of information flow. For example, it makes little sense to speak of the activation status of a non-referential NP. In example (28) below, the NP *profesora* (‘teacher’) is non-referential, occurring as a predicate nominal which is being used to classify a person. It is the first mention of *profesora* in this conversation, yet upon hearing *profesora*, the interlocutors will not create a new ‘file’ for this NP. Instead, *profesora* is information that is added to the information file about the subject NP (*ella*). Similarly, in (29) in the verb-object compound *tener vacaciones* there is no specific vacation that is activated in the minds of the participants.

(28) Non-referential NP

P: ... pero ella es **profesora**,

P: ... but she is a **teacher**.

(Café:853)

(29) Non-referential NP

C: .. tenemos [**vacación**].

A: [no,
pero si] tienen ca- --
no tienen **vacaciones** por dos meses,
o cómo es?

C: no,
tenemos **vacació=n** cada seis meses.

C: .. we have [**vacation**].

A: [no,
but if] you have ca- --
don't you have **vacation** for two months,
or how is it?
C: no,
we have **vacation** every six months.

(Clases:40-46)

However, discourse referentiality is also perhaps the most difficult parameter of information flow to operationalize, as the term has loosely been used in the literature to apply to different situations, and has not been well defined in a manner that permits the researcher to objectively code NPs for this factor in a way that can be replicated by future investigators. As Ewing (2005:149) points out, discourse referentiality involves speaker intention, and 'cannot be established by a single quantitative criterion such as, for example, number of mentions'.

One of the most satisfying descriptions of the differences between discourse referential and non-referential NPs is found in Du Bois and Thompson (1991), who use the terms *tracking* and *non-tracking*. Tracking referents are those which have 'continuity of identity' within the discourse (cf. Du Bois 1980); non-tracking NPs are those which fulfill a variety of other discourse functions, including predicating, orienting and classifying. Predicating NPs name activities, as illustrated in (29) above and (30) below, and in Spanish tend to occur in V-O compounds. Orienting NPs spatially or temporally situate another NP, as in (31), and are usually found as adverbials or as parts of prepositional phrases. Classifying NPs, as in (28) above and (32) below, are usually found as predicate nominals in Spanish.

(30) Predicating NP

F: se llevó unos para aprender **inglés**,
F: she got some to learn **English**,

(Birthday:260)

(31) Orienting NP

R: .. eso haces **todos los días**,

R: .. you do that **every day**,

(Jugo:879)

(32) Classifying NP

C: una laptop ya es **una computadora**.

C: a laptop is **a computer**.

(Clases:604)

Noun phrases were coded as discourse non-referential if they fulfilled one of the non-tracking discourse functions described by Du Bois and Thompson (1991) and illustrated above, as well as those which have a non-tracking grammatical role, such as interrogative pronouns and vocatives. The remainder were coded as discourse referential, except for a small number which were coded as unclear (n=2 from Pear Stories, n= 11 from conversations) because it was not apparent whether they fulfilled a non-tracking function or not.

4.2.6 Summary of factors related to information flow

This section described how NPs were coded for five factors related to information flow and why these factors are important. Three factors of these factors (activation state, distance from previous mention and realization of the previous mention) have to do with other mentions of the same referent in the surrounding discourse, and the inclusion of these factors in this study will help us understand how discourse at large shapes the grammar of a NP. The other two factors, specificity and discourse referentiality, are also known to influence coding of an NP and for this reason are included in this study.

4.3 Additional factors

It has been demonstrated (e.g., Fox 1987) that information flow alone cannot account for all instances of NP realization. For example, where theories of information flow may predict a full NP, Fox found that speakers may use pronouns to mark a continuation of a discourse topic, and even to return to a previous discourse topic. In addition, other studies find interactional functions for free NPs (Ford, Fox and Thompson 2002), such as marking the speaker's stance towards their speech. It is therefore reasonable to expect that interactional factors may help to explain the usage of NPs which do not conform to general patterns of information flow or syntactic integration.

While several studies have pointed to the relationship between interactional factors and NP usage, the scope of this effect is unknown, in part because it is often difficult to operationalize coding for so-called 'interactional' factors, and indeed difficult to define 'interactional factors.' For this study, the following factors which have been linked to interactional factors in the literature were coded for all NPs in the conversational data only, as this is where interactional concerns are more likely to be visible.

4.3.1 Increments

Following Ford, et al. (2002:16), NPs occurring in a 'non-main clause continuation after a possible point of turn completion' are coded as increments, as in (33). Increments can occur in different syntactic roles, commonly as objects of prepositions and NPs in the 'Other' syntactic role ('free NPs' in Ford et al.'s (2002) terms).

(33) Increment

el semestre se divide en tres bimestres.
.. **exámenes** cada bimestre.

'the semester is divided into three trimesters.
.. **exams** each trimester.'

(Caro:64-65)

4.3.2 *Repetition*

Tannen (1987:583-84) discusses many interactional functions of repetitions; among them strategies for managing turn-taking, showing listenership and providing back-channel responses, stalling, humor and play, and persuasive effect. For this study, any NP involved in a direct repetition of two IUs that occur with no more than 1 intervening IU, whether produced across two speakers as in (34) and (35) or by the same speaker as in (36), were coded as repetitions. Note that some instances involve the repetition of an entire clause, or in other cases, a free NP is repeated. NPs that were repeated across two separate clauses but other elements of the clause were different (see (37)) were not counted as repetitions.

(34) Clausal repetition across speakers

F: [y todavía] --
todavía hay la tarabita?
S: todavía hay la tarabita.

F: [and still] --
the cable car is still there?
S: the cable car is still there.

(Deportes:191-93)

(35) Free NP repetition across speakers

M: vamos a servirnos un rico café [ahora] que nos hace frío.
G: [XX]
un rico cafecito.
M: .. un rico cafecito.

M: let's have a nice cup of coffee [now] that it's cold.

G: [XX]
a nice little cup of coffee.
M: a nice little cup of coffee.

(Café:386-389)

(36) Self-repetition

R: ... toma juguito,
toma juguito.
R: ... drink juice,
drink juice.

(Jugo:863-64)

(37) Two occurrences of same NP not coded as a repetition

M: nos vamos con **ellos**,
regresamos con **ellos**,
M: we go with **them**,
we leave with them,

(Café:359-60)

4.3.3 Co-construction

Co-constructions are described by Lerner (1991) as a single sentence that is collaboratively produced by two speakers. Helasvuo (2001a) finds that full NPs in predicate nominal constructions provide an opportune place for participants to co-construct the conversation. NPs involved in co-constructions, such as example (38), were noted. NPs that are involved in co-constructions may exhibit patterns of realization or modification that are different from those predicted by information flow.

(38) Co-construction

M: y ha salido con o= --
con muy buenas eh,
G: **notas**.
M: and he ended up with o= --
with really good uh,
G: **grades**.

(Café:84-86)

4.3.4 Lists

The occurrences of three-part (or more) lists has been noted for a number of languages (cf. Helasvuo 2001a; Jefferson 1990; Sánchez-Ayala 2003), and it is clear that speakers orient toward lists in conversation. Following Sánchez-Ayala (2003), lists were identified as having at least 3 parts, each item on the list occupies its own intonation unit, and each part of the list exhibits a particular intonation contour, usually transitional for the first and second parts, final or rising for the final part. Lists found in this data were sometimes comprised of three parallel clauses, prepositional phrases, or simply NPs. Because the three parts of a list are all similar (in order for the listener to recognize it as a list), NPs involved in lists may be subject to additional constraints on modification and/or realization not predicted by information flow. Two lists are exemplified in (39) and (40) below.

(39) Three-part list comprised of parallel clauses

.. el arroz es carbohidrato,
las papas son carbohidratos,
el fideo es carbohidrato.

'.. rice is a carbohydrate,
potatoes are carbohydrates,
noodles are a carbohydrate.'

(Food:82-84)

(40) Three-part list comprised of prepositional phrases

entonces yo puedo presentar en **el KFC**,
en **la CH Farina**, ((CH FARINA = PIZZA PLACE))
en **donde yo quiera**.

'so I can present [a voucher] at KFC,
at CH Farina,
at wherever I want.'

(Jugo:166-68)

4.4 Summary

As the following chapters will illustrate, both information flow and interactional factors exhibit considerable influence on the coding and syntax of NPs. Chapter 5 examines the grammatical manifestations of discourse referentiality, and Chapter 6 investigates the factors that influence form and distribution of referential NPs, specifically how information flow concerns are reflected in the realization, modification and syntactic role of NPs. The same factors are examined for non-referential NPs in Chapter 7. Lastly, Chapter 8 provides an account of all of the NPs involved in the abovementioned interactional constructions (increments, repetitions, co-constructions and lists). In addition, this chapter profiles the interactional uses of the NPs in the syntactic role ‘Other’ and explores interactional motives for NPs that do not conform to the patterns of realization as predicted by theories of information flow.

5 Discourse referentiality

This chapter examines the grammatical manifestations of discourse referentiality. Recall from Chapter 4 that all NPs in this study were coded as either discourse referential or discourse non-referential following Du Bois and Thompson (1991). Here, the data (excluding the unclear tokens) are examined in order to look for correlations between discourse referentiality and formal properties (i.e., presence of modifier, grammatical role, persistence, etc.) that have been linked to referentiality in the literature. Then, the data were analyzed in order to determine which combination of formal factors is representative of the distinction between discourse referential and non-referential NPs.

5.1 Factors related to discourse referentiality

5.1.1 *Persistence*

The relationship between persistence (i.e., continuity of reference in discourse) and discourse referentiality has often been discussed in the literature (Du Bois and Thompson 1991; Helasvuoto 2001b; Torres Cacoullos and Aaron 2003), although the correlation between persistence and discourse referentiality is not straightforward. Torres Cacoullos and Aaron (2003) attempted to use persistence as a measure of tracking status in their study of lone English-origin nouns in Spanish, but found this to be unsatisfactory as speakers 'seem to present NPs as tracking and then do not track them' (2003:324), i.e., they do not persist.

In fact, when one examines spoken language, examples such as (41) below, in which *sol* ('sun') is mentioned several times but is clearly not referential, show that NPs may clearly persist in the discourse without necessarily being discourse referential.

(41) Persistent discourse non-referential referent

A: qué más,
Caro?
.. te vas a clases ya?

C: .. sí,
sí me voy ahora.

A: .. en serio?
pero está haciendo **basta=nte so=l**.
.. qué te vas a poner?

C: sí,
todos estos días ha hecho **mucho sol**,
pero ahora,
me voy a poner un poco abrigado,
porque estoy un poco con gripe.

A: .. (H) uh=,
pero te vas a calzina=r,
sí está haciendo **un sol**,
creo que estamos a unos treinta y cinco grados.

C: .. sí=,
debe se=r.
.. [ayer] también hizo **un sol** --

A: [uh].

C: .. bastante fuerte.

A: what else,
Caro?
.. are you going to class already?

C: .. yeah,
yeah I'm going now.

A: .. really?
but **the sun** is really strong.
.. what are you going to wear?

C: yeah,
these last few days there has been a lot of **sun**,
but I'm going to bundle up,
because I've got a little cold.

A: .. uh,
but you're going to fry,
yeah there's really a lot of **sun**,
I think it's around 35 degrees.

C: .. yeah,
must be.
.. yesterday also **the sun** was --

A: uh

C: .. really strong.

(Classes:1-22)

In Givón's (1995) sense, an NP is deemed 'persistent' if it occurs again in at least one of the following ten clauses³⁰. For this study, this measure was combined with Givón's measure of 'lookback,' whether or not a referent occurred in the preceding ten clauses. Therefore, referents in this study that were also mentioned in the preceding or following ten clauses were coded as persistent. In combining Givón's persistence and 'lookback' all occurrences of both previous and subsequent mentions in the immediately surrounding discourse are captured, whereas if persistence alone were used, only the subsequent mentions would be included. It should also be noted that Givón (1995) assigns referents a Topic Persistence value based upon the number of times that a referent occurs again in the following ten clauses; here only the binary distinction between persistent and not-persistent was made. There are certain NPs (vocatives, *nada* 'nothing', *nadie* 'no one', *alguien* 'someone', *todo eso* 'all that', *cuenta in darse cuenta* 'to realize' etc.) for which persistence is not applicable. Results for persistence are given in Table 4.

Table 4: Discourse Referentiality and Persistence in both genres (N=4405³¹)

Persistence	Non-referential		Referential		Total	
Persistent	634	25%	1869	75%	2503	100%
Not persistent	830	81%	189	19%	1019	100%
N/A	855	97%	28	3%	883	100%
Total	2319	53%	2086	47%	4405	100%

There clearly is a correlation between persistence and referentiality, but 25% of persistent NPs are still non-referential, and 19% of non-persistent NPs are referential. In

³⁰ Givón first introduces the notion of persistence in the Introduction to his 1983 volume, where he states that 'There is no maximal value assigned *by definition...*' (1983:15). It is in his 1995 work that he specifies that persistence is 'the number of times the referent recurs within the next (cataphoric) 10 clauses following its present occurrence' (1995:79).

³¹ Recall that of the 4418 NPs in the corpus, tracking status for 13 was not coded, resulting in the 4405 tokens coded in this section.

the following sections, the relationships between other properties of NPs and discourse referentiality are examined in order to determine if there are stronger indicators than persistence alone.

5.1.2 NP Realization

In a study of Javanese conversation, Ewing (2005) found a correlation between the form of NPs and discourse referentiality; specifically he found that the larger nominal expressions, such as headless relative clauses and full NPs, are more commonly associated with non-referential (non-tracking in his terms) referents, and that pronouns and evoked (unexpressed) mentions are more commonly associated with discourse referential NPs. Proper names tend to be non-referential, especially when used as vocatives, but may be tracking in other grammatical roles. Ewing's findings for Javanese (all grammatical persons, referring expressions only; his Table 4.6, 2005:150) are copied here for convenience in Table 5:

Table 5: Discourse Referentiality and Expression type in Javanese (Ewing 2005)

Expression		Non-referential		Referential		Total	
Headless Clause	Relative	7	88%	1	13%	8	100%
Lexical		100	65%	55	35%	155	100%
Proper Name		51	76%	16	24%	67	100%
Pronoun		10	15%	55	85%	65	100%
Unexpressed		1	1%	121	99%	122	100%
Total		169	41%	248	59%	417	100%

The coding scheme used in the study of Ecuadorian Spanish is slightly different, with two distinctions made between pronominal forms of expression (personal pronouns

and others³²), no separate category designated for proper names, which are included in the lexical category, as mentioned above, the exclusion of headless relative clauses, and the inclusion of non-referring expressions. The results for both genres of Spanish, shown in Table 6, are similar to those Ewing obtained for Javanese. Ewing finds that both personal pronouns and unexpressed (Ewing's 'evoked') mentions tend to be referential in Javanese, whereas in the Ecuadorian data personal pronouns are 84% referential and unexpressed mentions are 79% referential. The higher proportion of unexpressed non-referential mentions in the Spanish data can be attributed to the use of general (non-specific) third person plural, as in example (42).

(42) Non-specific unexpressed third person plural³³

F: ... (3.0) y tú necesitas ir a --
pasantía,
y contactarte con un asesor.
A: [ya].
F: [para] ver cuántos negocios **cierran** de cada cuántas personas.

F: and you need to go to --
internship,
and get in contact with an assessor.
A: yeah.
F: to see how many deals **(they) close** out of each number of
people.

(Birthday:115-119)

³² Demonstrative, interrogative and other pronouns were grouped together here because they have very similar patterns.

³³ It can be argued that non-specific third person plural referents could be excluded on the grounds that they function as impersonal statements and therefore have no subject; they were included in this study because as Lapidus and Otheguy (2005) point out, there are different degrees of nonspecificity, that is, in certain contexts, the subject does in fact refer to a set of possible referents or a designated group of people, and in some contexts, these 'impersonal subjects' are in fact expressed in Spanish.

Table 6: Discourse Referentiality and Expression type in both genres (N=4405)

Expression	Non-referential		Referential		Total	
Lexical	1704	64%	961	36%	2665	100%
Personal pronoun	83	16%	426	84%	509	100%
Other pronouns	355	90%	38	10%	393	100%
Unexpressed	177	21%	661	79%	838	100%
Total	2319	53%	2086	47%	4405	100%

The relationship between type of expression and referentiality is very strong for ‘other pronouns’. Non-personal pronouns and show a very strong tendency (90%) toward non-referentiality. Personal pronouns show a tendency to be referential (84%); however it will be shown later that a stronger correlation is found when combining expression with specificity. For the remaining types of expression (lexical and unexpressed) the relationship between referentiality and form of expression is even less strong, so properties of personal pronouns, lexical and unexpressed mentions were further analyzed.

5.1.3 Modification

In their analysis of bare English-Origin nouns in New Mexican Spanish, Torres Cacoullos and Aaron (2003) argue that non-referential (non-tracking) status of NPs in Spanish corresponds to lack of determination. Determiners here include all articles, quantifiers, demonstratives and possessive adjectives.

The results for presence of determiner (for lexical mentions, including NPs derived from adjectives only) and referentiality are displayed in Table 7.

Table 7: Discourse Referentiality and Determiner presence for Lexical mentions in both genres (N=2665)

Modifiers	Non-referential		Referential		Total	
Bare ³⁴	684	93%	54	7%	738	100%
Determiner	1020	53%	907	47%	1927	100%
Total	1704	64%	961	36%	2665	100%

Across the two genres bare NPs are almost categorically non-referential (see (43)), which is similar to Torres Cacoullos and Aaron's (2003) findings. However, for those NPs occurring with determiners³⁵, approximately half are referential (see (44)) the other half non-referential (see (45), where it occurs in a fixed expression). In fact, more than half of the non-referential lexical NPs appear with a determiner, 60%, 1020/1704. So, while lack of determiner may be a good indicator that a referent is non-referential, presence of determiner is not a good criterion identifying referentiality. Modification (or lack thereof) thus cannot be used as a measure of referentiality.

(43) Bare, non-referential NP

A: no te tengo **miedo**,

A: I'm not afraid of you (lit. I have no **fear** of you)

(Deportes:599)

³⁴ This category includes some NPs co-occurring with descriptive adjectives, e.g., *hacen cosas sucias* 'they do **dirty things**'.

³⁵ All determiners were collapsed here for convenience. However, determiner type was also examined, showing contradictory patterns between the two genres (NPs with demonstratives tended to be non-tracking in narratives, yet tracking in conversation, and NPs with quantifiers were more commonly tracking in conversations but non-tracking in conversations) or inconclusive patterns (NPs with definite articles, indefinite articles and possessive adjectives were closer to half referential, half non-referential).

(44) Referential NP with determiner

R: le diste **la hoja**?

R: did you give her **the paper**?

(Jugo:397)

(45) Non-referential NP with determiner

A: .. (H) jugando **a las escondidas**?

A: .. (H) playing **hide-and-go-seek**?

(Fumar:220)

5.1.4 *Specificity*

Du Bois and Thompson (1991) note that both particular and general NPs occur with both tracking and non-tracking functions, although there is a loose correlation between specificity (particularity) and tracking function, and generality and non-tracking function. This relationship was explored for the Ecuadorian Spanish data, and several patterns emerge. Table 8 shows the distribution of general and specific personal pronominal and unexpressed mentions according to referentiality. Recall from Table 6 that in the aggregate, personal pronouns and unexpressed are mostly referential (84% and 79%, respectively), which is not the case for lexical NPs, at 36%. General pronominal and unexpressed referents are very strongly non-referential (87.5%, see (46) and (47)), while specific pronominal and unexpressed referents are largely referential (97%, see (48)).

(46) General pronominal NP

J: yo no sé a quien **le** jurabas,

J: I don't know who (**le**) you swore to,

(Jugo:221)

(47) General unexpressed NP

C: pero creo que igual peligroso.
A: .. cla=ro,
ahora [∅ te ven no más],
C: [XXX]
A: y ∅ te pueden robar.

C: but I think that (it's) just as dangerous.
A: .. of course,
now [(they) can see you],
C: [XXX]
A: and (they) can rob you.

(Clases:536-540)

(48) Specific pronominal and unexpressed NP

F: una de mis compañeras tiene un hijito de unos cuatro o cinco años.
A: hm,
ya.
F: entonces **ella** se llevó o=tro,
∅ ya no se llevó este Baby Mozart.

F: one of my colleagues has a little son about three or four years old,
A: hm,
yeah.
F: so **she** took one,
(she) didn't take this Baby Mozart.

(Birthday:252-257)

Table 8: Discourse Referentiality and Specificity in both genres (among personal pronouns and unexpressed mentions; N=1347)

Specificity	Non-referential		Referential		Total	
General	231	87.5%	33	12.5%	264	100%
Specific	29	3%	1054	97%	1083	100%
Total	260	19%	1087	81%	1347	100%

This distinction is less clear cut, however, for lexical mentions, as seen in Table 9. While the same pattern is seen for general referents (strongly non-referential), specific referents are much more divided (only 55% are referential). For this reason, lexical mentions are further examined in the next section.

Table 9: Discourse Referentiality and Specificity in both genres (lexical NPs only; N=2665)

Specificity	Non-referential		Referential		Total	
General	989	92%	82	8%	1071	100%
Specific	715	45%	879	55%	1594	100%
Total	1704	64%	961	36%	2665	100%

5.1.5 Grammatical role

Thompson (1997) found that there is clearly a relationship between core arguments and tracking referents. Specifically, the majority of core arguments are tracking, and the majority of obliques are non-tracking. Therefore, the discourse referentiality and grammatical role of lexical NPs are examined here in Table 10 and Table 11, where the genres are examined separately because the proportion of referential NPs in the O role is quite different between the two. Across all syntactic roles, pronouns and unexpressed mentions were consistently referential when specific and non-referential when non-specific. It was therefore not necessary to look at grammatical role separately for these NPs. The data present similar results for the core vs. oblique distinction in Spanish in the narratives (95% of A, 81% of S, and 72% of objects are referential, with figures ranging from 20% to 66% for arguments in non-core roles), but the distinction is less obvious in the conversational data (84% of A, 64% of S, 37% of objects, and figures for arguments in non-core roles ranging from zero (vocatives) to 39%). Indirect objects are included with core arguments here because they show similar properties to core arguments: specifically, they are often human, as are A and S (Du Bois 1987), whereas obliques are most often non-human. They are collapsed with direct objects because they pattern similarly.

Table 10: Discourse Referentiality and grammatical role in Spanish Narratives (lexical mentions only; N=693)

Role	Non-referential		Referential		Total	
Core roles						
A ³⁶	2	5%	41	95%	43	100%
S ³⁷	34	19%	142	81%	176	100%
Objects (direct and indirect)	59	28%	153	72%	212	100%
Non-core						
Prepositional phrases, Adverbial phrases and Predicate nominals		80%		20%		100%
'Other'	177		45		222	
Extension of previous syntactic frame	11	34%	21	66%	32	100%
	5	63%	3	37%	8	100%
Total	288	42%	405	58%	693	100%

Table 11: Discourse Referentiality and grammatical role in Spanish Conversation (lexical mentions only; N=1972)

Role	Non-referential		Referential		Total	
Core roles						
A	7	16%	37	84%	44	100%
S	114	36%	201	64%	315	100%
Objects (direct and indirect)	231	63%	138	37%	369	100%
Non-core						
Prepositional phrases, Adverbial phrases and Predicate nominals		91%		9%		100%
'Other'	749		72		821	
Vocative ³⁸	173	65%	94	35%	267	100%
Extension of previous syntactic frame	120	100%	-	-	120	100%
	22	61%	14	39%	36	100%
Total	1416	72%	556	28%	1972	100%

In order to look for stronger patterns in the Spanish conversational data, in which grammatical role did not pattern as expected, grammatical role and referentiality were

³⁶ Following Dixon (1979), grammatical role A refers to the subject of a transitive clause, and S is the subject of an intransitive clause.

³⁷ Subjects of reflexive verbs and single arguments were included in this category.

examined in conjunction with specificity: the results are given in Table 12 for the conversational and Table 13 for the narrative data.

Table 12: Specificity, Grammatical role and Discourse Referentiality of full NPs in Spanish Conversation (N=1972)

Specificity	Role	Non-Referential	Referential	Total			
General	A	5	56%	4	44%	9	100%
	S	83	70%	35	30%	118	100%
	Objects (direct and indirect)	195	93%	15	7%	210	100%
	Prep. phrases, Adv. phrases and Predicate nominals	435	98%	10	2%	445	100%
	'Other'	135	91%	13	9%	148	100%
	Extension of previous syntactic frame	17	94%	1	6%	18	100%
General Total		870	92%	78	8%	948	100%
Specific	A	2	6%	33	94%	35	100%
	S	31	16%	166	84%	197	100%
	Objects (direct and indirect)	36	23%	123	77%	159	100%
	Prep. phrases, Adv. phrases and Predicate nominals	314	84%	62	16%	376	100%
	'Other'	38	32%	81	68%	119	100%
	Extension of previous syntactic frame	5	28%	13	72%	18	100%
	Vocative	120	100%	-	-	120	100%
Specific total		546	53%	478	47%	1024	100%
Grand total		1416	72%	556	28%	1972	100%

We can see from Table 12 that when specificity and grammatical role are combined, the patterns are much stronger than when either specificity or grammatical role is examined alone. A clear core-oblique distinction is observed among the specific NPs, with the core arguments correlating with referentiality and non-core NPs strongly

³⁸ Vocatives are categorically non-referential; while the NPs in vocatives are often referring (i.e., they refer to a real world entity), functionally, they are not referential.

being non-referential and all general NPs pattern the same (non-referential). Recall that when objects were examined independently of specificity, nearly two thirds were non-referential and one third referential. When specificity is taken into consideration, 93% of general objects are non-referential, and 77% of specific objects are referential (see (49) and (50)). This is a very important finding and supports the approach taken here that looking at factors in conjunction with one another yields better results than looking at factors independently. Overall, 36% of S tokens are non-referential and 64% are referential, but 70% of general NPs in the S role are non-referential, and 84% of specific S are referential (see (51) and (52)). Specific A is very strongly referential (94%), whereas for general A the tokens are only slightly over 50% non-referential. This is likely due in part to the extremely low token count (only 9 general, lexical A), and the fact that lexical A itself is somewhat unusual, given the 'non-Lexical A Constraint' (Du Bois 1987:829).

(49) General, non-referential object

.. lo= llamaron **la atención** con un .. silbido,
 `.. they called **his attention** with a .. whistle,'

(PS43:88)

(50) Specific, referential object

y le botó **el sombrero**.
 `and (she) knocked off **his hat**.'

(PS29:42)

(51) General, non-referential S

A: (H) qué rápido que pasa **el tiempo**,
 A: (H) how quickly **time** passes,

(Clases:70)

(52) Specific, referential S

R: .. solo **la Daisy** viene.

R: .. only **Daisy** is coming.

(Jugo:518)

Now that NPs in core roles (A,S and O), vocatives, objects of prepositions and adverbial phrases have been accounted for, we are left with NPs in the role 'Other' and 'extension of previous syntactic frame'. Interestingly, NPs in the 'Other' syntactic role appear to pattern more like core arguments, with the majority of general 'Other' referents being non-referential (91%) and the majority of the specific 'Other' NPs being referential (68%). This is an important finding because these NPs have often been left out of analyses, and when they have been studied it has been largely in isolation rather than in terms of how they might pattern like other NPs. NPs that are an extension of a previous syntactic frame pattern like the NPs in the 'Other' syntactic role in the conversations, but as will be seen below, do not follow the same patterns in the narratives.

The core-oblique split for specific NPs observed in Table 12 for conversational data is also seen in the narrative data, shown in Table 13. Specific NPs in the core argument roles (88-98%) and 'Other' role (75%) are strongly referential; and the majority of the non-core arguments are still non-referential (72%). Again, all general NPs have a very strong tendency across syntactic roles to be non-referential (100% in all roles except for objects, where 90% are non-referential). The notable differences between the genres are a) a higher proportion of specific NPs that are referential in the narratives (70% of specific NPs in the narratives are referential, versus 47% in the conversations), b) general NPs make up close to half of the conversational data (48%, 948/1972) but only about

one-fifth of the narrative data (18%, 123/693), and c) the different patterning of NPs in the ‘extension of a previous syntactic frame’ role.

Table 13: Specificity, Grammatical role and Discourse Referentiality of full NPs in Spanish Narratives (N=693)

Specificity	Role	Non-Referential	Referential	Total			
General	A	1	100%	-	-	1	100%
	S	13	100%	-	-	13	100%
	Objects (direct and indirect)	38	90%	4	10%	42	100%
	Prep. phrases, Adv. phrases and Predicate nominals	63	100%	-	-	63	100%
	'Other'	4	100%	-	-	4	100%
	Extension of previous syntactic frame	-	-	-	-	-	100%
General Total		119	97%	4	3%	123	100%
Specific	A	1	2%	41	98%	42	100%
	S	21	13%	142	87%	163	100%
	Objects (direct and indirect)	21	12%	149	88%	170	100%
	Prep. phrases, Adv. phrases and Predicate nominals	114	72%	45	28%	159	100%
	'Other'	7	25%	21	75%	28	100%
	Extension of previous syntactic frame	5	63%	3	38%	8	100%
Specific total		169	30%	401	70%	570	100%
Grand total		288	42%	405	58%	693	100%

5.2 Summary of grammatical manifestations of discourse referentiality

In sum, it was found that a combination of factors is the best predictor of discourse referentiality. For some referents, realization alone proved to be a sufficient criterion—all NPs realized as demonstrative, interrogative and other pronouns, were categorically strongly non-referential. For others, it was a combination of realization and specificity—all specific personal pronouns and unexpressed mentions were strongly

referential, all general personal pronouns and unexpressed mentions were strongly non-referential. For the remaining lexical mentions, it was a combination of syntactic role and specificity—specific lexical mentions as core arguments or in the syntactic role ‘Other’ were strongly as referential, specific lexical mentions that were not core arguments were largely non-referential, and all general lexical mentions were very strongly non-referential. Note that the only specific, non-referential NPs are those that are not core arguments or in the syntactic role ‘Other’.

The least clear patterns were found for NPs in the grammatical role ‘extension of previous syntactic frame’, which do not show clear patterning. For this reason, it was decided to recode their grammatical role according to the role they performed in the previous syntactic frame, for example, in example (21), repeated here as example (53) for convenience, *una parte* is a general direct object, and is recoded as a non-referential referent.

(53) Extension of previous syntactic frame

S: .. vi una parte,
en mi casa,
F: ya.
S: de ahí **una parte** en el aeropuerto,

S: .. I saw part [of the soccer game],
at my house,
F: yeah.
S: and then **another part** at the airport,
(Deportes:83-86)

5.3 Recoding referentiality

Based on the patterns that were observed in this chapter, for the remainder of the study, the data were recoded for referentiality according to the table presented in Table 14.

Table 14: Coding for discourse referentiality³⁹

Discourse Referential	Discourse Non-referential
Specific subjects realized as full NPs, personal pronouns and unexpressed mentions (N=1176)	All other subjects (N=420)
Specific objects (D and IO) realized as full NPs or personal pronouns (N=638)	All other objects (N=357)
Specific NPs in syntactic role ‘Other’ (N=161)	Predicate nominals (N=114)
Extension of previous (referential) syntactic frame (N=19)	Objects of prepositions (N=929)
	Adverbials (N=129)
	Vocatives (N=120)
	General NPs in syntactic role ‘Other’ (N=233) ⁴⁰
	Grammatical words (<i>qué, cuánto</i> , etc.) (N=83)
	Extension of previous (non-referential) syntactic frame (N=39)

It is through the identification of the factors that correlate with discourse referentiality that the NPs in this study can be coded for referentiality in an objective and replicable way. This is important because it allows other scholars to identify referentiality in the same way in order to more readily compare across investigations. It builds upon previous work in which many of these factors were associated with referentiality and lays the framework for future investigations that examine the interaction between referentiality and grammatical patterns.

³⁹ An alternate coding scheme could recode all bare NP (excluding proper names) as non-referential.

6 Discourse Referential NP Results

In this chapter, the effects of information flow on form and distribution of referential NPs in Spanish are examined. The set of NPs examined here are those summarized in Table 14, e.g., all specific lexical and pronominal NPs in core roles and specific NPs in the syntactic role ‘Other’. The data show many patterns that are directly related to information flow. New and given referents show significantly different morphosyntactic patterns. Additionally, a number of genre differences are noted, which highlight the different resources that speakers use when confronted with different levels of information flow pressure, along with different audiences.

The chapter is organized in the following manner. First, an overview of referentiality and activation status of all NPs in each genre is presented. Second, referential NPs are examined. In section 6.2, the realization, modification and syntactic roles of new and given referents are compared. The following section, 6.3, presents additional considerations for referential NPs and in Section 6.4 there is a discussion of how these results relate to Preferred Argument Structure. Lastly, section 6.5 summarizes the findings presented within this chapter.

6.1 Overview of NPs

The overview of NPs presented in Table 15 shows that the distribution of referential NPs in the two genres is strikingly different. Slightly more than one quarter of the NPs in the narratives are non-referential, whereas nearly two thirds of the NPs in the conversational data are non-referential.

Table 15: Overview of NPs by discourse referentiality (N=4418)

	Narratives		Conversations	
Referential	881	72%	1113	35%
Non-referential	347	28%	2077	65%
Total	1228	100%	3190	100%

Chi-square = 484.795, $p < 0.0001$

Why is there a significantly higher proportion of non-referential NPs in the conversations? The much higher proportion of non-referential NPs in the conversations can be explained in part by the higher overall proportion of non-core arguments (which are largely non-referential, 86% in narratives and 91% in conversations). As seen in Table 16, there is nearly twice the rate of non-core NPs in the conversations (48%) than in the narratives (25%).

Table 16: Overall distribution of core vs. non-core NPs (N=4418)

	Narratives		Conversations	
Core NPs	923	75%	1668	52%
Non-core NPs	305	25%	1522	48%
Total	1228	100%	3190	100%

Chi-square = 191.298, $p < 0.0001$

Specific uses of non-referential NPs are described in section 6.3. Non-referential NPs fulfill discourse functions in the conversations that are not present in the narratives (including vocatives) or are less frequent in the narratives (predicate nominals, object-verb compounds, and non-referential NPs in the 'Other' syntactic role, which will be discussed further in Chapters 7 and 8). Non-referential NPs in the narratives are more limited to specific functions such as orienting and situating referential NPs (especially as objects of prepositions). The distribution of NPs by referentiality is then a genre difference—non-referential NPs are more common in the conversations because they are

used not only to situate and orient referents, but they also perform a number of interactional functions, as will be seen in Chapter 8.

Not only are there genre differences with respect to distribution of discourse referential versus non-referential NPs, there are also differences in the distribution of referential tokens by activation status (Table 17).

Table 17: Overview of discourse referential NPs by activation status⁴¹ (N=1861)

	Narratives		Conversations	
Given	660	84%	748	70%
New	130	16%	323	30%
Total	790	100%	1071	100%

Chi-square = 45.614, $p < 0.0001$

Note that there is a significantly higher proportion of given referents in the narratives than in the conversations (84% vs. 70%). This suggests a higher rate of persistence in the narratives.⁴² When the rates of persistence of new NPs between the two genres are examined, as presented in Table 18, it becomes clear that new referents in the narratives are significantly more likely to be persistent (that is, to be referred to subsequently) than new referents in the conversations (62% vs. 46%). That is, having introduced a new referent, speakers are more likely to refer to it again with the next 10 clauses in a narrative than a conversation. However, given the low number of tokens of new referents (130 in narratives, 323 in conversations), it is not evident that the

⁴¹ Note that the total here is smaller than the total number of referential NPs given in Table 15. A subset of tokens were coded for different activation statuses, discussed in 6.3.2.

⁴² Recall from Chapter 4 that persistence was coded independently from referentiality and information flow. Therefore, persistent NPs, that is, those that were referred to within the previous or following ten clauses, may be new, given or non-referential. Not all new NPs are persistent, nor are all given NPs persistent (in the event that their previous or subsequent mention occurred at a distance of greater than 10 clauses). Non-referential NPs were coded as persistent or not based upon both previous and subsequent mention.

difference is sufficient to account for the much higher proportion of given referents in the narratives.

Table 18: Persistence of new referents (N = 453)

	Narratives		Conversations	
Persistent	81	62%	149	46%
Not persistent	49	38%	174	54%
Total	130	100%	323	100%

Chi-square = 9.706, $p < 0.01$

When the data are examined more closely it is evident that a particular group of referents within the narratives are referred to repeatedly. The 81 persistent new referents in the narratives are subsequently referred to again an average of nearly 10 times (806/81), while in the conversations, the average is about five times (802/149). Therefore it is not just the overall rate of persistent NPs that accounts for the higher proportion of given referents in the narratives, but also the 'longevity' of their topicality.⁴³ The following example, (54), illustrates the high topicality (or persistence) of the boy who steals the pears in one of the narratives. In this example, the boy is referred to as the subject of 10 out of 11 consecutive main clauses. Compare with (55) (in which all referential NPs are bolded), from the conversational data, where over nine clauses, there are five NPs with different referents.

⁴³ It is important to keep in mind that this study focuses only on the third person. The rate of persistence (or topicality) of other referents (first and second person) may be much higher for conversation. However, Travis (2007) reports that there are higher rates of topic continuity for 1s in narratives than in conversations. Together with the results present here, this suggests that there are more topic shifts in conversations than in narratives for all grammatical persons.

(54) Topic continuity in narrative

pasa **un niño** en una bicicleta,
... y=
∅ pensó= .. coger .. primero una o dos peras.
pero ∅ decidió al final,
coger todo el canasto que estaba lleno.
... (2.0) ∅ cogió,
y ∅ subió a su bicicleta,
y ∅ se fue por el --
y ∅ siguió su camino.
... (2.0) de ahí,
bueno,
pues,
... (2.0) ∅ se distrajo al ver una niña ... pasar por a- --
... pasar por su --
... (2.0) por su delante,
∅ se distrajo y ∅ se chocó con una piedra,
tropezó= .. la bicicleta,
y **él** se cayó.
Se regó todos los --
todos los frutos,
todas las peras,
.. el canasto,

a boy passes by on a bicycle,
and,
he thought about taking first one or two pears.
but (**he**) ultimately decided,
to take the whole basket that was full.
... (2.0) he took (it),
and (**he**) got onto his bicycle,
and (**he**) went by the --
and (**he**) went on his way.
... (2.0) from there,
well,
well,
... (2.0) (**he**) got distracted upon seeing a girl ... pass by a- -
... pass by his --
... (2.0) in front of him,
(**he**) got distracted and (**he**) crashed into a rock,
the bike .. faltered,
and **he** fell.
Out spilled all of the --
all of the fruits,
all of the pears,
.. the basket,

(PS12: 37-58)

(55) Topic continuity in conversational data

R: **el Dino** juega bien,
[**el Santi** también].
J: [**ese Dino** juega bien],
chucha.
el= --
quién más,
el Zuneo solo.
R: **el Lucho** no.
el Lucho se [da muchas vueltas].
J: [**el Lucho**] --
<X y viste el X> **Lucho**,
∅ es alegón,
me da iras,
porque ∅ es alegón.
R: ah ah.
.. es bien alegón **el Lucho**,
y **el Roro** también,
ya estaba igual.

R: **Dino** plays well,
[**Santi** too].
J: [**that Dino** plays well],
shit.
the= --
who else,
just **Zuneo**.
R: not **Lucho**.
Lucho [runs around a lot].
J: [**Lucho**] --
<X and you saw X> **Lucho**,
(**he**) is a cheater,
it makes me mad,
because (**he**) is a cheater.
R: uh huh.
Lucho really is a cheater,
and **Roro** too,
was the same.

(Jugo: 774-791)

In sum, we see marked differences in the overall distribution of the NPs in the two genres. In the conversations there are more non-referential NPs, partly due to interactional roles of non-core NPs. In conversations there more new NPs and lower persistence of those NPs as speakers switch topics (cf. Travis 2007)

The following sections examine formal differences among referential NPs—realization, definiteness and syntactic role.

6.2 Formal characteristics of discourse referential NPs

As outlined in Chapter 4, all referential NPs were assigned an information status code of either new or given, as opposed to non-referential NPs, for which information status does not apply. In addition, according to the coding scheme in this study, they are also all specific NPs. The aims of this section are to examine and contrast the forms that new and given NPs take (realization and definiteness), what syntactic roles are most conducive to presenting new referents in Spanish, and which syntactic roles are productive for given referents. A comparison between the two genres is made in each subsection, showing how different strategies for introducing new referents and given referents are used when information flow pressures are high (narratives) and when they are low (conversation).

6.2.1 *Discourse referential NPs and realization*

6.2.1.1 *New NPs and realization*

Unsurprisingly, as Table 19 shows, new referents are introduced almost categorically as full NPs (95% in the narratives, 90% in the conversations, the difference between the genres is not statistically significant). Of the few that were coded pronominally, they are either highly marked (and may be analyzed by some as Full NPs, or redundant (see (56)), and the unexpressed new mentions are highly inferable from the context.

Table 19: New referents and realization (N = 453)

	Narratives		Conversation	
Full NPs	123	95%	291	90%
Pronouns	5	4%	21	5%
Unexpressed⁴⁴	2	1%	11	3%
Total	130	100%	323	100%

In the narratives, all five pronominal new referents were part of the double object construction or 'redundant pronouns' (cf. Butt & Benjamin 2004), in which the referent is also referred to with a full NP, as in (56). This structure is also common in the conversations (accounting for 11/21 new pronouns). Another five in the conversation were modified pronouns, as in (57), and four were 'numeric pronouns' (which may or may not be modified), as in (58). The remaining one was part of a joke in which the speaker made a play on words. Note that subject personal pronouns (*él, ella, ellos*, etc.) were not found to introduce new referents, which is not surprising.

(56) Redundant pronoun⁴⁵

∅ **le** encuentra a **unos chicos**,

(**le**) (he) meets **some boys**,

(PS42: 37)

(57) Modified pronoun⁴⁶

R: y **los de la basura** nunca arreglaron la puerta,

R: and **the ones from the garbage** (i.e., the garbage men) never fixed the door,

⁴⁴ This only refers to unexpressed subjects, as unexpressed objects were not included (cf. Chapter 4). Subjects = 44% (537/1228) of all narrative NPs and 32% (1026/3190) of conversational NPs. Of these, 54% (291/537) are unexpressed in the narratives and 52% (538/1026) in the conversations.

⁴⁵ Co-referential clitic pronouns and full NPs were coded as two distinct mentions to capture the broad range of uses of the pronouns. In many contexts, as in this example, the use of the pronoun is variable. For example, in the conversational data, the following example is found, also with a specific, human direct object, yet the redundant pronoun is not found.

(i) viste el Zuneo?,
did you see Zuneo? (Jugo:237)

In this example, the accusative *a* is also lacking, and it is found that the use of accusative *a* is also variable when the direct object co-occurs with the pronoun *le*, as in:

(ii) y ∅ le encuentro ahí la Carolina,
and (I) (le) find there la Carolina, (Fumar:290)

Note that in (56) there is a lack of agreement between the pronoun and the co-referential full NP (*le* is singular, *unos chicos* is plural), which suggests that the pronoun is not just a copy of the full NP, or an agreement marker, rather it has a separate function, perhaps to signal to the interlocutor that new information will be presented.

⁴⁶ *Los* was coded as a pronoun, however it could also be considered to be the definite article standing alone (i.e., not modifying a head noun). For the purpose of this analysis it is considered a pronoun, albeit an infrequent form, which shows that pronouns that introduce new information are different from pronouns that signal given information.

(58) New, numeric pronoun⁴⁷

A: (H) ahí está **una de mis mejores amigas,**

A: (H) there is one of my best friends,

(Deportes:125)

Both of the unexpressed mentions of new referents in the narratives were produced by the same speaker, who may have been struggling with the high information pressures of introducing so many new referents into the narrative or perhaps with the task of narrating the story itself. One of these is illustrated below, in (59). In this example, taken from the beginning of the transcript, the speaker first introduces the man who picks pears. He then goes on in the twelfth line to refer to the young man who walks past with a goat (*ø estaba halando una oveja* ‘(he) was pulling a sheep’). In line 12 of this example and again in lines 16 (*.. ø estaba halando por ahí* ‘(he) was pulling over there) and 17 (*ø asoma* ‘(he) appears’) the subject is unexpressed. It is not until line 18, a full 6 IUs after the first mention, that the referent is realized as a full NP (*y asoma otro joven* ‘and another young man appears’).

⁴⁷ All cases of *uno/una/unos/unas* were coded as pronouns in this study. These cases could be argued to be part of a morphological scale between a determiner and a pronoun, much like the head noun in (57) could be argued to be a pronoun or a determiner.

(59) Unexpressed first mention in narrative

1 ... (H) ahí asoma un campesino,
2 ... (2.0) (H) que está subiéndose la escalera,
3 .. a cosechar la fruta.
4 ... esa fruta es= .. peras.
5 ... (3.0) entonces,
6 ø está cosechando las frutas,
7 ya ø cogió una canasta muy= ... grande.
8 la canasta es grande.
9 .. y ø está cosechando,
10 ya ø cogió dos,
11 .. (H) asoma --
12 .. ø estaba halando una oveja.
13 ... o un chivo.
14 .. que es de la misma .. raza,
15 no?
16 (H) .. ø estaba halando por ahí,
17 ø asoma,
18 y asoma **otro joven**,
19 ... otro campesino ahí asoma con la oveja.

1 ... (H) and a peasant appears,
2 ... (2.0) (h) that is climbing up the ladder,
3 .. to harvest the fruit.
4 ... that fruit is= .. pears.
5 ... (3.0) so,
6 (he) is harvesting the fruits,
7 (he) already got a really big basket.
8 the basket is big.
9 .. and (he) is harvesting,
10 (he) already got two,
11 .. (H) appears -
12 .. (**he**) was pulling a sheep.
13 ... or a goat.
14 .. that is the same .. species,
15 no?
16 (H) .. (he) was pulling over there,
17 (he) appears,
18 and **another young man** appears,
19 ... another peasant appears with the sheep.

For the conversational data, in all the instances of new referents being 'introduced' as unexpressed mentions, the referent is clearly inferable from the context. Interestingly, of the 11 tokens of new unexpressed mentions, six refer to someone's parents, as in (60). Note in the example that using an unexpressed mention for the first reference to someone's parents does not cause a problem, as evidenced by Rodrigo's response in (60).

(60) Inferable unexpressed mention

A: mañana=na no va a haber nadie,
 ø se van a Cuenca.
 R: .. pero no ø saben que ø van a hacer la fiesta?
 A: sí sabe la mamá.

A: tomorrow no one is going to be there,
 (**they**) are going to Cuenca.
 R: .. but (they) don't know that (they're) going to throw the
 party?
 A: yeah the mom knows.

(Fumar: 431-434)

6.2.1.2 *Given NPs and realization*

Given referents, by contrast, have been mentioned previously in the discourse, so it follows that they are, to some degree, topical. It is often claimed that continuous/accessible NPs will receive less grammatical coding than their newer, discontinuous/inaccessible topics (cf. Givón 1981).

Table 20: Given referents and realization⁴⁸ (N = 1408)

	Narratives		Conversation	
Full NPs	223	34%	201	27%
Pronouns	155	23%	205	27%
Unexpressed	282	43%	342	46%
Total	660	100%	748	100%

⁴⁸ The counts for given referents do not include referents that were mentioned previously as a group ('member of a set') or those plural referents whose members were previously introduced individually. These referents are examined separately below.

It is clear, however, from Table 20, that a significant proportion of given referents (34% in narratives and 27% in conversation⁴⁹) receive substantial grammatical coding (i.e., full NPs). Givón's (1981) claim is partially true; when the distribution of given referents expressed as pronouns and unexpressed mentions is compared with that of new referents (cf. Table 19), it is clear that a much higher proportion of given referents receive less grammatical coding than new referents. Yet, while the number of given referents which are encoded as a full NP is much smaller than the proportion of new referents encoded as Full NPs (in the narratives, 34% vs. 95%, and in the conversations, 27% vs. 90%, or approximately one third in each), it is still higher than what may be predicted by theories of information flow and referent realization. In order to determine if these given NPs are being reintroduced into the discourse after a period of time and are therefore being treated as 'new' again, the distance from the previous coreferential mention was measured in clauses, and the results are presented in Table 21 (note that figures in this Table represent a previous mention in all syntactic roles, not just subjects).

Table 21: Distance in clauses from previous mention for given full NPs (N=419)⁵⁰

	Narratives		Conversation	
0	67	30%	88	45%
1-2	54	24%	45	23%
3-4	24	11%	15	8%
5-10 clauses	38	17%	14	7%
10+	40	18%	34	17%
Total	223	100%	196	100%

⁴⁹ The proportion of given referents which are full NPs in the narratives is significantly higher than the number of given referents that are full NPs in the conversations (Chi-square = 7.644, $p < .01$).

⁵⁰ Five referents were excluded from the conversational data because they were third person references to participants in the conversation and the previous realization was first or second person.

As Table 21 shows, NPs that were referred to in the previous three clauses (i.e. 0-2 clauses) do account for a large portion (54% in narratives, and 68% in conversation⁵¹) of the given NPs that are realized as full NPs. Even more surprising is that nearly half (45%) of the given full NPs in conversations were produced in the following clause, as in (61), and approximately a quarter (23%) at a distance of 1-2 clauses, as in (62). In (61) it is the same speaker who produces the same full NP in two consecutive clauses; it appears as though the emphasis as evidenced by the use of *mismo* ‘herself’ may have motivated the use of the expression (versus unexpressed) for the second mention. In (62), three speakers are collaboratively narrating an event and additional factors such as repetition as and narrative structure techniques to signal change in setting show clear influence on the form of the NP *Marcolín*. It is interesting that two speakers (Rodrigo and Gerardo) produce identical IUs almost simultaneously, both of which contain a given full NP that was referred to as a full NP at a distance of just one clause away. Clearly, distance from previous mention, and by extension, topic accessibility/continuity is not the only factor related to the realization of given referents (as noted in Fox (1987)). Additional factors that contribute to referent realization not predicted by information flow are further examined in Chapter 8.

(61) Full NP followed by another Full NP with no intervening clauses

R: .. pero no ø saben que ø van a hacer la fiesta?

A: sí sabe **la mamá**.

... (2.0) **la mamá mismo** le dijo.

R: .. but (they) don't know that (they're) going to throw the party?

⁵¹ It is interesting that these data show higher topic continuity for 3rd person subjects in the conversations; this is contrary to what Travis (2007) found for 1st person singular subjects in a comparison of conversational and narrative data. This may be an artifact of the Pear Stories, which have 5 salient human actors, as opposed to the narratives data in Travis' study, in the stories are mostly told from a first person point of view.

A: yeah **the mom** knows.
... (2.0) **the mom herself** told her.

(Fumar: 433-435)

(62) Given full NP followed by another given full NP at a distance of 1 clause

G: no=,
el **Marcolín** no más.
ese día estábamos aquí viendo el fútbol,
pues.
R: ajá,
y justo [llegó el **Marcolín**].
G: [y justo llegó] el **Marcolín**,
que venía de la --
M: .. de la fiesta.
así fue.

G: no=,
just **Marcolín**.
that day we were here watching soccer,
anyway.
R: aha,
and [**Marcolín** arrived just then].
G: [and **Marcolín** arrived] just then,
he was coming from the --
M: .. from the party.
that's what happened.

(Café: 1137-1146)

A number of studies (Cameron 1994; Flores-Ferrán 2002; Travis 2005, 2007) have found that priming, or parallel syntactic structures, exhibits a consistent influence upon variable first and second person subject expression in Spanish. Priming effects for realization of third person have been less studied in Spanish (however, see Cameron 1994 and Flores Ferrán (2002) for analyses that include third person), and is perhaps more complicated, as third person referent realization has three forms (Full NPs, pronouns and unexpressed) and is subject to different constraints (particularly information flow) than first or second person subjects. The distribution of the realization of all given full NPs (not just subjects) by form of previous mention is given in Table 22. As Table 22 shows for the conversations, when there is a lexical mention of the referent in the immediately preceding clause (cf. (61)), the majority of full NPs were also previously realized as a full

NP (66%), and pronouns and unexpressed mentions also show a tendency toward mirroring the same previous realization (46% of pronominal given mentions that were referred to in the previous clause were also realized pronominally in the previous mention, and 50% of unexpressed mentions that were mentioned in the previous clause were also realized as unexpressed mentions in the previous clause). The trend goes beyond just the immediately preceding clause for lexical mentions and unexpressed mentions—regardless of distance, if the previous mention was a full NP, a subsequent NP is always preferred, and if the previous mention was unexpressed, it is also likely that the subsequent mention will be unexpressed, even at a distance of 3-4 clauses. Table 23 shows similar trends in the narratives, where there is no distance between two mentions of the same referent, 58% of full NPs were previously realized as a full NP, 37% of pronominal mentions were previously realized pronominally, and 59% of unexpressed mentions were previously unexpressed.

Table 22: Realization of Given NP, Distance in Clause from Previous Mention and realization of previous mention in conversations (N=734)⁵²

Realization	Distance			Totals
	in Clauses	PM Full NP	PM Pronoun	
Full NP	0	66% (58)	18% (16)	16% (14) (88)
	1-2	69% (31)	7% (3)	24% (11) (45)
	3-4	87% (13)	-	13% (2) (15)
	5-10	64% (9)	7% (1)	29% (4) (14)
	10+	62% (21)	18% (6)	21% (7) (34)
Total		67% (132)	13% (26)	19% (38) (196)
Pronoun	0	28% (34)	46% (56)	26% (32) (122)
	1-2	23% (10)	31% (14)	44% (19) (43)

⁵² 14 NPs were excluded because either the NP referred to a participant in the conversation (3rd person mention, not a vocative), therefore coding for previous mention is irrelevant as the previous mention was either first or second person, or because the recorded portion of the conversation began after the referent had already been introduced, and the previous mention was therefore absent from the transcript, i.e., if target occurred in first 10 clauses.

	3-4	33%	(5)	27%	(4)	40%	(6)	(15)
	5-10	20%	(2)	50%	(5)	20%	(3)	(10)
	10+	50%	(5)	10%	(1)	40%	(4)	(10)
Total		28%	(56)	40%	(80)	32%	(64)	(200)
Unexpressed	0	35%	(76)	15%	(32)	50%	(108)	(216)
	1-2	33%	(27)	23%	(19)	44%	(36)	(82)
	3-4	14%	(3)	23%	(5)	64%	(14)	(22)
	5-10	38%	(3)	13%	(1)	50%	(4)	(8)
	10+	20%	(2)	30%	(3)	50%	(5)	(10)
Total		33%	(111)	18%	(60)	49%	(167)	(338)
Grand Total		41%	(299)	23%	(166)	37%	(269)	(734)

Table 23: Realization of Given NP, Distance in Clause from Previous Mention and realization of previous mention in narratives (N=655)⁵³

Realization	Distance		PM Full NP		PM Pronoun		PM Unexpressed		Totals
	in Clauses								
Full NP	0	58%	(39)	21%	(14)	21%	(14)	(67)	
	1-2	59%	(32)	24%	(13)	17%	(9)	(54)	
	3-4	54%	(13)	13%	(3)	33%	(8)	(24)	
	5-10	68%	(26)	18%	(7)	13%	(5)	(38)	
	10+	78%	(31)	13%	(5)	10%	(4)	(40)	
Total		63%	(141)	19%	(42)	18%	(40)	(223)	
Pronoun	0	30%	(30)	37%	(37)	32%	(32)	(99)	
	1-2	29%	(12)	26%	(11)	45%	(19)	(42)	
	3-4	50%	(4)	38%	(3)	13%	(1)	(8)	
	5-10	33%	(1)	33%	(1)	33%	(1)	(3)	
	10+	100%	(1)	-	-	-	-	(1)	
Total		31%	(48)	34%	(52)	35%	(53)	(153)	
Unexpressed	0	27%	(60)	14%	(31)	59%	(130)	(221)	
	1-2	45%	(22)	12%	(6)	43%	(21)	(49)	
	3-4	33%	(2)	17%	(1)	50%	(3)	(6)	
	5-10	33%	(1)	33%	(1)	33%	(1)	(3)	
	10+	-	-	-	-	-	-	-	
Total		30%	(85)	14%	(39)	56%	(155)	(279)	
Grand Total		42%	(274)	20%	(133)	38%	(248)	(655)	

⁵³ Five tokens were excluded because the previous realization was truncated

In addition, in nearly all cases, regardless of form of preceding mention, the greater the distance between two mentions, the more likely an NP is to be realized as a full NP. However, even when there is a distance of more than ten clauses, there is still variation in form of expression. The figures in Table 22 and Table 23 suggest that there is a lexical priming effect for Full NPs, but can neither confirm nor deny that structural priming is a relevant factor influencing third person referent realization in Spanish. In order to consider structural priming, a more comprehensive analysis that separates the different syntactic roles in order to account separately for unexpressed mentions, which here only refer to subjects, would have to be undertaken, which is beyond the scope of the current study. However, the figures in the above tables do indicate that information flow alone, particularly the notion that given and topical referents (here operationalized as those which have been previously mentioned in recent discourse) receive lesser grammatical coding, is not able to explain all instances of referent realization. Chapter 8 will focus on interactional factors that influence referent realization and will help account for some of the patterns evident in Table 22 and Table 23.

6.2.2 *Discourse referential NPs and modification*

This section explores the differences in modification, both determiners and relative clauses, between new and given referents.

One of the most frequent types of modification in Spanish is with an article, either definite or indefinite. The terms 'definite' and 'indefinite' have two different, but related, meanings. In the discourse-pragmatic sense, there is a difference between definite referential nouns, or those which are 'assumed by the speaker to be uniquely identifiable to the hearer' (Givón 1978:296) and indefinite referential nouns, which are 'not so assumed' (Givón 1978:296). In the descriptive sense, some languages make such a

distinction in the morphosyntax. Spanish (like English) has definite and indefinite articles, which correspond to a degree to the discourse-pragmatic meaning, yet also have different uses. In this section, the occurrence of new NPs with definite and indefinite markers is examined in detail.

Articles in Spanish are linked to a number of different functions. The definite article is commonly used with entities that the speaker assumes are known or identifiable to their interlocutor (Bosque 1989; Ortiz Ciscomani forthcoming), as well as with generic (referring to an entire class, i.e. they are a sub-class of non-specific; see Torres Cacoullos & Aaron 2003 for evidence for different behavior from other non-specific NPs) entities (Butt & Benjamin 2004; Ortiz Ciscomani, forthcoming). Ortiz Coscomani (forthcoming) lists the following ways in which a speaker may assume that a referent is known to the hearer 1) previous mention, 2) previous mention of an associated entity or an evoked frame, 3) parts of a whole, including body parts, 4) identifiable through 'shared cultural world', 5) nouns of unique reference, 6) presence in the physical environment of the speech occasion. In addition, Butt and Benjamin (2004:31) note that an article (definite or indefinite) is required before nouns modified by a qualifier, as in (63), presumably because the modification is a pathway to identifiability or serves to mark the NP as a unique referent. Ortiz Ciscomani (forthcoming) also notes that the definite article may be used to convey subjectivity and topicality. In addition, several dialects of Spanish use definite articles before proper names (De Mello 1992), which is extremely common in this variety of Ecuadorian Spanish.

(63) Definite article preceding modified NP

G: creo que tiene que divorciarse por **las leyes de= Estados Unidos,**

G: I think that they have to get divorced because of **the laws of the United States,**

(Gerardo: Café, 1416)

In general, the indefinite article in Spanish is similar to the indefinite article in English, albeit less frequent (cf. Torres Cacoullós & Aaron 2003).⁵⁴ It is said to mark new, unidentifiable information (Garachana Camarero, forthcoming), as well as indefiniteness. Unlike English, Spanish makes use of a plural indefinite article, as in (64).

(64) Plural indefinite article

B: **unos cinco tornillos** están ahí.

B: **some five screws** are in there.

(Hermanos: 102)

Garachana Camarero (forthcoming) sums up the difference between the definite and indefinite articles as essentially being tied to the identifiability of the entity—the use of the definite article implies identifiability, the use of the indefinite article implies unidentifiability, and by extension, subjectivity and topicality.

6.2.2.1 *New NPs and modification*

A summary of definiteness of new NPs is presented in Table 24. Once again, there are clear differences between the genres. The majority of new referents in both genres are modified, i.e., are not bare, be that definite or indefinite or indefinite marking, (100% in narratives, 92% in conversations). A significantly higher proportion is definite in the conversations (80%) than in the narratives (57%) (Chi-square = 18.902, $p < 0.0001$).

⁵⁴ There are, however, some notable differences between the languages. Spanish omits the indefinite article preceding nouns referring to profession, occupation, sex, social status, and as the complement of certain verbs (including *ser* 'to be' and *tener* 'to have').

Du Bois (1980:207) finds a smaller proportion of definite first mentions in English narrations of the Pear Film (34%). New mentions which are marked by the indefinite article are much more frequent in narratives (43%) than in conversation (12%).

Table 24: New NPs and definiteness (lexical NPs only; N = 414)⁵⁵

	Narratives		Conversation	
Definite marker ⁵⁶	57%	70	80%	234
Indefinite article	43%	53	12%	34
Bare	-	-	8%	23
Total	100%	123	100%	291

Table 25: New NPs and type of determiner (lexical NPs only; N = 414)

	Narratives		Conversation	
Definite article	33%	41	58%	170
Indefinite article	43%	53	12%	34
Possessive adjective	11%	14	14%	41
Demonstrative adjective	1%	1	3%	10
Descriptive adjective ⁵⁷	5%	6	2%	5
Quantifier	7%	8	3%	8
Bare	-	-	8%	23
Total	100%	123	100%	291

Overall, similar proportions of new nouns are introduced with an article (definite or indefinite) between the two genres (76% in the narratives and 70% in the conversations, see Table 25). However, the definite article is much more productive in the conversations than in the narratives. The results presented in Table 24 and Table 25 suggest that a higher proportion of the new referents in the narratives may be

⁵⁵ Pronominal and unexpressed mentions were excluded from this count, Table 25 and Table 26.

⁵⁶ Definite markers include the definite article, quantifiers, possessive adjectives, demonstrative adjectives and *otro* (coded as ‘other determiner’)

⁵⁷ Recall from the Methodology chapter that some nouns occur with both descriptive adjectives and other determiners (e.g., *mi mejor amiga* ‘my best friend’). These cases (n=30 in the narratives and n=65 in the conversations) were coded for the other determiner (e.g., possessive pronoun) and not the descriptive adjective. These tokens all refer to the adjective *otro* or an ordinal number (e.g., *le están viendo otro*

unidentifiable, given the suggested link between new/unidentifiable information and the indefinite article (Garachana Camarero forthcoming), whereas in the conversation, the data suggest that the majority of new referents are identifiable, due to their definite marking.

One factor that contributes to the higher proportion of new nouns with indefinite articles in the Pear Stories is a genre effect. Of all the new human referents in the Pear Stories, 55% (32/58) are introduced with an indefinite article; only 16% are introduced with a definite article (9/58). On the other hand, in the conversations, only 5% (7/133) of the new human referents are introduced with an indefinite article, as in (65), and 48% (65/133) are introduced with a definite article. Of the new human referents in the conversations, nearly one third (32%, 42/133) are first names, as in (66), which are nearly always modified by a definite article (95% [40/42] of first mentions of names are modified with a definite article in these data)⁵⁸. First names (and other proper nouns) are not used in the Pear Stories, which helps to explain why fewer overall new referents are introduced with a definite article in this genre. However, when the first names are excluded from the analysis, as Table 26 shows, there is still a significantly higher proportion of new referents marked with the definite article in the conversations (51%) versus the narratives (33%). The proportion of non-human referents introduced with indefinite articles is twice as high in the narratives than in the conversations (30% [21/70] in the narratives, 15% [28/181] in the conversations).

colegio ‘they are looking into another school’), as no descriptive adjectives occurred without another determiner. For example:

⁵⁸ Note that first names can be modified by other definite determiners in this variety of Spanish: demonstrative and possessive adjectives, e.g., “pero yo le veo muy cagado a **este Diego**” (Jugo: 1366) and “ahí quería ponerle a **mi Ricardo**” (Café: 693).

(65) New referent modified by indefinite article

F: eh,
una de mis compañeras tiene **un hijito de unos cuatro o cinco años.**

F: eh,
one of my colleagues has **a son about four or five years old.**
(Birthday: 251-253)

(66) First name with definite article

J: .. qué más que quieres que te cuente Rodriguito?
...(2.0) ah,
ahora también me llamó a la oficina **la .. Claudia.**

J: .. what else do you want me to tell you Rodriguito?
...(2.0) ah,
today **Claudia** called me at the office.

(Jugo: 300-302)

Table 26: Proportion of new referents introduced with definite article (excluding proper names) (N = 372)

	Narratives		Conversations	
Definite article	41	33%	129	52%
Other forms of modification or bare	82	67%	120	48%
Total	123	100%	249	100%

Chi-square = 10.591, $p < 0.01$

It should be noted that the use of the definite article with first names is not limited to first mentions; nearly all first names in the conversational data (89%, 127/142, excluding 58 vocative uses, where articles do not occur) are accompanied by the definite article, regardless of activation status⁵⁹.

Other common uses of the definite article for first mentions in the conversations include qualified NPs (n = 27, see (67)), the introduction of a relative of an already mentioned speaker (i.e., anchored referents, cf. Prince 1981) (n = 12, see (68)), body parts (n = 9, see (69)), lists, entities of unique reference (n = 9, see (70)). Of the

remaining new definite NPs, the majority were frame-evoked NPs, as in example (71) below, where the topic of discussion is a soccer game. Of these, only body parts were found in the narratives (n=2), and they were also introduced with the definite article.

(67) Qualified first mention

E: oh=,
te acuerdas cuando jugábamos carnaval con el huevo podrido?
R: .. (H) no,
no recuerdo.
E: qué a=sco,
Rodrigo,
[pusieron] --
R: [y yo te puse]?
E: no=,
le lanzaron a la Verónica Jaramillo,
el balde de agua con el huevo podrido.

E: oh=,
do you remember when we played carnival with the rotten egg?
R: .. (H) no,
I don't remember.
E: how disgusting,
Rodrigo,
[you put] --
R: [and I put (it) on you]?
E: no=,
you (pl.) threw (it) at Verónica Jaramillo,
the bucket of water with rotten egg.

(Viejos Amigos: 278-288)

(68) First mention of a relative

P: ... y **el abuelito de él**,
qué es?
P: ... and **his grandfather**,
what does he do?

(Café: 766-767)

(69) First mention of body part

J: ... me duele **el pie**.
J: ... **my foot** hurts.

(Jugo: 23)

⁵⁹ Of the remaining 11%, 2 of these are modified with a demonstrative adjective and the remaining are bare,

(70) NP of unique reference

P: entonces usted por ver **el mundial** no se ha ido.

P: so you didn't go because you were watching **the World Cup**.

(Café: 1150)

(71) Frame-evoked NP

R: pero,
sí nos salvamos algunas,
yo metía la --
la pierna justito,
[y le] quitaba **la pelota**.

R: but,
we did save some,
I just stuck out my --
my leg,
[and] I took **the ball** away from him.

(Jugo: 1204-1208)

Possessive adjectives are another way to mark definiteness. The proportion of new nouns introduced with possessive adjectives is similar between the two genres—11% of all new referents in the Pear Stories, and 14% in the conversations. In the conversations, possessive adjectives are commonly used (20/41 new nouns modified by possessive adjective) when introducing family members, as in (72). In the Pear Stories, body parts and clothing account for six of the 14 new referents modified by possessive adjectives, and the fixed phrase *siguió su camino* ('he went on his way') accounts for another four.

(72) Introduction of family member

A: ... ya quiero saber si ya están aquí para irle a visitar a **mis sobrinos**.

A: ... I want to know if they are here yet to go see **my nephews**.

(Fumar: 56)

as in (76) below.

Like possessive adjectives, demonstrative adjectives also mark definiteness. New nouns modified by demonstrative adjectives make up a small proportion of the data in each genre (1% in the narratives, and 3% in the conversations). In all instances, they refer to nouns which are clearly identifiable to the interlocutor, either by presence in the physical environment, as in (73) in which the speaker is referring to a telephone lying on the table, or due to shared information, as in (74), where the participant is describing the movie they just viewed and were asked to narrate.

(73) New NP in surrounding environment

R: no ves que está súper descargado **ese teléfono**?

R: can't you see that **that telephone** is super uncharged?

(Jugo: 819)

(74) New 'shared information' NP

esta película comienza muy de mañana,

this movie starts early in the morning,

(PS38: 5)

New, bare nouns are absent in the narratives, but make up 8% of the new, referential nouns in the conversations. Of the 23 tokens of bare, new nouns in the conversation, approximately one third ($n = 8$) are introduced with the verbs *tener* and *haber* (four each), as in (75), and another eight were proper nouns, as in (76). Three were in the syntactic role 'Other', as in (77), where the speaker is offering sugar to another participant for her coffee.

(75) New, bare noun introduced with *tener*

F: pero a las dos tengo en cambio **examen de ventas**=.

F: but on the other hand at two o'clock I have **sales exam**.

(Birthday: 168)

(76) New, bare proper noun

R: ... (2.0) como presta el carro **Lenín**,
R: ... (2.0) how **Lenín** lends out his car,

(Jugo: 262)

(77) New, bare noun in 'Other' syntactic role

G: señora Manuelita,
... **azuquitar**?
G: señora Manuelita,
... **sugar**?

(Gerardo, Café, 383)

New nouns modified by quantifiers are somewhat more common in the narratives than in the conversations (7% of new referents in the narratives, 3% in the conversations)—this difference is more likely due to the design of the Pear Film than a true genre difference between narratives and conversation. All of the new nouns modified by quantifiers in the Pear stories refer to either the three boys or the two baskets of fruit.

New nouns modified by descriptive adjectives that do not co-occur with determiners (e.g., *otro colegio* ‘another school’, not *el otro colegio* ‘the other school’) are a very small proportion of new NPs overall (5% in the narratives, 2% in the conversations). However, the descriptive adjectives among the new referents are primarily limited to one token type: all six tokens from the narratives are of the adjective *otro* (‘other’) + NP⁶⁰, and four of the five in the conversations are the adjective *otro*. The remaining conversational token is an ordinal adjective (*séptimo* ‘seventh’), which is also a definite marker.

Not only are nouns modified by determiners, which syntactically precede the noun, but they can also be modified by relative clauses, which follow the noun. NPs

which are modified by relative clauses may also be modified by determiners (see (78) below). The overall proportion of referential NPs that are heads of relative clauses is given in Table 27 below. Note that the occurrence of relative clauses is very low overall (only 3% of all referential NPs are modified by a relative clause). There is a significantly higher overall proportion of NPs modified by relative clauses in the narratives (5%) versus the conversations (2%). This may very well be an indicator of the higher information flow pressures of the narratives, as it reflects a more 'condensed' form of packaging the information.

Table 27: Overall proportion of referential NPs (lexical NPs and pronouns) modified by relative clauses (N=1093)⁶¹

	Narratives		Conversations	
Not modified by relative clause	322	95%	737	98%
Head of relative clause	17	5%	17	2%
Total	339	100%	754	100%

Chi-square = 5.031, $p < 0.05$

Looking now specifically at the new NPs, as Table 28 shows, as for the overall figures, there remains a significantly higher proportion of new referents in the narratives that are modified by a relative clause, as seen in (78). Again, this likely reflects the higher information flow pressures of the narratives and the need for more information to be presented in a smaller unit⁶².

⁶⁰ This suggests that *otro* is grammaticizing as a determiner rather than functioning as a descriptive adjective.

⁶¹ Excludes unexpressed mentions.

⁶² Thompson (1984) finds a higher rate of subordination in more formal written versus informal written discourse, although relative clauses were excluded from her study. This suggests that formality may be a significant factor in the rate of subordination.

Table 28: New referents and relativization (N=440)⁶³

	Narratives		Conversation	
Not modified by relative clause	113	88%	299	96%
Head of relative clause	15	12%	13	4%
Total	128	100%	312	100%

Chi-square =8.688, $p < 0.01$

(78) New referent modified by relative clause

existía primeramente **un= campesino que estaba cosechando peras.**

'at first there was **a= peasant that was harvesting pears.**

(PS41:6)

6.2.2.2 Given NPs and modification

As mentioned earlier, one of the uses of the definite article in Spanish is to signal given information. As seen in Table 29, the high percentages of given referents that are modified by definite articles in these data (78% in narratives, 64% in conversations) confirms that given referents are frequently accompanied by the definite article. When these proportions are compared to the proportion of new referents modified by the definite article (33% in narratives and 58% in conversations, cf. Table 25) the robustness of this use is evident. As opposed to new referents, however, where the definite article is more common in conversations, for the given referents, the definite article is more common among in the narratives. This is likely due to the higher frequency of other definite markers (e.g., possessive and demonstrative adjectives) in the conversations; when these are collapsed, we see that 82% of given referents have some kind of definite marking in the conv, 87% in the narratives.

⁶³ Excludes unexpressed mentions.

Table 29: Given referents and type of modification (N=423)⁶⁴

Modifiers	Narratives		Conversation	
Definite article	174	78%	129	64%
Bare	4	2%	24	12%
Possessive adjective	14	6%	20	10%
Demonstrative adjective	6	3%	16	8%
Indefinite article	9	4%	8	4%
Quantifier	14	6%	3	1%
other determiner	2	1%	1	-
Total	223	100%	201	100%

However, we do still have a higher rate of indefinite given NPs in conversation, as 12% of the given referents in conversation are bare vs just 2% in the narratives. An examination of these NPs shows that 16 of the 24 (67%) are proper nouns (but not necessarily first names), four are parts of lists (all first names), four are repetitions (one of which is a proper noun), and two are co-constructions. This suggests that bare given NPs perform somewhat specialized functions within the conversations. Properties of NPs performing interactional work will be examined in greater detail in Chapter 8. The proportion of given referents which are modified by possessives and demonstratives is higher in the conversations (raising the overall rate of definite NPs in this genre), and the proportion of NPs modified by quantifiers and descriptive adjectives is similar between the two genres.

The distribution of given referents modified by a relative clause (6% in narratives and 1% in conversations, see Table 30) is very similar to the distribution of all referents and relativization (5% in narratives and 2% in conversations, see Table 27) above, and shows the same trend (i.e., more modification in the narratives) as the distribution of new referents modified by a relative clause (12% in narratives and 4% in conversations, see

⁶⁴ As with modification of new referents, only full NPs are analyzed here.

Table 28). Again, the higher proportion of NPs in the narratives that are part of a relative clause reflects the higher information pressures of the narratives. Not only is it a more condensed form of packaging, but in the examples below, it allows speakers to syntactically treat referents as given (in the sense that they are modified with a definite article), yet give additional information about the referent to remind the interlocutor of the relevance of this particular referent. In (79) and (80) the relative clauses serve to help identify referents that had not been mentioned in recent discourse.

(79) Given referent modified by a relative clause

mientras tanto,
el campesino que estaba en la cosecha,
 .. (TSK) (H) baja y=
 se da cuenta de que no hay una canasta.

meanwhile,
the peasant that was in the harvest,
 .. (TSK) (H) goes down and=
 realizes that one of the baskets is not there.

(PS52: 69-72)

(80) Given referent modified by relative clause

los tres niños siguieron también su camino,
 (H) y= encontraron **el sombrero que se había .. caído,**
 el sombrero del niño.

the three boys continued on their way,
 (H) and found **the hat that had .. fallen,**
 the boy's hat.

(PS 36: 72-72)

Table 30: Given referents modified by relative clause (lexical NPs and pronouns; N=784)⁶⁵

	Narratives		Conversation	
Not modified by relative clause	355	94%	402	99%
Head of relative clause	23	6%	4	1%
Total	378	100%	406	100%

Fisher's exact test, two-tailed $p < 0.0001$

6.2.3 Discourse referential NPs and syntactic role

6.2.3.1 New NPs and syntactic role

From previous studies of Preferred Argument Structure (cf. Ashby & Bentivoglio 1993; Du Bois 1987, 2003; Du Bois et al. 2003) we know that the preferred grammatical roles for new referents are S (subject of an intransitive verb) and O (less agentive argument of a transitive verb). In this study, 87% of new referents in the narratives are in the S and O roles, and 69% of the new referents in the conversation are in the S and O roles (see Table 31).

Table 31: New referents and syntactic role (N = 423)

	Narratives		Conversation	
A	9	7%	22	7%
S	68	52%	110	34%
Direct Object	45	35%	114	35%
Indirect Object	-	-	10	3%
Other	5	4%	60	19%
Continuation of previous syntactic frame	3	2%	7	2%
Total	130	100%	323	100%

One of the interesting trends seen in Table 31 is that a relatively large number of new referents in the conversations are introduced in the 'Other' syntactic role, as in (81) and (82). In (81), Feidy is a family member, and in (82) Santi and Lucho are other friends that the speakers regularly play soccer with.

(81) New referent in 'Other' syntactic role

P: mhm=.
... (H) y **la Feidy**?
.. ella también ya se graduó este año?
...(3.0) y ella,
en qué?
no va a seguir la universidad?

P: mhm=.

⁶⁵ Excludes unexpressed referents

... (H) and **Feidy**?
... did she graduate this year too?
...(3.0) and she,
in what?
(she's) not going to go to college?

(Café: 796-801)

(82) New referents in 'Other' syntactic role

R: el Dino juega bien,
[**el Santi** también].
J: [ese Dino juega bien],
chucha.
el= --
quien más,
el Zuneo solo.⁶⁶
R: el Lucho no.
el Lucho se [da muchas vueltas].

R: Dino plays well,
[**Santi** too].
J: [that Dino plays well],
shit.
the= --
who else,
just Zuneo.
not **Lucho**.
Lucho runs around a lot.

(Jugo: 774-782)

Although these 'Other' or free NPs have been largely ignored in other large scale studies of the effects of information flow (cf. Asby & Bentivoglio 1993; Du Bois 1987, 2003, etc.), several previous works have pointed out that free NPs in a variety of languages can introduce new referents (Bentivoglio 1993; Croft, 1995, 2007; Kumpf 2003; Matsumoto 2003; Ono and Thompson 1994; Tao, 1992). With nearly one fifth of the new referents in the conversation occurring in the 'Other' syntactic role, this is clearly a very productive strategy in Spanish for introducing new referents. It should also be noted that the much higher proportion of 'Other' new referents in the conversations than in the narratives is likely due to the fact that they are highly identifiable to the speakers.

⁶⁶ 'Zuneo' had already been introduced previously in the conversation.

Of the 60, nearly one half (N=25) are proper nouns (Prince's (1981) 'unused'), and another eleven are 'anchored' (Prince 1981) with possessive adjectives, as in (83).

(83) Anchored new referent

R: y **tu pelada**?
... [la recepcionista]?

R: and **your girlfriend**?
... [the receptionist]?

(Jugo: 1089-1090)

6.2.3.2 *Given NPs and syntactic role*

The distribution of given referents and syntactic role is given in Table 32. The two genres are similar. The preferred roles for given referents are subject of intransitive verbs, followed by subject of transitive verb, direct object, and indirect object. Syntactic roles 'Other' and extension of a previous syntactic frame make up less than 10% of the data in each genre. Activation state and syntactic role are considered together in Section 6.4.

Comparing the narratives and conversations, in the narratives we see a higher number of given referents as subjects of transitive verbs (26% vs. 15%, Chi-square = 25.840, $p < .0001$) and direct objects (20% vs. 15%, Chi-square = 6.904, $p < .01$), and a lower proportion of subjects of intransitive verbs (40% in narratives, 50% in conversations, Chi-square = 12.554, $p < .001$). This may be another genre effect, namely, that the narratives are overall higher in transitivity, at least for third person subjects.

Additionally, there are more given NPs in the syntactic role 'Other' in the conversational data than in the narratives (8% vs. 2%, Chi-square = 25.712, $p < .0001$). Over one third of these (22/61) are performing interactive functions, which will be examined in Chapter 8.

Table 32: Given referents and syntactic role (N=1408)

	Narratives		Conversation	
A	171	26%	110	15%
S	266	40%	373	50%
Direct object	132	20%	109	15%
Indirect object	75	11%	93	12%
Other	13	2%	61	8%
Continuation of previous syntactic frame	3	-	2	-
Total	660	100%	748	100%

6.2.4 *Summary of new NPs*

Not only are there proportionally nearly twice as many new referents in the conversations than in the narratives (30% vs. 16%)⁶⁷, but there are also measurable differences between the genres in the way in which new NPs are modified, the number of new NPs that are heads of relative clauses, and the preferred syntactic roles of new NPs. There are no significant differences between the genres for realization.

New NPs are overwhelmingly expressed as full NPs, and tend to be modified by either a definite or indefinite article, and are unlikely to occur bare in form. New NPs in the narratives are more likely to be introduced with indefinite articles, and the definite article is more common in conversation, reflecting the higher rate of identifiability of referents in conversation and the increased shared information between the speakers. The higher rate of new NPs that are modified by relative clauses in the narratives likely reflects the higher information flow pressures of the genre. Regarding syntactic roles, the narratives and conversations show approximately the same rate of new NPs as subjects of transitive verbs, direct objects, and continuations of previous syntactic frames, but there

⁶⁷ This alone suggests that there is higher information pressure in the conversations, however it is important to recall that first and second person referents are not included in this study, which would affect the overall proportion of new referents in the conversation, likely showing a much lower proportion of new referents overall.

are more new NPs in the syntactic role ‘Other’ in conversations, and more new NPs as subjects of intransitive verbs in the narratives, revealing different strategies for referent introduction across genres.

6.2.5 *Summary of given referents*

Given referents make up 84% of the NPs in the narratives and 70% of the referents in the conversations. Despite the difference in proportion between the two genres, the patterns for realization are very similar. There are slight differences in the patterns of modification, relativization and syntactic role.

We found a higher proportion of bare given NPs in the conversations, which has a two-fold explanation. On the one hand, several of the bare given NPs are proper nouns, which are absent from the narratives. On the other hand, it appears that these bare given NPs are also fulfilling interactional functions, which are clearly more prevalent in the conversations than the narratives. Aside from that, there are striking similarities in the patterns of type of determiner of given, full NPs between the two genres.

As was mentioned earlier, the higher information flow pressures of the narratives result in more given NPs being modified by relative clauses, and the likelihood of given NPs being subjects of transitive verbs or direct objects may be an indicator of higher overall transitivity of the narratives, which may also be a result of higher information flow pressures. Higher transitivity may also be related to the higher rate of foregrounding in the narratives (cf. Hopper & Thompson 1980).

6.3 Additional considerations for referential NPs

This section explores two phenomena that are related to information flow, yet direct comparisons between new and given referents cannot be made. First, the semantic

class of verbs that introduce new referents is examined, followed by a section detailing the patterns of referents which are neither truly new or given under the criteria used in this study (i.e., members of a set, re-named mentions, or first group mentions).

6.3.1 *New referents and semantic class of verb*

For Spanish, Ashby and Bentivoglio (1997) find that 'truly new referents' (i.e., not previously mentioned and not identifiable) are commonly introduced with the existential verb *haber* and the verb *tener* ('to have'). Sánchez-Ayala (2005) finds that *haber* has an almost exclusively presentative function in Spanish, but that *ser* ('to be')⁶⁸, *tener*, and motion verbs are also commonly used as presentatives.

The distribution of new referents and semantic class of verb is given in Table 33. In these data, *haber* is scarcely used to present a new referent into discourse. However, copulas (11% in narratives and 19% in conversations) and motion verbs (29% in narratives and 8% in conversations) appear to be rather productive resources for introducing new referents, and perception verbs (14%) are clearly useful in the narratives. *Tener* is more frequently used to introduce new referents in the conversations (8%) than in the narratives (4%). As with preferred syntactic role for new referents, we also different verb types being used in the two genres.

⁶⁸ Yet Ashby and Bentivoglio (1993) report that copulas strongly disfavor new referents in their data.

Table 33: New referents (in A, S, DO and IO roles) and semantic class of verb (N=378)⁶⁹

	Narratives		Conversations	
Other ⁷⁰	50	41%	145	57%
Copula	13	11%	49	19%
Motion	35	29%	21	8%
Perception	17	14%	13	5%
<i>Tener</i>	5	4%	21	8%
<i>Haber</i>	2	2%	7	3%
Total	122	100%	256	100%

While the distribution in Table 33 shows the frequency of different verb classes with new referents, this may also reflect the overall frequency of these verb classes in the data. In order to compare which verb classes favor new referents to given referents, Table 34 and Table 35 show the overall frequency of these verbs, along with proportions of given and new referents. As Table 34 shows, copulas (Chi-square = 15.812, $p < 0.0001$ ⁷¹), perception verbs (Chi-square = 6.686, $p < 0.01$), and *haber* (Chi-square = 5.661, $p < 0.05$) all take significantly higher proportions of new referents in the narratives when compared to the given referents. Neither the proportion of *tener* (Chi-square = 1.178, $p = 0.2777$) nor motion verbs is significantly higher than the average (Chi-square = 3.295, $p = 0.0695$). The case is different for conversations. As Table 35 shows, only *tener* (Chi-square = 12.366, $p < 0.01$) and *haber* (Chi-square = 14.118, $p < 0.01$) show significantly higher proportions of new referents.

⁶⁹ Only core arguments (A, S, DO and IO) are included. New referents in the 'Other' syntactic role do not occur with a verb and are therefore not included in this count.

⁷⁰ Includes cognitive process, speech act and other stative, impersonal and experiencer verbs

⁷¹ Chi-square tests in this section compare the distribution of given and new referents among the different verb types between the genres.

Table 34: Distribution of verb type for Given and New referents in narratives (N=754)

	Given		New		Total	
Other	89%	420	11%	50	100%	470
Copula	58%	18	42%	13	100%	31
Motion	79%	134	21%	35	100%	169
Perception	72%	44	28%	17	100%	61
<i>Tener</i>	75%	15	25%	5	100%	20
<i>Haber</i>	33%	1	67%	2	100%	3
Total	84%	632	16%	122	100%	754

Table 35: Distribution of verb type for Given and New referents in conversations (N=912)

	Given		New		Total	
Other	73%	389	27%	145	100%	534
Copula	72%	124	28%	49	100%	173
Motion	82%	93	18%	21	100%	114
Perception	70%	30	30%	13	100%	43
<i>Tener</i>	48%	19	53%	21	100%	40
<i>Haber</i>	13%	1	88%	7	100%	8
Total	72%	656	28%	256	100%	912

It is not surprising that perception verbs, as in (84) would be more productive in the narratives than in the conversations, as part of the task was for the participants to narrate what they saw. In addition, the nature of the Pear Film, with new characters constantly arriving, as in (85), the relatively high proportion of motion verbs in the narratives is to be expected, although motion verbs are overall more frequent in this genre (22%, 169/754, of referential NPs in the narratives occur with a motion verb, whereas only 12.5%, 11/912, in conversation do).

(84) Referent introduction with verb of perception

acabo de ver **una película**,
 I just saw **a movie**,

(PS13, 3)

(85) Referent introduction with motion verb

luego vinieron .. tres niños más que .. pasaban por ahí,

then three little boys that were passing through came along,

(PS41, 28)

6.3.2 Other types of 'given' referents

A small set of given referents, including individuals who had previously been introduced as belonging to a group (cf. Du Bois' (1980) 'member of a crowd phenomenon'), groups whose members had previously been introduced separately ('first group mention'), and individuals who were renamed is examined in this section. As Du Bois (1980) points out, these types of mentions may be treated differently than other given mentions because the individual referent has not been explicitly referred to, and therefore may not be activated individually. In Table 36 the distribution of these referents is given.

Table 36: Other types of given referents (N=133)

	% of referential		% of referential	
	Narratives	NPs	Conversation	NPs
Member of a set	58	7%	21	2%
Renaming	29	3%	13	1%
First group mention	4	<1%	8	1%
Total	91	10%	42	4%

'Members of a set' and 'renaming' referents are more common in the narratives than in the conversations, which is likely due to the design of the Pear Film (in particular, the higher proportion of members of a set is mostly due to the group of three boys that help the main character collect his pears). However, in each genre, 'members of a set' is

the most common subtype of given referents, followed by renaming, and lastly, 'first group mention.'

As Table 37 shows, individuals whose immediately prior mention had been as members of a group are most likely to be realized as a full NP (66% in the narratives, 76% in the conversations), as in (86). A substantial portion of 'members of a set' in each genre (34% in narratives, 24% in conversations) were realized pronominally. Close analysis reveals that all of the pronominal mentions were numerical pronouns (see (87)).

Table 37: Realization of 'members of a set' (N=79)

	Narratives		Conversation	
Full NP	38	66%	16	76%
Pronoun	20	34%	5	24%
Total	58	100%	21	100%

(86) 'Member of a set' as Full NP

(H) tenía tres canastos.
 .. (H) por ahí pasó una --
 un señor con una cabrita,
 pasaban viendo las peras,
 pero siguieron.
 (H) luego viene un niño,
 .. con una bicicleta.
 (H) entonces es- --
 al ver a la --
 la --
dos canastas llenas de --

(H) he had three baskets.
 .. (H) along came a --
 a man with a goat,
 they went by looking at the pears,
 but they continued on.
 (H) then a boy comes along,
 .. with a bicycle.
 (H) so it's- --
 upon seeing the --
 the --
two baskets full of --

(PS40: 24-34)

(87) ‘Member of a set’ realized pronominally

y con todas las canastas,
.. y parecía que le faltaba **una**,

and with all the baskets,
.. and it seemed that he was missing **one**,

(PS38:90-91)

This indicates that speakers do treat these referents differently than other given referents, where pronominal and unexpressed mentions are the preferred method of coding (67% in the narratives, 73% in the conversations, cf. Table 20). As seen in Table 38, referents which had previously been introduced as members of a set were most likely to be modified by a definite marker (55%) in the narratives, and in the conversations, they were overwhelmingly (94%) introduced with a definite marker.

Table 38: Modification of 'members of a set' (N=54)⁷²

	Narratives		Conversation	
Definite marker	21	55%	15	94%
Indefinite article	16	42%	1	6%
Bare	1	3%	-	-
Total	38	100%	16	100%

It is interesting that in the narratives modification patterns for individuals previously introduced as members of a set are similar to those for new NPs (cf. Table 24). With respect to definiteness, the overall rate of definite marking is similar between both given and new referents in the conversations (80% of new referents and 83% of given referents are definite), and the individuals previously introduced as members of a set in the conversations shows and even higher proportion of definite marking (94%). This

⁷² Only those referents realized as a full NP were included in this count.

shows that the genre difference regarding ‘members of a set’ reflects the overall patterns of modification for new referents in each genre.

(88) Member of a set with indefinite article

```
.. mientras subía y bajaba con los canastos,  
(H) eh=,  
y el árbol,  
.. que= --  
mientras --  
ya cuando llena una canasta,  
  
.. while he went up and down with the baskets,  
(H) eh=,  
and the tree,  
.. that --  
while --  
when he fills a basket,
```

(PS43:26-31)

(89) Member of a set with definite article

```
P: entonces,  
ya  $\emptyset$  le han matriculado en el Colegio Americano,  
pero,  
...(2.0) pero el Robin dice que no quiere que vaya a ese  
colegio,
```

```
P: so,  
(they've) already registered him in the American High School,  
but,  
...(2.0) but [the] Robin says that he doesn't want him to go to  
that school.
```

(Café: 678-681)

Renaming is moderately productive in the narratives. The majority (19/29) of the 'renamed' referents are the pears – referred to sometimes as just fruit (*frutos, frutas*), other times as apples (*manzanas*), and at times just as the harvest (*la cosecha*). Another nine 'renamed' referents refer to different characters. For example, some speakers vacillate between the labels *campesino* ('peasant') and *agricultor* ('farmer') when referring to the man who collects the pears and the man pulling the goat is sometimes *un joven* ('a young man') and at other times a *campesino*. Two speakers experience uncertainty about what

type of animal is being pulled, suggesting both *oveja* ('sheep') and *chivo* ('goat') or *cabra* ('goat') as possibilities. Lastly, one speaker refers to the road as both *camino* ('road') and *sendero* ('path'). In the conversations, renaming often occurs to further specify a referent, as in (90), but giving two different labels to the same referent, as in the narratives, is also found.

(90) Renamed referent

A: y tu ñaño,
 el --
 el --
 el de pelito largo,
 me --
 cómo se lla=ma,

A: and your brother,
 the --
 the --
 the one with the long hair,
 me --
 what's his na=me,

(Adán, Chicas, 422-427)

All of the 29 renamed referents in the narratives are full NPs, as are 10 of the 13 renamed referents in the conversations. The other three are modified pronouns as in (90). Nearly all renamed referents are definite (i.e., treated as given), except for new referents for which the speakers had trouble coming up with an appropriate label, as in (91).

(91) Indefinite renamed NP

.. estaba halando una oveja.
 ... o **un chivo**.
 '.. he was pulling a sheep.
 ... or **a goat**.'

(PS 52:27-28)

The last subset of given mentions, those which had previously been introduced individually and are later referred to as a group, is also the least frequent of the three types, with only four tokens in the narratives and eight in the conversations. They tend to

be unexpressed, indicating that the speaker feels that their interlocutor(s) will have no problem identifying the referent in question. An example follows in (92). In this example, the speaker had previously introduced the boy who steals the pears (the subject of *está yendo en la bicicleta* ‘is going on the bicycle’) and the girl on the bicycle (*pasa una niña* ‘comes a girl’), and subsequently refers to the two of them as a plural subject (*a lo que Ø están pasando juntos* ‘when (they) are passing by each other’).

(92) First group mention

Ø está yendo en la bicicleta,
 eh,
 .. en el que misma --
 en el mismo camino,
 pero en el sentido contrario,
 pasa una niña,
 en otra bicicleta.
 .. (H) a lo que Ø **están pasando** juntos,
 eh,
 .. el sombrero del niño vuela,
 (H) y el niño,
 por regresar a ver el sombrero,
 no ve una piedra grande y se choca.

(he) is going on the bicycle,
 eh,
 .. on the same one that --
 on the same road,
 but on the other side,
 comes a girl,
 on another bicycle.
 .. (H) when (**they**) are passing by each other,
 eh,
 .. the boy's hat flies,
 (H) and the boy,
 upon looking back at the hat,
 doesn't see a big rock and he crashes.

(PS100 131-143)

In sum, the only genre difference noted is in the modification of members of a set, which mirrors the differential treatment of new referents in each genre. The difference in frequency of these referents between the genres (namely that they are more common in the narratives) has more to do with the design of the Pear Film, where there are several

groups (the baskets, the pears, the three boys) that are alternately mentioned as groups or individuals. The 'members of a set' show a tendency to be treated as new referents in each genre, whereas renaming and 'first group mention' are nearly always treated as given/identifiable referents.

6.4 Preferred Argument Structure

Repeated studies of Preferred Argument Structure across many different languages (cf. Du Bois 1987, 2003) have shown universal tendencies. Specifically, there tends to be only one new argument per clause, and new arguments are typically found either as subjects of intransitive verbs or objects. In addition, NPs that are subjects of transitive verbs are typically not expressed as full NPs.

As seen in Table 31 above, 87% of new referents in the narratives are in fact found in the S or O role, but the trend is much less strong in the conversations, with only 69% of new referents falling into those two roles. The number of new referents presented in the 'Other' syntactic role is surprisingly high in the conversational data (19%) (see (81) and (82) above), yet low in the narratives (4%), indicating that there are genre differences with respect to preferred syntactic roles for introducing new referents. However, both genres are equal in the low rate of new A (7%), which is comparable (albeit slightly higher) to what other studies of Preferred Argument Structure have found (see Du Bois 2003:70).

These findings do not invalidate the constraints that have been proposed when describing Preferred Argument Structure, rather they indicate that Preferred Argument Structure is perhaps a little narrow in focus, and by excluding non-core, referential arguments, a large piece of the picture (i.e., the high number of new referents in the

'Other' syntactic role) has been missing from our understanding of Preferred Argument Structure. Additionally, this shows the need to look beyond narrative data and include different genres in the study of PAS.

Another way of looking at PAS is to examine the distribution of the total number of referents in each syntactic role in relation to activation status. Table 39 shows the proportion of given and new referents in each syntactic role for narratives and Table 40 for conversations.

Table 39: Syntactic role and activation status in Narratives (N=790)

	New		Given		Total	
A	5%	9	95%	171	100%	180
S	20%	68	80%	266	100%	334
Direct object	25%	45	75%	132	100%	177
Indirect object	-	-	100%	75	100%	75
Other	28%	5	72%	13	100%	18
Continuation of previous syntactic frame	50%	3	50%	3	100%	6
Total	16%	130	84%	660	100%	790

Table 40: Syntactic role and activation status in Conversations (N=1071)

	New		Given		Total	
A	17%	22	83%	110	100%	132
S	23%	110	77%	373	100%	483
Direct object	51%	114	49%	109	100%	223
Indirect object	10%	10	9%	93	100%	103
Other	50%	60	50%	61	100%	121
Continuation of previous syntactic frame	78%	7	12%	2	100%	9
Total	30%	323	70%	748	100%	1071

Here, the principles of PAS are evident in both the narratives and the conversations. In each genre a low number of NPs in the A role are used to present new information (5% in the narratives and 17% in the conversations). The S role is strikingly similar between the genres (20% new in the narratives and 23% new in the

conversations), yet the O role shows a considerable genre difference, with 51% of direct objects in the conversations being new, yet only 25% in the narratives. This suggests that while both roles are productive for introducing new referents, the larger number of new referents in the conversations is being introduced as O (and ‘Other’), not S. This also indicates is that in the conversational data, where information pressures are low, subjects of both transitive and intransitive verbs pattern similarly in their proportion of new and given referents, but in the narratives, with higher information flow pressures, the S and the O roles are more similar. This is consistent with Du Bois’ (1987) prediction that higher information flow pressures result in a stricter adherence to PAS, as evidenced in the more ergative patterning.

Regarding the 'Avoid lexical A' constraint (Du Bois 2003:71), as Table 41 shows, nearly a quarter of subjects of transitive verbs in the narratives are lexical, and 27% in the conversations are lexical. While it is true that overall, most subjects of transitive verbs are not lexical, there do appear to be frequent contexts in which lexical A is permitted, which are further examined.

Table 41: Form of expression for subjects of transitive verbs (N=312)⁷³

	Narratives		Conversations	
Full NP	40	22%	35	27%
Pronominal or unexpressed	140	78%	97	73%
Total	180	100%	132	100%

One recurring context (15/40) for lexical A in the conversations is with a pronominal, given object, as in (93). This is also seen in the narratives (10/35 lexical As

⁷³ This shows only given and new referents.

follow this pattern). These types of lexical As do not violate the 'avoid more than one lexical core argument' (Du Bois 2003:71) constraint.

(93) Lexical A with pronominal object

J: **el Roddy** me estaba entrevistando porque soy <VOX famoso VOX>.

J: **Roddy** was interviewing me because I am <VOX famous VOX>.

(Jugo, 390)

Another common context in which a lexical A is permitted is when the object is non-referential, as in (94), where the NP *laptop* is non-referential, and in Verb-Object Compounds (cf. Thompson and Hopper 2001:46), as in (95). This context is more common in the conversations (21/40 lexical As) than in the narratives (8/35). This reinforces the importance of distinguishing between discourse referential and non-referential NPs, and perhaps the constraint could be modified to indicate that speakers tend to avoid 'more than one lexical *referential* core argument'.

(94) Lexical A with non-referential object

A: **el Fabri** también quiere una laptop,

A: **Fabri** also wants a laptop,

(Clases, 526)

(95) Lexical A with V-O Compound

... (2.0) y **el niño sigue su camino**.

... (2.0) and the boy **goes on his way**.

(PS13:56)

Lastly, a closer look at the IUs containing lexical As in the narratives reveals that a significant proportion of these IUs (17/40, 43%) are truncated (see (96), the truncation is indicated by two dashes '--'). It is hypothesized that truncation may be another measure of high information flow pressures. An additional eight of these lexical As appear in a

different IU than the direct object, indicating that it is difficult for speakers to produce a lexical A and direct object in the same IU. When the proportion of lexical As that are found in truncated IUs (17/40, 43%) is compared to the proportion of other NPs in the narrative data that occur in truncated IUs (214/1188, 18%) it shows that the number of lexical As occurring in truncated IUs is significantly higher (Chi-square = 7.656, $p < 0.01$). This supports the notion that lexical As are a manifestation of higher information flow pressures.

(96) Lexical A in truncated IU

(H) **un niño** lleva sus frutos en --
.. en una bicicleta,

(H) **a boy** carried his fruit on --
.. on a bike,

(PS14:34-35)

In sum, this section has shown that, in general, the constraints of Preferred Argument Structure are evident in both genres under consideration here, but that violations of the proposed constraints are different between the genres. Specifically, a higher proportion of new arguments are introduced in and O and 'Other' roles in the conversations than in the narratives, O and S pattern similarly in the conversations and S and A in the narratives, and the 'violations' of the 'Avoid lexical A' constraint demonstrate different characteristics between the genres.

6.5 Summary

6.5.1 *Given vs. new referents*

The first difference noted in this chapter between given and new referents was found in their overall frequency—given referents are more than five times more frequent than new

referents in the narratives (84% of all referential NPs in this genre are given), and nearly two and a half times more frequent than new referents in the conversations (70%). The data also show marked differences in the preferred grammatical roles of new and given referents. New referents are unlikely to be introduced as subjects of transitive verbs (only 7% of new referents in each genre), but as they persist in the discourse, they often go on to occur in this syntactic role (26% of given referents in the narratives 15% of given referents in the conversations are subjects of transitive verbs). In addition, many more new referents are direct objects (35% in both the narratives and the conversations) than given referents (20% in the narratives and 15% in the conversations). On the other hand, there are more given indirect objects (11% in the narratives and 12% in the conversations) than new indirect objects (no new indirect objects in the narratives, and only 3% of the new NPs are indirect objects in the conversations).

Lastly, Table 39 and Table 40 compare the proportion of referential NPs according to information status and syntactic role in each genre. In each genre, when compared to the overall proportion of new referents, direct objects, syntactic role 'other' and continuation of previous syntactic frame show higher than expected proportions of new activation status and subjects of transitive verbs and indirect objects show lower than expected proportions of new referents. A few genre differences do emerge. In the narratives, subjects of intransitive verbs show a slightly higher than expected proportion of being new, but are less likely to introduced new referents in the conversations. This shows that no syntactic role is exclusive to given or new referents, but that certain syntactic roles (direct object, other, continuation of previous syntactic frame and subject of intransitive verbs [only for narratives]) are more conducive to new referents, whereas

given referents are more likely to appear in other syntactic roles (subject of transitive verb, indirect object, and subject of intransitive verbs [in conversations]).

Regarding the subjects of intransitive verbs, there is a clear genre difference. There are more new subjects of intransitive verbs than given subjects of intransitive verbs in the narratives (52% of new referents are subjects of intransitive verbs, and only 40% of given referents occupy the same syntactic role), yet there are more given subjects of intransitive verbs in conversations than new subjects of intransitive verbs (50% of given referents in the conversations and 34% of new referents). This may be an indication of higher transitivity in the narratives, where more given referents go on to become subjects of transitive verbs.

The syntactic role 'Other' is twice as common among new referents than in given referents. This is related to the presentative function of this syntactic role, which is particularly common in the conversations.

Regarding the modification of new versus given referents that are full NPs, the indefinite article is preferred for new referents in the narratives, yet the definite article is preferred in the conversations (Table 25), which is linked to the higher identifiability of the referents in conversation. The definite article is the most common form of referent modification among given referents, and bare referents (given or new) are highly uncommon in both genres.

6.5.2 Genre-specific differences

Throughout this analysis, attention has been drawn to several genre-specific differences.. Many of the observed genre differences are very clearly a result of different information flow pressures and transitivity. Other observed differences are more accurately

reflections of different distributions of types of referents between the genres (i.e., more proper nouns in the conversations, more referential NPs in the narratives), than genre differences themselves.

A number of formal differences were observed that appear to reflect the higher information flow pressures. The much higher proportion of referential NPs in the narratives (see Table 15) reflects the speaker's discourse goals in the narratives—to introduce and track entities through the text, and the lower rate of referential NPs in the conversations reflects the lower information pressures of this genre. In the conversational data, as will be seen in the following two chapters, speakers use NPs for many purposes that are not related to the introduction and tracking of discourse entities. This fundamental genre difference is also seen in the rates of persistence of referents across the genres. New referents are significantly more likely to be persistent in the narratives than in the conversations (Table 18), although there is a lower overall proportion of new referential NPs in the narratives than in the conversations (Table 17), and more given referents in the narratives. This is linked to the greater topic continuity of narratives than conversational data (cf. Travis 2007). It was also seen that resources for tracking given referents, such as renaming them and referring to members of a set are more productive in the narratives than in the conversations (Table 36), and that the need for more condensed 'packaging' of the information in the narratives is evidenced in the higher proportion of third person NPs that are modified by relative clauses in the narratives (see Table 27).

Argument structure and the distribution of NPs across different syntactic roles are also different between the genres. The preferred syntactic role for new referents is

different—nearly one out of five new referents in the conversational data is introduced in the 'Other' syntactic role, whereas the more traditional subject of intransitive verb (52%) and direct object roles (35%) are preferred in the narratives. The 'Other' role proves more productive for given referents in the conversations (8%) than in the narratives (2%; see Table 32) as well, and will be further explored in Chapter 7. Section 6.4 showed that the narrative data demonstrate a much stricter adherence to PAS in terms of lower proportions of both New and Lexical A. In addition, ergative patterning (similarities between S and O roles) is evident in the narratives, but nominative/accusative patterning (similar S and A) is evident in the conversations. This suggests that under higher information pressures, the S role is more productive for introducing new information than under lower information pressures.

The presentative verbal lexicon is different between the genres. Copulas are frequently used as presentatives in the conversations; whereas motion verbs are the most common presentative verbs in the narratives (which may be an effect of the design of the film itself, rather than a true genre difference) (see Table 33).

The genres also differ in the frequency at which different types of referents are found. The conversations are higher in identifiable referents: the higher proportion of new referents modified by definite markers in the conversations (80% versus 57% in the narratives) reflects the type of referents present in the genre – specific, known, people (first names are often modified by the definite article in this dialect of Spanish, and mentions of relatives are also often accompanied by the definite article), and the higher rate of the indefinite article with new mentions (43% of new referents in the narratives are introduced with the indefinite article, vs. 12% in the conversations, see Table 24).

Transitivity differences were also uncovered and will be further explored in 7.7. The narratives appear to be higher in transitivity than the conversations, given the higher number given third person NPs in the A and O roles in the narratives (see Table 32) In sum, the different discourse needs and pressures that result from the two different genres lead to a number of measurable differences in the rates at which different grammatical resources are used (i.e., relative clauses, different argument structures, definite marking, etc.) and the rate at which referential NPs occur and persist; however there is no evidence that the linguistic conditioning of forms (i.e., the uses of full NPs vs. pronouns, use of articles) differs from genre to genre.

7 Discourse Non-referential NPs

Non-referential NPs make up nearly two thirds of the conversational data and slightly more than one quarter of the narrative data. Recall from Chapter 5 that non-referential NPs take many forms and syntactic roles—demonstrative, interrogative and other pronouns, as well as representing most non-core arguments and all general NPs (see Table 14). This chapter will show that non-referential NPs are markedly different in a number of ways from referential NPs, and in addition, non-referential NPs are not as homogenous of a class of NPs, as referential NPs are. This is reflected syntactically in the distribution of non-referential NPs across more syntactic roles than referential NPs. Therefore, the grammar of non-referential NPs is more varied; and in order to best understand the patterns, they are examined here according to syntactic role. Most of the comparisons in this are made across genres, but where relevant, non-referential NPs are compared to referential NPs. Most previous research on information flow has failed to take into consideration this fundamental difference between referential and non-referential NPs, and as will be seen here, non-referential NPs perform very different functions from referential NPs, which is reflected in their grammar in many ways; and in a comprehensive analysis of NPs, they should necessarily be considered separately.

The organization of this chapter is as follows. The first section is dedicated to non-referential NPs in core syntactic roles, and the following sections look at the different non-core roles (objects of prepositions, predicate nominals, adverbials and vocatives). Then, Section 7.6 provides a summary of non-referential NPs. In 7.7 the relationship between referentiality and transitivity is explored.

7.1 Discourse non-referential NPs in core roles

Thompson's 1997 study of the core vs. oblique distinction in English indicated that the majority of non-tracking referents are obliques, and the majority of tracking referents are core. Following her observation, we can expect that core non-referential NPs will be less frequent than core referential NPs, and given the fact that core non-referential NPs are less frequent, it may be the case that they perform somewhat specialized discourse functions. Table 42 appears to support Thompson's (1997) observation: the majority of non-referential NPs in the narratives (76%) and conversations (67%) are non-core arguments, although this may be an artifact of the coding of referentiality. However, it is interesting to note that a substantial number of non-referential NPs in each genre are core arguments (i.e., A, S, DO, IO), 24% in the narratives and 33% in the conversations. The proportion of non-referential NPs in the conversations that are core is significantly higher in the conversations (33%) than in the narratives (24%).

Table 42: Core vs. Non-core non-referential NPs (N=2424)

	Narratives		Conversation	
Core	85	24%	692	33%
Non-core	262	76%	1385	67%
Total	347	100%	2077	100%

Chi-square = 10.223, $p < 0.01$

As discussed in Chapter 5, non-referential core arguments are also general. An examination of the non-referential core arguments by syntactic role shows several trends (figures given in Table 43). Table 43 shows that the two genres pattern the same in terms of the proportion of non-referential core arguments that fall into each of the syntactic roles (A, S and O); in each genre the highest proportion of non-referential core NPs are Direct Object, followed by S, then A, and Indirect Objects.

Table 43: Non-referential core NPs by syntactic role (N=777)

	Narratives		Conversation	
A	1	1%	127	18%
S	32	38%	260	38%
Direct Object	52	61%	275	40%
Indirect Object	-	-	30	4%
Total	85	100%	692	100%

One noteworthy difference between the genres is the frequency of non-referential A—this is nearly non-existent in the narratives (1 token, and 1% of the non-referential core NPs), yet one in five of the non-referential core argument NPs in the narratives falls into this role (18%).⁷⁴ The majority of non-referential subjects of transitive verbs are unexpressed (85%, 108/127) examples of non-specific third person plural, as in (97), where the speaker is talking about being crowned as the winner of a pageant. A smaller number (7%, 9/127) are generic subjects, as in (98).

(97) Unexpressed non-referential A

A: me **van a hacer** la coronación.

A: (**they**) are going to crown me.

(Deportes: 135)

(98) Generic subject of transitive verb

E: **los hombres** me hacen sufrir a mí.

E: **men** make me suffer.

(Viejos Amigos: 91)

This genre difference seen in the different proportions of non-referential A may be related to the specific discourse needs of the conversational speakers. In conversations, people are more likely to make generalizations (cf. Thompson & Hopper 2001) and make

⁷⁴ While this appears to indicate that conversations are higher in transitivity than the narratives, it will later be argued that non-referential NPs are not as adequate of a measure of transitivity as referential NPs.

more use of the non-specific third person plural with transitive verbs. It should also be noted that in each of the above examples, it could be argued that the first person (object) is more topical than the third person (subject). These non-specific third person subjects may be more useful in the conversations, used to relate to the first or second persons, whereas in the narratives of the Pear Story third person specific (e.g., referential) subjects are typical and there is less of a need for non-specific (e.g., non-referential) third person subjects.

Non-referential subjects of intransitive verbs occur at the same proportion in each genre, 38%. As in the non-referential subjects of transitive verbs in the conversations, the non-referential NPs in the S role also serve to form generalizations, as evidenced in (99) and (100). In the narratives, a frequent function of non-referential S is to set the stage for the narrative and provide the listener with background information, as in (101) and (102)

(99) Non-referential full NP subject of intransitive verb

A: pero están caras **las laptop**.

A: but **(the) laptops** are expensive.

(Amanda, Clases, 562)

(100) Non-specific unexpressed subject of intransitive verb in conversation

A: ponte que allá **vendan** más barato.

A: suppose that **(they)** sell (it) cheaper there.

(Amanda, Clases, 706)

(101) Non-specific unexpressed subject of intransitive verb in narrative

el canto de los --
del gallo,
me da la impresión de que \emptyset es eh,
.. tipo seis de la mañana.

the song of the --
the rooster,
gives me the impression that **(it)** is eh,
.. like six in the morning.

(PS43:12-16)

(102) Impersonal subject of intransitive verb in narrative

∅ se escuchaba que venía un ternerito,
(**one**) heard a calf coming, (lit. ∅ was heard that a calf was
coming)

(PS36:19)

A typical function of the non-referential direct objects in these both genres of these data is to form part of naming an event, activity or situation, as in (103) and (104). These direct objects commonly occur with highly frequent yet semantically 'weak' (and highly polysemous) verbs, such as *dar* ('give'), *hacer* ('make, do'), *poner* ('put') and *tener* ('have'). In fact, nearly one half (45%, 125/275) of all non-referential direct objects in the conversational data are objects of one of these four verbs, and a similar proportion (46%, 24/52) in the narratives are objects of these verbs. These are what Thompson and Hopper (2001:46) refer to as 'V-O' Compounds, and it could be argued that they are not in fact direct objects, but are rather part of an intransitive predicate.

(103) Direct object naming an event in conversation

E: .. y como no hacemos **ejercicio**,

E: .. and since we don't do **exercise**,

(Eduardo, Food, 315)

(104) V-O compound in narrative

.. (H) éste bajó del árbol y ∅ se dio **cuenta** recién en ese momento que le faltaba una cesta.

.. (H) he climbed down from the tree and (he) realized (lit. gave himself **account**) at that moment that he was missing a basket.

(PS41:48-49)

Another interesting genre difference emerges when the realization of subjects is examined. The conversations show proportionally over two times as many unexpressed non-referential subject NPs as the narratives (21% and 47% respectively, see Table 44). The majority (70%, 128/182) of the unexpressed subjects in the conversations are non-specific third person plural, as in (105). This is a clear genre difference, as the non-specific third person plural is very common in conversation, yet is less so in narratives.

Table 44: Subject realization of non-referential NPs

	Narratives		Conversations	
Full NP	14	42%	125	32%
Pronoun	12	36%	80	21%
Unexpressed	7	21%	182	47%
Totals	33	100%	387	100%

(105) Non-specific third person plural

- A: ... cla=ro --
 es que ahora vienen --
 en una misma cámara,
 o sea es ca- --
 es filmadora y [cámara] digital.
- C: [mhm],
 claro.
 viene con todo.
- A: pero \emptyset me han dicho que no es tan bueno,
- A: ... of course --
 it's that now they com --
 in just one camera,
 well it's ca- --
 it's a video recorder and digital [camera].
- C: [mhm],
 of course.
 it comes with everything.
- A: but (**they**) have told me that it's not so good,
- (Clases, 447-455)

When the realization of object non-referential NPs is examined, as in Table 45, we see that full NPs are the preferred form in each genre, 81% in the narratives and 68%

in the conversations. This is very different than what was seen for the subject non-referential NPs, where the minority were full NPs (42% in the narratives and 32% in the conversations, see Table 44). The genre difference is not significant; it is very likely that most of the non-referential objects are part of V-O compounds, in which they are unlikely to appear in reduced form.

Table 45: Realization of object non-referential NPs

	Narratives		Conversations	
Full NP	42	81%	208	68%
Pronoun	10	19%	97	32%
Totals	52	100%	305	100%

Chi-square = 2.774, $p = 0.10$ (not significant)

The modification of non-referential NPs in core syntactic roles (Table 46) shows that there are no important genre differences in the modification of full non-referential NPs for these referents. The most notable finding emerges when they are compared to referential NPs. There is a very high proportion of bare non-referential NPs in core syntactic roles (43% in the narratives and 45% in the conversations) when they are compared to referential NPs, which show a low overall instance of bare marking (2-12% for given NPs, 0-8% for new NPs). This supports Hopper and Thompson's (1984) observation that NPs that perform discourse functions that are not prototypical (i.e., NPs that do not track entities and are non-referential) will show less of the grammar associated with NPs (i.e., determiners).

Table 46: Modification of full non-referential NPs in core syntactic roles

	Narratives		Conversations	
Definite marker	27	48%	141	42%
Indefinite article	5	9%	46	14%
Bare	24	43%	151	45%
Totals	56	100%	338	100%

In sum, this section on non-referential NPs in core syntactic roles has shown that there are measurable genres differences and that the discourse functions performed by these NPs are different from the discourse functions of referential NPs in the same syntactic roles. Whereas S and O are productive roles for introducing new referential NPs and align in that sense, non-referential S and O diverge in their functions, and in conversations, there is more alignment in non-referential S and A (e.g., NPs in these roles are useful for formulating generalizations). Between the genres, it was observed that non-referential A is nearly non-existent in the narratives, but is productive in the conversations for making generalizations. When the realization of subjects was examined, it was found that unexpressed subjects are more frequent in the conversations than in the narratives, with the third person plural impersonal subject being a very frequent construction in the conversations. Lastly, no genre differences emerged when modification of full NPs was examined, but when compared to referential NPs, it is clear that bare NPs are much more common among the non-referential NPs.

7.2 Objects of prepositions

Of the non-core non-referential NPs, the majority in each genre are objects of prepositions (see Table 47). However, the proportion of non-referential NPs that are objects of prepositions is much higher in the narrative data (88% vs. 50%). This is a reflection of the higher information flow pressures of the narratives, as prepositions are seen as a grammatical resource for speakers to consolidate high amounts of information into a smaller number of units (cf. Biber 1988, Chafe 1982, 1985).

Table 47: Non-core non-referential NPs and syntactic role (N = 1647)

	Narratives		Conversation	
'Other'	11	4%	222	16%
Continuation of previous syntactic frame	2	1%	37	3%
Prepositional Phrase	231	88%	698	50%
Adverbial Phrase	11	4%	118	9%
Interrogative	-	-	83	6%
Vocative	-	-	120	9%
Predicate Nominal	7	3%	107	8%
Total	262	100%	1385	100%

These data were examined by preposition to see if different patterns of usage would emerge that would explain the difference in proportion, and the frequency of different prepositions is given in Table 48. As the table shows, the distribution of each preposition is similar between the two genres. It appears then, that the much lower rate of non-referential NPs that are objects of prepositions in conversations may be attributed to two factors—lower information flow pressures (and therefore, fewer NPs modified by prepositional phrases, cf. Chafe 1982, 1985) and the fact that non-referential NPs are more widely distributed among the different syntactic roles in the conversation than in the narratives.

Table 48: Frequency of prepositions in Narratives and Conversations (N=929)

	Narratives		Conversation	
<i>a</i> ('to, at')	34	15%	133	19%
<i>con</i> ('with')	47	20%	81	12%
<i>de</i> ('from, of')	48	21%	164	23%
<i>en</i> ('in, on')	73	32%	177	25%
<i>para</i> ('for')	3	1%	29	4%

<i>por</i> (‘for, by’)	18	8%	60	9%
other	8	3%	54	8%
Total	231	100%	698	100%

Functionally, NPs that are objects of prepositions primarily serve for orientation, spatially or temporally, as in (106) and (107), but may also perform other functions, including (but not limited to) indicating accompaniment ((108)) or cause ((109)).

(106) Spatially orienting object of preposition

le pone en **la parte de adelante**,

(he) puts it in **the front part**,

(PS13:26)

(107) Temporally orienting object of preposition

a **las doce del día** ∅ les decimos chau chau.

at **twelve noon** (we) tell them bye bye.

(Birthday: 150)

(108) Object of preposition expressing accompaniment

.. ∅estoy con **el Rodri**,

.. (I) am with **(the) Rodri**,

(Fumar:555)

(109) Object of preposition expressing cause

L: [el Cristian],
mi compañero,
se ha intoxicado.

.. dos días ∅ no ha venido.
[intoxicado por la chuleta].

E: [∅ debe ser por la comida de] ahí,

L: (the) Cristian,

my colleague,
 got sick.
 .. two days (he) hasn't come.
 [poisoned by **the pork chop**].
 E: [(it) must be due to the food] there,

(Cafe:1416)

Grammatically, we can see that NPs that are objects of prepositions tend to be full NPs (Table 49, 90% in narratives and 88% in conversations) and that of those full NPs, most are modified by a definite marker (see Table 50; 72% in narratives, 59% in conversations). Interestingly, there does appear to be a genre difference in the proportion of NPs that are objects of prepositions that are bare (36% in the conversations, yet only 10% in the narratives) and that are modified with an indefinite article (18% in the narratives and only 5% in the conversations). The higher proportion of bare NPs in the conversations can partially be explained by the frequent occurrence of phrases with bare NPs such as *de ley* ('for sure', lit. of law), *por favor* ('please', lit. by favor), and *por ejemplo* ('for example').

Table 49: Realization of object of preposition NPs

	Narratives		Conversations	
Full NP	207	90%	615	88%
Pronoun	24	10%	83	12%
Totals	231	100%	698	100%

Table 50: Modification of full NPs that are objects of prepositions

	Narratives		Conversations	
Definite marker	167	72%	362	59%
Indefinite article	42	18%	32	5%
Bare	22	10%	221	36%
Totals	231	100%	615	100%

To summarize, genre differences are evident in the overall frequency of NPs that are objects of prepositions (more common in the narratives) and in the modifiers that occur with these NPs.

7.3 Predicate nominals

Predicate nominal NPs occur primarily with the Spanish copula *ser* (the other copula, *estar*, is more frequently found with predicate adjectives), and function to classify or define referential NPs, as in (110). NPs in this role are more frequent in the conversations than in the narratives, but are not very frequent overall. Predicate nominals make up only 3% (N=7) of the non-core non-referential NPs in the narratives and only 8% in the conversations (N=107). The higher frequency in the conversations may reflect the subjective nature of conversations (cf. Scheibman 2001).

(110) Predicate nominal with *ser*

P: ... (3.0) pero el Tebi también,
buen estudiante ha sido.

P: ... (3.0) but (the) Tebi too,
 has been **(a) good student**.

(Café:793-794)

Grammatically, we can see from Table 51 that the preferred form of expression for predicate nominals is a full NP (100% in the narratives and 89% in the conversations). In terms of their modification, approximately half are modified and half are bare in each genre, with those that are modified being split nearly equally between definite and indefinite markers (see Table 52).

Table 51: Realization of predicate nominals

	Narratives		Conversations	
Full NP	7	100%	95	89%
Pronoun	-	-	12	11%

Totals	7	100%	107	100%
--------	---	------	-----	------

Table 52: Modification of predicate nominals

	Narratives		Conversations	
Definite marker	2	29%	20	21%
Indefinite article	2	29%	23	24%
Bare	3	43%	52	55%
Totals	7	100%	95	100%

In sum, the only notable genre difference with respect to predicate nominals is the frequency with which they occur; they are more frequent in the conversations than the narratives, reflecting the subjective nature of conversations. Grammatically, they are similar between the genres, showing a strong tendency to be expressed as full NPs, approximately half of which are bare and the remainder evenly divided into those which are definite and indefinite.

7.4 Adverbials

NPs that are syntactically found in adverbial phrases, such as (111) and (112), make up 4% of the non-core non-referential NPs in the narratives and over twice that amount (9%) in the conversations (cf. Table 47). They are realized almost exclusively as full NPs (Table 53), the majority of which are modified (Table 54).

(111) Adverbial NP in conversation

las dos se fueron **este fin de semana,**

both went **this weekend,**

(Deportes:73)

(112) Adverbial NP in narrative

`el momento que él estaba distraído recogiendo,`
`pasó un niño.`

`the minute that he was distracted picking (pears),`
`a boy passed by.`

(PS29:27-28)

Table 53: Realization of adverbial NPs

	Narratives		Conversations	
Full NP	9	82%	114	97%
Pronoun	2	18%	4	3%
Totals	11	100%	118	100%

Table 54: Modification of full adverbial NPs

	Narratives		Conversations	
Definite marker	5	56%	85	75%
Indefinite article	2	22%	17	15%
Bare	2	22%	12	11%
Totals	9	100%	114	100%

Functionally, these NPs orient the listener, and for this reason it is not surprising that they pattern more like NPs that are objects of prepositions (tendency to be full, definite NPs) than non-referential NPs that serve other discourse functions (i.e., predicate nominals, which tend to be bare, or non-referential NPs in the core-roles, which are less likely to be full NPs).

7.5 Vocatives

Unsurprisingly, vocatives are absent in the narratives, but relatively frequent, at 9% of the non-referential non-core NPs (cf. Table 47), in the conversations. Vocatives are uniformly full NPs, and largely bare (Table 55), although 10% are modified by a possessive adjective, as in (113). Note that not all of the vocative NPs modified by

possessive adjectives are not proper nouns; there are an additional 50 bare vocatives that are not proper nouns, as in (114).

Table 55: Modification of vocatives

	Conversations	
Possessive adjective	12	10%
Bare	108	90%
Totals	120	100%

(113)⁷⁵ Vocative NP modified by possessive adjective

F: ... hola **mi vida**.
 cómo está,
 cómo [le fue ahora]?
 A: [hola **mi amorcito**].
 bie=n,
 y a ti?

F: ... hello **my life**.
 how are you,
 how [did it go today]?
 A: [hello **my love**].
 goo=d,
 and you?

(Birthday:4-9)

(114) Bare vocative NP (not proper noun)

y quién llegó primero?
 no me acuerdo,
loco.
 and who arrived first?
 I don't remember,
dude. (lit. 'crazy')

(Fumar:201-203)

Perhaps one of the most interesting comparisons can be seen when we look not just at vocatives, but at all of the proper names in the corpus. As Table 56 shows, referential proper names are largely marked with a definite marker (92%), and non-

⁷⁵ Note that the use of the determiner is not optional here.

referential proper names are largely bare (73%). Examples (115) and (116) exemplify this difference.

Table 56: Modification of proper names

	Referential		Non-referential	
Definite marker	108	92%	23	27%
Bare	10	8%	61	73%
Totals	118	100%	84	100%

(115) Bare non-referential NP

.. qué más que quieres \emptyset que \emptyset te cuente **Rodriguito**?
 .. what else do (you) want me to tell you **Rodriguito**?

(Jugo:300)

(116) Definite referential NP

hoy \emptyset le llamé a **la Alejandra**,
 today (I) called **(the) Alejandra**.

(Jugo:214)

7.6 Summary of non-referential NPs

When non-referential NPs are examined by syntactic role, it becomes clear that non-referential NPs are not a homogeneous class of NPs that contrasts with referential NPs, rather it is a set of many subgroups of NPs that perform different discourse functions.

Grammatically, we see differences among the different types. Overall, pronominalization is rare in non-referential NPs, except for subject NPs, where we see approximately one fifth (21%) to one third (36%, cf. Table 44) are realized as pronouns, and to a lesser degree non-referential object NPs (19% and 32% pronominal in the narratives and conversations respectively, cf. Table 45). Patterns of determiner type and presence also reveal differences between the types of non-referential NPs. On the one

extreme are the vocatives, which are fully 90% bare (cf. Table 55), and on the other extreme are the adverbial NPs, which are only 22% (narratives) and 11% (conversations) bare (cf. Table 54), and objects of prepositions (10% bare in the narratives, 36% bare in the conversations, cf. Table 50). The non-referential NPs in other syntactic roles (i.e., subjects, objects and predicate nominals) show a more even distribution between those that occur with and without determiners.

The distribution of non-referential NPs among different syntactic roles is also interesting. Two thirds of non-referential NPs in the narratives are concentrated in one syntactic role, objects of prepositions (67%), whereas the same syntactic role accounts for only one third (34% of the non-referential NPs in the conversations (cf. Table 57). The number of prepositional phrases has been linked to information flow pressures (cf. Biber 1988, Chafe 1982, 1985) therefore this is another indication of the higher information flow pressures of the narrative genre. Table 57 also shows that non-referential NPs have a wider variety of different discourse functions in the conversations (i.e., the larger number of non-referential subjects, vocatives) than in the narratives (where they mainly serve as props or to modify other NPs).

Table 57: Non-referential NPs and syntactic role (N = 1647)

	Narratives		Conversation	
A	1	<1%	127	6%
S	32	9%	260	13%
O	52	15%	305	15%

'Other'	11	3%	222	11%
Continuation of previous syntactic frame	2	<1%	37	2%
Prepositional Phrase	231	67%	698	34%
Adverbial Phrase	11	3%	118	6%
Interrogative	-	-	83	4%
Vocative	-	-	120	6%
Predicate Nominal	7	2%	107	5%
Total	347	100%	2077	100%

7.7 Referential NPs, non-referential NPs, and transitivity

In Chapters 5 and 6 and in this chapter we have seen how referential and non-referential core NPs pattern and how this relates to information flow and other discourse needs. In this section, direct comparisons between referential and non-referential NPs are made in order to examine the relationships between transitivity and genre, and transitivity and referentiality. The first two tables in this section, Table 58 and Table 59, provide different ways to measure the transitivity between the two genres. The difference in transitivity among third person subjects (both referential and non-referential) between the genres can be seen in the proportion of subjects which are subjects of transitive versus intransitive verbs between the genres; this is an indication of the higher degree of agentivity of 3rd person subjects in the narratives. As Table 58 shows, a significantly higher number of subjects in the narratives are subjects of transitive verbs than in the conversations (32% vs. 26%).

Table 58: Third person referential and non-referential subjects and verb transitivity in narratives and conversations (N = 1596)

	Narratives		Conversations	
Transitive	185	32%	263	26%
Intransitive	385	68%	763	74%
Total	570	100%	1026	100%

Chi-square = 265.144, $p < 0.0001$

It is possible that conversations and narratives are equally high in transitivity, but that transitivity is more common among first and second person subjects in the conversations than in third person. For this reason, the proportion of third person NPs that are direct objects in each genre was also examined, in order to account for third person direct objects in the conversations that may have first or second person subjects. As Table 59 shows, the NPs (referential and non-referential alike) in the narratives are more likely to be a direct object, which is another measure of transitivity (22% of all NPs in narratives are direct objects, versus 16% in conversations). These results combined then support a higher degree of transitivity in the narratives.

Table 59: Proportion of direct objects by genre (N=4418)

	Narratives		Conversations	
DO	275	22%	506	16%
Other syntactic role	953	78%	2684	84%
Total	1228	100%	3190	100%

Chi-square = 25.550, $p < 0.0001$

As Hopper and Thompson (1980) and Thompson and Hopper (2001) note, transitivity should be viewed as scalar, and the transitivity of a clause depends on a number of different factors. One of these factors is the referentiality of the NPs within the clause. Hopper and Thompson state (1980:253) that referential objects are more highly

affected (i.e., an indicator of higher transitivity) than non-referential objects. It follows that the referentiality of the subject of the clause may also have some bearing on the degree of the transitivity of the clause; specifically, a referential agent is likely to be an indicator of a higher degree of transitivity. Table 60 and Table 61 show that the NPs in transitive clauses (both the As and the Os) in the narratives are much more likely to be referential than in the conversations, but this is in part because the narratives are overall higher in referential NPs than the conversations (72% of NPs in narratives are referential, 35% of all NPs in conversations are referential). However, the conversations are clearly lower in transitivity, with less referential A and O (52% and 54% referential) than S (66% referential, see Table 61), again supporting the hypothesis that the narratives are higher in transitivity, and by extension, information flow pressure.

Table 60: Referentiality of NPs in transitive clauses in narratives (N=923)

	Referential		Non-referential		Total	
Subject of transitive verb	184	99%	1	1%	185	100%
Subject of intransitive verb	353	92%	32	8%	385	100%
Direct object	223	81%	52	19%	275	100%
Indirect object	78	100%	-	-	78	100%

Table 61: Referentiality of NPs in transitive clauses in conversations (N=1668)

	Referential		Non-referential		Total	
Subject of transitive verb	136	52%	127	48%	263	100%
Subject of intransitive verb	503	66%	260	34%	763	100%
Direct object	231	54%	275	46%	506	100%
Indirect object	106	78%	30	22%	136	100%

8 Additional Interactional Factors Results

In this chapter, the effects of interactional factors on form and distribution of NPs in the data are examined. There are two main objectives—to show the functions of NPs that go beyond information flow and the exchange or sharing of information, and secondly, to provide a comprehensive account of NPs in the ‘Other’ syntactic role, which account for a significant proportion of the conversational data (11%, 350/3190). First, a detailed analysis of NPs that perform ‘interactional functions’ not related to information flow as identified in the Methodology chapter is presented in Sections 8.1 through 8.4. These NPs were chosen to represent interactional functions because they occur in constructions that can be identified on formal grounds. In no way is it suggested that these are the only NPs that play an important role in interaction, rather these NPs were selected because they can be identified objectively, and therefore lend themselves to quantification and empirical study. In the following section, the functions of NPs in the syntactic role ‘Other’ from the data are examined, and in 8.6 there is an analysis of given full NPs which contradict patterns of realization predicted by information flow. All NPs in this chapter come from the conversational data, the narrative genre is not included here. The implications of the findings presented in this chapter are discussed in the last section.

8.1 Increments

Recall from 2.6 and 4.3.1 that ‘increments’ describe constituents that speakers add to their turn after reaching a point in which the turn could be considered prosodically, syntactically and interactionally complete (cf. Ford et al. 2002, Schegloff 1996). In (117) and (118), the increments *exámenes cada bimestre* (‘exams each trimester’) and *del eme*

pe tres a la computadora ('from MP3 to the computer') both occur after IUs that have final intonation (indicated by the period), showing that they occur after a turn is prosodically complete. The IUs preceding the increments, *el semestre se divide en tres bimestres* ('the semester is divided into three trimesters') and *para quemar* ('to burn') are also syntactically and pragmatically complete, that is, they do not need additional information for the interlocutor to understand what is being said. It is interesting to note in (118) that the IU *para quemar* ('to burn') is also an increment.

(117) Free constituent Increment

C: el semestre se divide en tres bimestres.
 .. **exámenes cada bimestre.**

C: 'the semester is divided into three trimesters.
 .. **exams** each trimester.'

(Clases:64-65)

(118) Extension Increment

R: ... sí.
 ... (2.0) como así te prestó esa cosa la Nancy?
 .. que Ø se estaban bajando música?

A: hm,
 eh --
 unas canciones.

R: .. [de dónde]?

A: [para quemar].
 ... **del eme pe tres a la computadora.**

R: ... yes.
 ... (2.0) how is it that Nancy lent you that thing?
 .. that (they) were downloading music?

A: hm,
 eh --
 some songs.

R: .. [from where]?

A: [to burn].
 ... **from MP3 to the computer.**

(Fumar:113-121)

Ford, Fox and Thompson (2002) identify two types of increments: extensions, which syntactically follow from the prior utterance (see (118)), and free constituents

((117)), which do not syntactically follow from the previous syntactic frame. These two types of increments serve different interactional functions in American English. Interactionally speaking, Ford et al. (2002) find that both kinds of increments serve to create new opportunities for recipient turn uptake in a conversation, that is, they are a resource that speakers employ when confronted with difficulties in turn taking. However, while extensions 'function as action continuations, adding more to the same turn action' (2002:33), free constituents (particularly Unattached NPs) typically accomplish a new action, usually an assessment or an indicator of stance that can provide clues to the interlocutor about what a preferred response may be. However, cross-linguistic research has shown that increments perform different functions in other languages. Kim (2007) finds that in addition to eliciting turn-uptake, increments in Korean are also used as an interactional buffer between speaker turns. Field (2007) shows that increments in Navajo have little to do with recipient (*or* interlocutor) turn-uptake; instead it is culturally expected that periods of silence may occur in a conversation without another speaker pursuing a turn.

Ford et al. (2002) also find an iconic relationship between the form and function of increments in American English. Extensions, which are syntactically fitted to the previous utterance, also continue the action of the utterance. Free constituents, on the other hand, which do not have a syntactic relationship with the prior utterance, perform a new action.

This study seeks to enhance our understanding of grammatical increments by investigating the form, semantic properties, and functions of NPs found in increments in Ecuadorian Spanish. It is not intended to be a comprehensive analysis of increments in

Spanish, rather it is focused on the role that NPs play in increments. As a result, increments that do not include NPs, as in (119), are not included in the analysis.

(119) Increment without NPs

M: [la pierna],
que me duele atrás.
(H) como estoy poniéndome las rodilleras,
entonces,
me parece que eso --
.. le impide a las venas.
.. que tengo variz.
.. **atrás.**

M: [the leg],
that hurts me behind.
(H) since I'm putting bandages on my knees,
so,
it seem like that --
.. impedes the veins.
.. that I have varicose veins.
.. **behind.**

(Café:142-149)

A total of 109 NPs involved in increments were identified in the conversational data. The overall proportion of NPs that are found in increments is rather small, 3.4% of all NPs are found in increments (109/3190), however they are relatively common compared to other interactional phenomena.

As noted above, cross-linguistically, increments have been found to perform a number of different functions (Ono & Couper-Kuhlen 2007). I am unaware of any work that has been done on functions of increments in Spanish, but I found that the functions of increments in these data do fit in with functions of increments that have been described for other languages—specifically to elicit turn-uptake, to indicate stance or give an assessment, and to perform a repair (cf. Ford et al. 2002, Couper-Kuhlen & Ono 2007). There is one additional, albeit rare, function—referred to here as 'speech act'.

An overwhelming majority of increments are geared toward eliciting turn-uptake (n=93, 85% of all increments). There are several ways in which this is realized. In some cases, as in (120), the speaker prompts an answer to a question they have posed when an answer was not forthcoming by the interlocutor (note the short pause following E.'s first line), or in other cases, as in (121), they answer their own question (though note in this case in overlap with the interlocutor). In other instances, the increments serve to elaborate upon or clarify something previously said, as in (122). Other increments serve to specify a referent, apparently in order to help their interlocutor identify the referent, exemplified in (123). Lastly, a number of repetitions were found, as in (124), which appear to function as stalling mechanisms or place-holders, while the speaker waits to see if another speaker will take a turn or if s/he should continue.

(120) Increment prompting an answer

E: quieres casarte con ella?
.. **de ley?**

R: .. claro.

E: do you want to marry her?
.. for sure? (lit. 'of law')

R: .. of course.

(Viejos Amigos:121-123)

(121) Increment answering speaker's own question

A: ya vino ayer,
no te acuerdas?
... que estábamos sentados?
... qué día fue que estábamos sentados?
R: [antes de ayer].
A: [**el lunes**].
.. ahí llegó.

A: he came yesterday,
don't you remember?
... we were sitting down?
... what day was it that we were sitting down?
R: [the day before yesterday].
A: [**Monday**].
.. that's when he arrived.

(Chicas:491-497)

(122) Elaborating increment

J: verás,
.. llegué,
y= los jefes de las agencias tenían que hacer una publicidad.
.. **de= hm,**
... **de la calificación del banco,**
que es doble A.

J: look,
.. I got there,
and the heads of the agencies had to do some publicity.
.. **of= hm,**
... **of the rating of the bank,**
which is double A.

(Jugo:104-109)

(123) Specifying increment

G: .. ah=.
... pero bueno,
que ya termine,
no?
.. **el colegio.**

G: .. ah=.
... but good,
that he is finishing,
right?
.. **high school.**

(Café:463-467)

(124) Repeated increment

- F: [pero=
de] ahí en si=
cada Baby Einstein,
.. tiene un nombre.
- A: [ya].
- F: [que] le diferencia.
hay un montó=n,
le cuento.
un [montón].
- A: [ya].
- F: .. yo elegí éste porque están para más niñitos,
- F: [bu=t,
so] from there,
each Baby Einstein,
.. has a name.
- A: [yeah].
- F: [that] distinguishes it.
there are a to=n,
I tell you.
a [ton].
- A: [yeah].
- F: .. I chose this one because they are more for little kids,
(Birthday:306-316)

The second most frequent type of increment in these data are those that indicate stance or give an assessment (n=10), as in (125) and (126). As seen in (126), the NP *mentira* (lit. '(a) lie') is used in this variety of Spanish to indicate that the previous utterance by the speaker is intended as a joke, and that the interlocutor should not take offense. In this case, Rodrigo is teasing Julio, whose coworker has a crush on him, but the feeling is not mutual.

(125) Assessing increment

- J: me pasé hecho el loco.
... **es que,**
qué hecho pedazos.
.. (H) oye,
Roddy,
- J: I was totally crazy.
... **it's just that,**
what a mess.
.. hey,
Roddy,

(Jugo:207-211)

(126) Stance-displaying increment

R: y tu pelada?
... [la recepcionista]?
J: [cuál pelada]?
R: <@ **no,**
mentira @>.
J: no es mi pelada,
es mi --
es mi [empleada].
R: [es mi] recepcionista.
no,
mentira.

R: and your girlfriend?
... [the receptionist]?
J: [what girlfriend]?
R: <@ **no,**
just kidding @>.
J: she's not my girlfriend,
she' my --
she's my [employee].
R: [she's my] receptionist.
no,
just kidding.

(Jugo:1089-1099)

A small number of increments (n=4) were shown to be a form of self-repair (cf. Couper-Kuhlen & Ono 2007) in which the speaker either replaces or corrects a previous referent. This is exemplified in (127), in which Señora Manuelita is talking about the kind of visa a relative needed in order to travel to the United States and is having trouble coming up with the correct wording. Paca first finishes Señora Manuelita's clause with the prepositional phrase *de visita* (interactionally speaking, this is a co-construction), then changes her wording to *de paseo*, which is an increment through which she continues her own turn following a point of prosodic, syntactic and pragmatic completion.

(127) Repairing increment

M: tenía que irse de ehm --
P: ... de visita.
 [**de paseo**].

M: she had to go on ehm --
P: ... on a visit.
 ... **on a trip**.

(Café:1240-1242)

The last type of increment found in these data are formulaic speech acts, with only two tokens, shown in (128), performing a thanking function, and (129), performing a welcoming function. Note that they both appear after pauses, and may be another strategy of extending a turn at talk to avoid long pauses, in addition to politeness strategies.

(128) Speech act increment

R: quieres juguito?
J: .. bueno.
 .. **gracias Rodriguito**.
 se te agradece con toda el alma.

R: do you want juice?
J: .. sure.
 .. **thanks Rodriguito**.
 I appreciate it from the bottom of my heart

(Jugo:73-76)

(129) Speech act increment

M: Gerardito,
 ... venga,
 qué gu=sto,
 .. qué gusto tenerles,
 como siempre.
 .. **bienvenidos a mi casa**.

M: Gerardito,
 ... come,
 what a pleasure,
 .. what a pleasure to have you all,
 as always.
 .. **welcome to my house**.

(Café:5-10)

8.1.1 *Forms of NPs in increments in Ecuadorian Spanish*

Ono and Thompson (1994) found that in a study of unattached NPs in English, the majority are non-referential, especially those which occur after the end of a turn, as increments do. While not all the NPs in increments in these data are unattached, they are largely non-referential (86%, 94/109) (as is evident from the preceding examples), which shows that referentiality is important not only for information flow considerations, but also for interactional purposes. Recall from Chapter 6 that 64% of all NPs in the conversations are non-referential (Table 15), so the rate for NPs in increments of 86% is significantly higher than would be predicted given the distribution of referential and non-referential NPs in the corpus. The increments containing referential NPs do not appear to perform different interactional functions than the increments containing non-referential NPs. The NPs in the increments in the previous examples are all non-referential; the NP *la recepcionista* ('the receptionist') in (130) is an example of a referential NP in an increment; in this case the function is to further specify a NP.

(130) Referential NP in an increment

R: y tu pelada?
... **la recepcionista?**

R: and your girlfriend?
... **the receptionist?**

(Jugo:1089-90)

The following tables, Table 62 through Table 64, examine the distribution of NPs in increments by syntactic role, realization and modification in order to examine patterns of use. The figures for NPs in increments by syntactic role are given in Table 62.

Table 62: Increments by syntactic role (N=109)

Object of preposition	43%	47
‘Other’	39%	42
All other roles	18%	20
Total	100%	109

The majority of NPs in increments are either objects of prepositions (43%, as in (131)) or unattached NPs (syntactic role ‘Other’⁷⁶, 39%; see (132)). A smaller percentage (18%) is found distributed among many other syntactic roles, including subjects, direct objects, adverbial phrases and vocatives. Note that the overall proportion of non-referential NPs which are objects of prepositions is 33% (682/2057), so the 43% here is much higher. The proportion of NPs in increments that are in the syntactic role ‘Other’ (39%) is also higher than expected given the overall proportion of NPs in the ‘Other’ syntactic role (11% [222/2057] of non-referential NPs are found in the ‘Other’ syntactic role and 11% [121/1071] of referential NPs are also found in this role). Therefore, a higher than expected proportion of both objects of prepositions and NPs in the syntactic role ‘Other’ are found in increments. This is not surprising, however, since by definition increments are non-main-clause continuations (Ford et al. 2002:16), which limits the number of core arguments that would fit the criteria to be an increment.

(131) Increment NP as object of preposition

A: [hm=].
 ... (2.0) pero en cuál terminal estará?
 ... (2.0) en **el de la Marín**?

A: [hm=].
 ... (2.0) but in what terminal could she be?
 ... (2.0) in **the one at la Marín**?

(Fumar:413-415)

⁷⁶ It should be noted that all unattached NPs fall into the syntactic role ‘other’, but not all NPs in syntactic role ‘Other’ are unattached NPs.

(132) Increment NP in syntactic role ‘Other’

E: .. yo también andaba loca,
oye,
por cambiarme,
cuando entré a la Universidad,
porque queda tan lejos.
.. [hora y media de viaje].

E: .. I was also going around crazy,
you know,
to move,
when I entered the University,
because it's so far away.
.. [hour and a half trip].

(Viejos Amigos:635-640)

The majority (88%) of NPs in increments are realized as full NPs, as shown in Table 63. This is likely related to high proportion of NPs in increments which are non-referential, which tend to be realized as full NPs (cf. Chapter 7)

Table 63: Increments by realization (N=109)

Full NP	88%	96
Non-personal pronouns ⁷⁷	7%	8
Personal pronoun	4%	4
Unexpressed	1%	1
Total	100%	109

The patterns presented in Table 64, which shows type of determiner for NPs in increments, are similar to the patterns seen in Chapter 7 for determiners of non-referential NPs in the conversations. The majority show definite marking (51%) or are bare (44%); the higher rate of bare NP is consistent with what was found for most types of non-referential NPs (when compared to overall rates of bare-marking for referential NPs).

⁷⁷ Includes demonstrative, neuter, interrogative and ‘other’.

Table 64: Determiners and full NPs in increments (N=96)

Definite marker	51%	49
Bare	44%	42
Indefinite article	5%	5
Total	100%	96

In summary, increments tend to be non-referential full NPs that do not differ in determiner presence from other non-referential NPs. The syntactic roles most associated with increments are objects of prepositions and the syntactic role ‘Other’. It was earlier reported (Chapter 6) that one important function of NPs in the syntactic role ‘Other’ was to present new information; here the diversity of this syntactic role is evident, as it is also a very productive role for interactional purposes.

8.1.2 Summary of NPs in increments

While there is a growing body of literature on increments (cf. Field 2007; Ford et al., 2002; Kim 2007; and Ono & Couper-Kuhlen 2007, among others), little work has been done on the form and function of NPs in increments. In a study of unattached NPs in conversational English, Ono and Thompson (1994) find that unattached NPs that come at the beginning of a turn tend to be referential, and unattached NPs that come at the end of a turn tend to be non-referential. As the data in this study show, the majority of NPs in increments, which by definition come after a place in which a change of turn is relevant, if not necessarily at the end of a turn, are non-referential.

The distribution of NPs in increments by syntactic role and definite marking is not surprising. As Ford et al. (2002) found, NPs in the syntactic role ‘Other’ (their ‘Unattached NPs’) are a common form for increments. The distribution of determiner

presence for increments is similar to that of non-referential NPs in general, which is attributed to the large proportion of increment NPs in these data that are non-referential.

Regarding the function of increments, the Spanish data presented here show that NPs are a resource for the speakers in these data to accomplish a variety of conversational actions, from performing speech acts to indicating stance and pursuing turn-uptake. Unlike Ford et al.'s (2002) study of increments, no relationship between form and function was observed in these data; that is no specific functions were associated more with either extensions or unattached NPs; the functions are distributed between the two. From examining non-prototypical NPs (e.g., non-referential NPs that are not core arguments) as has been done here, we gain a greater understanding of how NPs contribute managing the interactional pressures of conversation and see how they go beyond managing the flow of information.

8.2 Repetitions

There were 47 NPs involved in direct repetitions (i.e., the same IU was produced twice with no modifications) in the conversational data⁷⁸. The repetitions themselves were found to perform a wide variety of functions. Self-repetitions (N = 20) were often found when competing for a turn or where there was overlap as in (133), or for emphasis (see (134)); and self-repetition in the form of an increment to provide another CTRP (Complex Transition Relevance Place, cf. Ford and Thompson 1996) and signal that the turn is over, as in (135). Inter-speaker repetition (N = 27) is more likely to reflect social actions, especially agreement, acknowledgement or solidarity, as in (136).

⁷⁸ Note that the occurrence of repetition is certainly much higher in conversation than what is shown here. These data are not representative of the overall frequency of repetition in the corpus because repetitions

(133) Self-repetition competing for a turn

R: .. no me acuerdo.
... pero quién más vivía aquí?
.. habíamos --
éramos bastantes,
de chiquitos.
A: el [Caloy].
R: [es que el Lenin] también vivía aquí.
A: **el Caloy.**
R: .. el Lenin,

R: .. I don't remember.
... but who else lived here?
.. we were --
there were a lot of us,
little kids.
A: [Caloy].
R: [but Lenin] also lived here.
A: **Caloy.**
R: .. Lenin,

(Fumar:340-348)

(134) Self-repetition for emphasis

L: ... si tú comes esa comida,
∅⁷⁹ no es fea.
E: ... ya.
L: .. ∅ no es fea,

L: ... if you eat that food,
(it)'s not bad.
E: ... yeah.
L: .. (**it**)'s not bad,

(Food:345-348)

(135) Self-repetition to signal end of turn⁸⁰

A: (H) más le gustó lo que se ha ido a --
a Disney World.
F: a [Disney World]?
A: [y ha] --
y= --
.. y ha filmado todito.
ha filmado todito.

that do not include NPs are excluded, as are repetitions that are not produced exactly the same by the (two) speaker(s), and those that occur with more than one intervening IU.

⁷⁹ Recall that unexpressed NPs are included in this study.

⁸⁰ The repetition of *a Disney World* in this example was not included in these data because tokens were not extracted from Francisco's speech in this conversation. This inter-speaker repetition appears to serve as a means of referent negotiation.

A: what he liked the most was going to --
to Disney World.
F: to [Disney World]?
A: [and he] --
and --
.. and he filmed everything.
he filmed everything.

(Deportes:45-51)

(136) Repetition showing acknowledgement

E: ... (2.0) cómo se llama?
L: .. gastroenteritis.
E: .. **gastroenteritis.**
o la otra,
que te pones amarillo?
L: .. hepatitis.
E: .. **hepatitis.**

E: ... (2.0) what's it called?
L: .. gastroenteritis.
E: .. **gastroenteritis.**
or the other,
when you turn yellow?
L: .. hepatitis.
E: .. **hepatitis.**

(Food:616-622)

Regarding the form and semantic properties of NPs in repetitions, a few interesting patterns emerge. The number of non-referential NPs in repetitions (N = 19, 40%) is lower than expected given the overall proportion of non-referential NPs in the conversational data (64%); this indicates that information that is more important to the discourse (i.e., referential NPs) is more likely to be repeated. A significant proportion of repetitions is found in the 'Other' syntactic role (Table 65), but NPs in repetitions are also commonly found as subjects and objects of prepositions. The majority (66%) occur as full NPs (Table 66).

Table 65: NPs in repetitions by syntactic role (N=47)

'Other'	34%	16
Subject	34%	16
Object of preposition	19%	9
Object	6%	3
Adverbial	4%	2
Interrogative	2%	1
Total	100%	47

Table 66: NPs in repetitions by realization (N=47)

Full NP	66%	31
Unexpressed	23%	11
Pronominal	11%	5
Total	100%	47

8.3 Co-constructions

Co-constructions are utterances that are jointly constructed by two speakers. Wilkes-Gibbs (1986; see also Clark and Schaefer 1989 and Learner 1987) describes the observed steps that speakers follow when collaborating to construct an utterance. One speaker produces a syntactically incomplete utterance which may also have verbal cues (i.e., truncation) that indicate that they are experiencing difficulties in completing the utterance. At this time another speaker offers a completion, which the original speaker may later confirm or reject as acceptable. In order for a speaker to be able to complete another's utterance, speakers must be able to orient to the projectability of the sentence in progress. As Lerner (1991) points out, there are many conversational schemas that lend themselves to co-constructions, such as conditional sentences, in which the first speaker produces the *if*-clause and the second speaker chimes in with the *then*-clause. Burgos (2007) finds that speakers rely on shared cognitive schemas as well as syntactic (incompleteness) and prosodic cues (truncation) to identify opportunities for co-constructions in Spanish. Lerner describes co-constructed sentences as 'interactional

achievement(s)' (1991:441) and they may be produced under a variety of interactional circumstances, including, but not limited, to joint story-telling (see. (137)), helping a speaker find a word (example (138)), or even form a type of question-answer sequence, as illustrated in (139).

(137) Story-telling co-construction

A: ... no,
cogimos bus hasta .. <@ Tumbaco @>.
R: cogimos bus y nos bajamos en --
claro,
[en Tumbaco].
A: [**Tumbaco**].
y de ahí [seguimos trotando].

A: ... no,
we caught a bus to .. <@ Tumbaco @>.
R: we caught a bus and we got off in --
of course,
[in Tumbaco].
A: [**Tumbaco**].
and from there [we continued jogging].

(Fumar:188-194)

(138) Co-construction helping speaker find a word

M: y ha salido con o= --
con muy buenas eh,
G: **notas**.

M: and he ended up with o= --
with really good uh,
G: **grades**.

(Café:84-86)

(139) Question-answer co-construction

A: ya pasas a?
C: .. séptimo semestre.

A: now you're headed into?
C: ... seventh semester.

(Clases:67-68)

In each of these examples, speakers were able to predict that a NP would syntactically follow from where the first speaker left off. In .

(137) and (139) the speakers oriented to the preposition, which are followed by NPs. In (138), part of the NP was already produced, as seen in the adjective *buenas*, where the gender (feminine) and number (plural) help the interlocutor identify an appropriate referent.

There were 11 NPs involved in co-constructions in these data. The majority are full NPs (N = 8, see (140)), followed by pronouns (N = 3, see (141)). Over half (N = 6) are objects of prepositions, and of the remaining, four are in the syntactic role ‘other’ and one is a subject. Eight are non-referential (which is not independent of the fact that 6 are objects of prepositions), and three are referential (two given, one new).

It is also noted that co-constructions occur in what can be described as interactionally rich portions of conversation. In several examples from these data, the second portion of the co-construction is followed by a repetition of the final segment by the speaker who produced the first segment, as in (140). The lengthy pause (of two seconds) in this example also seems to indicate that Rodrigo expects Adrián to finish the sentence, as he makes no attempt to repair his truncated IU. In (141), a series of co-constructions are seen, with three different speakers collaborating in their description of Feidy.

(140) Full NP in co-construction

A: a mí me acuerdo que me gustaba la Carolina.
R: .. la Carolina un día se asoma por la terraza del --
A: ... (2.0) de **la casa de ella**?
R: de la casa de la Sara.

A: I remember that I used to like Carolina.
R: .. one day Caroline appears on the terrace of --
A: ... (2.0) of **her house**?
R: of Sara's house.

(Fumar:472-475)

(141) Pronoun in co-construction

P: .. @@
 la Feidy es [más] --
M: [la] --
P: más amigable,
 [más así].
M: [más amigable],
 [2más alegre2],
G: [2más alegre2],
 [3má=s3] --
M: [3más conversona3],
 más --
G: más suelta es **ella**.

P: .. @@
 (the) Feidy is [more] --
M: [(the)] --
P: more friendly,
 [more like that].
M: [more friendly],
 [2more happy2],
G: [2more happy2],
 [3mo=re3] --
M: [3more conversational3],
 more --
G: **she** is more open.

(Café:1303-1314)

8.4 Lists

Lists were identified in these data as three or more IUs produced in succession that are syntactically similar and form a coherent unit. The cross-linguistic existence of three-part lists has interesting cognitive and interactional implications.

Sánchez-Ayala (2003) examines Spanish list construction from both a cognitive and an interactional perspective. On the cognitive side, he argues that lists are a grammatical resource for the cognitive task of enumerating. On the more interactional side, he finds two types of interactional tasks that lists can complete—framing lists that establish shared background between speakers, and demonstrative lists that reinforce a

previous claim made by the speaker⁸¹. Jefferson (1990) and Lerner (1991) have shown how speakers orient toward and project a third component of a list, which aids in turn-taking and provides opportunities for co-constructions. The effects of list construction on NP realization are considered here.

There were 47 NPs identified as parts of lists. Nearly two thirds are non-referential (62%, 32/47), one fifth are given (17%, 8/47), and the remainder are new (15%, 7/47). These figures approximate the overall distribution of non-referential, given and new NPs in the data set, therefore there is no evidence that referentiality or activation state are correlated with the interactional task of constructing a list. The most notable trend seen in the data is continuity of expression of NPs in lists. Within a list, NPs show the same realization—unexpressed, as in (142), pronouns or full NPs, as in (143).

(142) List with unexpressed subjects

- A: ah=.
.. ø iba corriendo y shu=,
ø se resbala,
y ø se pega una matada,
y ø se raspó la rodilla,
te acuerdas?
- A: ah=.
.. (she) was running and shu=,
(**she**) slips,
and (**she**) takes a fall,
and (**she**) scrapes her knee,
do (you) remember?

⁸¹ There are other constructions that perform these interactional tasks as well, they are not exclusively found in lists.

(143) Lists with pronouns and full NPs

J: tú pones Julio dice **esto**,
Roddy dice **esto**.
.. **Junior** dice **esto**.
... **Junior** es vago.
R: más o menos.
... [ah=].
J: [Junior no] quiere nada de la vida,

J: you put Julio says **this**,
Roddy says **this**.
.. **Junior** says **this**.
... **Junior** is lazy.
R: more or less.
... [ah=].
J: [Junior doesn't] want anything out of life.

(Jugo:93-99)

The forms of expression of the unexpressed and pronominal NPs in (142) and (143) are not surprising, but the repeated full NP in (143) is unexpected, and will be discussed in further detail in section 8.6. Givón's (1983) finding of lesser coding for NPs with subsequent mentions where there is little distance between mentions does not apply to this example, rather this NP shows the more coding in the two IUs that follow the first mention. While it has been shown, here and elsewhere, that NP realization is often a reflection of information status, examples such (143) as indicate that NPs realization is also sensitive to interactional pressures, here, the formation of a list and the need for the speaker to indicate that they are creating a list by repeating a full NP over three consecutive clauses. As evidenced in examples (142) and the pronouns in (143), there is not always a conflict between information flow pressures and interactional needs, as the realization of the NPs in these examples are predicted by both information flow and

interactional factors, but the full NPs in (143) show that information flow and interactional factors can in fact ‘compete’ in conversation.

8.5 NPs in ‘Other’ syntactic role

NPs in the syntactic role ‘Other’ (i.e., those which are syntactically independent) are nearly four times as likely as NPs in core and oblique roles to perform the functions identified above as interactional (Table 67). The proportion of NPs in core and oblique roles that are involved in the interactional constructions identified here ranges from only 0-2% (indirect and direct objects) to 9% (objects of prepositions). The proportion of NPs in the ‘Other’ syntactic role participating in these interactional constructions is 20%. These have been discussed in 8.1 through 8.4, but what about the remaining 80% of NPs occurring in the ‘Other’ syntactic role that do not occur in these interactional constructions? Do they also perform interactional functions? Given that these NPs constitute a large proportion of the NPs in the conversational data (11% of all NPs), they should be accounted for in a full analysis.

Table 67: Proportion of NPs fulfilling interactional functions in conversations (N=3190)

	Interactional ⁸²		Non-interactional		Total	
‘Other’	20%	70	80%	280	100%	350
All other syntactic roles	5%	144	95%	2696	100%	2840

Chi-square = 108.602, $p < 0.0001$

The short answer is that most referential NPs in the syntactic role ‘Other’ serve to regulate information flow and the majority of non-referential NPs in the syntactic role ‘Other’ are performing interactional functions. As Table 68 shows, the majority (65%) of the remaining NPs in the syntactic role ‘Other’ are non-referential. As we saw in

Chapters 5 and 6, referential and non-referential NPs perform very different functions in conversation, and therefore merit separate analyses. The following two sections look at referential and non-referential NPs in the ‘Other’ syntactic role.

Table 68: Distribution of NPs in syntactic role ‘Other’ by referentiality (N=280)⁸³

Syntactic role ‘Other’		
Referential	35%	99
Non-referential	66%	181
Total	100%	280

8.5.1 Referential ‘Other’ NPs

When the referential ‘Other’ NP tokens are examined, nineteen (23%, 29/128) have interactional functions previously identified (one co-construction, nine increments and nine repetitions). A significant number of referential NPs in the ‘Other’ syntactic role present new information (47%, 60/128, see Table 69). We have seen that ‘Other’ is a key role for introducing new information in conversation, and it appears to be one of the major functions of this role, followed by the interactional functions. The remaining NPs (N=49) are discussed below.

Table 69: Activation status of referential NPs in ‘Other’ syntactic role (N=128)

Syntactic role ‘Other’		
New	47%	60
Given	48%	61
Renaming	5%	6
'Member of a set'	1%	1
Total	100%	128

⁸² Includes NPs involved in increments, lists, repetitions and co-constructions.

⁸³ ‘other’ NPs involved in previously identified interactional constructions (increments, repetitions, co-constructions and lists were excluded from this table.

Broadly speaking, the functions of these NPs are more related to information flow than to social or interactional needs. Two categories of discourse functions were identified: referent negotiation (N = 36) and answering a question (N = 9). For four tokens, the discourse function was unclear. The majority of NPs in the ‘Other’ syntactic role which were determined to be involved in referent negotiation serve the following purposes in discourse—interaction between speakers in order to further define or request clarification of a referent as in (144) or to elaborate upon a referent as in (145).

(144) ‘Other’ NP asking for clarification⁸⁴

J: me toca hace=r --
 ... eso de la carrera del so=l.
 R: .. **qué carrera del sol?**
 J: I have to do --
 ... that the race of the sun thing.
 R: .. **what race of the sun?**

(Jugo:890-892)

(145) ‘Other’ NP elaborating⁸⁵

R: [cuándo se va]?
 E: [@@@@@@@]
 en próxima semana,
 ya no sé.
 [2falta una semana2].
 R: [2qué fecha2]?
 E: el veinticuatro.
 R: .. cuándo cae veinticuatro?
 E: .. la próxima semana,
 el lunes.
 R: .. lu=nes?
 E: ajá,
el próximo lunes,
 ya,
 .. se va.
 ... y yo voy a llorar.
 R: [when is he leaving]?

⁸⁴ The NP in this example can stand alone as a question, yet is not syntactically integrated into a clause and therefore fits the criteria for consideration as a NP in the syntactic role ‘other’.

⁸⁵ This example contains more than one NP in the ‘Other’ syntactic role. The NP *la próxima semana* (‘next week’) is answering a question, *el lunes* (‘Monday’) is specifying and *lunes* (‘Monday’) is negotiating a referent.

E: [@@@@@@@]
 next week,
 I don't know.
 [2in a week2].
 R: [2what day2]?
 E: the twenty-fourth.
 R: .. when is the twenty-fourth?
 E: .. next week,
 Monday.
 R: .. Monday?
 E: uh huh,
next Monday,
 yeah,
 .. he leaves.
 ... and I'm going to cry.

(Viejos Amigos:535-550)

The next most common function is found in question/answer sequences, which accounts for nine referents. This is exemplified in (146).

(146) ‘Other’ NP in question/answer sequence

F: .. cuál de sus hermanas está en Baños?
 A: **las dos.**

F: .. which of your sisters is in Baños?
 A: **both.**

(Deportes:70-71)

As these examples illustrate, referential NPs in the syntactic role ‘Other’ are clearly related to the exchange of information—they clarify referents, give additional information about referents, and answer questions about referents, in contrast to the more interactional functions served by the non-referential NPs examined in the following section.

8.5.2 Non-referential ‘Other’ NPs

Of the non-referential NPs in the ‘Other’ syntactic role, approximately one fifth (18%, 41/223) occur in one of the previously identified interactional constructions, leaving 182 for further analysis. Approximately one third (36%, 66/182) of these have

functions that are related to information flow, such as referent negotiation and question/answer sequences, similar to the given NPs in the syntactic role ‘Other’. For the remaining 116, a number of interactional functions were identified, and no clear function was evident for eight NPs.

Many of the interactional functions appear to be related to regulating the flow of the conversation. The nouns *todo* ‘everything’ (n=20) and *nada* ‘nothing’ (n=12) very commonly perform this function, as seen in (147) and (148) below. In (148) Gerardo uses *todo eso* (‘all that’) to show that the topic is closed. Note that *todo eso* also functions similarly to a tag question with the respect that they both play important roles in turn-taking. Schegloff (1996) points out the important role that tag questions play in American English to signal the end of a turn, and *todo eso* here functions much in the same way—note that following *todo eso*, Gerardo’s turn can be considered syntactically, intonationally and pragmatically complete (or CTRP, cf. Ford and Thompson 1996), hence an ideal place for a turn-transition to occur (although a turn transition does not in fact occur in this example). In (148) *nada* is found in a question-answer sequence that is used to guide the topic of a conversation.

(147) Interactional *todo*

G: ah,
que la gente cuando va,
... (H) primero creo que va al cementerio allá,
... a conoce=r,
y **todo eso**.
.. ah ah.
.. ah,
entonces,
para el sab- --,
pero usted se va el sábado,
y regresa cuándo?

G: ah,
so when the people go,
... (H) first I think they go to the cemetery there,
... to see it,
and **all that**.
.. ah ah.
.. ah,
so,
on Sat- --,
but you leave on Saturday,
and come back when?

(Café:308-318)

(148) Interaccional *nada*

J: .. ø te van a pagar?
R: .. n=ø sé.
J: .. sí,
ø te han de pagar.
R: qué más Julito?
J: .. [nada].
R: [ø te encontraste] con la pelada del Juano?

J: .. are (they) going to pay you?
R: .. I don't know.
J: .. yes,
they are sure to pay you.
R: what else Julito?
J: .. [**nothing**].
R: [did (you) run into] Juano's girlfriend?

(Jugo:188-194)

Eso (especially in the collocation *eso sí* 'that's true') is another common non-referential NP found in the 'other' syntactic role (n=21). It is often used to show agreement or disagreement, as in (149).

(149) *Eso* showing agreement

- C: claro,
es .. una carrera de bastante sacrificado pero,
.. bien pagado,
también dice.
ojalá.
[@@]
- A: [no],
eso sí,
sí es l- --
bien pagada es esa carrera.
para qué también.
- C: .. [ojalá,
mhm].
- A: [tienes que ubicarte] en una buena empresa,
nada más.
- C: sí,
eso también.
- C: sure
(it) is .. a career with a lot of sacrifices but,
.. well paid,
they also say.
I hope.
[@@]
- A: [no],
that's true,
yet (it) is l- --
well paid that career.
with good reason too.
- C: .. [I hope,
mhm].
- A: [you have to get into] a good company,
that's all.
- C: yes,
that too.

(Clases:193-209)

Other non-referential NPs in the ‘Other’ syntactic role serve to show speaker stance, attitude, or reaction toward a previous statement. Many of these function similarly to increments, occurring at the end of a turn, but they do not necessarily follow final intonation, as increments do. The NP *mentira* (‘lie’), which was shown in (126) as an increment, is also often found after continuing intonation, where it is used to show that a speaker has said something sarcastically or jokingly, as in (150). In this example the speakers are discussing laminate wood flooring, in Spanish called a ‘floating floor’, when

R characterizes a laminate floor as a floor that floats, then in a laughing voice says *mentira* 'lie'. The formula *qué* + NP ('what a NP') is also a common device to show attitude or stance, as in (151).

(150) *Mentira*

R: qué es piso flotante?
J: o sea,
es un piso así que vos [le mojas],
R: [que flota].
no,
<@ **mentira** @>.

R: what is floating floor?
J: well,
it's a floor that you [get it wet],
R: [that floats].
no,
<@ **lie** @>.

(Jugo:427-432)

(151) *qué* + NP

E: oh=,
te acuerdas cuando jugábamos carnaval con el huevo podrido?
R: .. (H) no,
no recuerdo.
E: qué a=sco,
Rodrigo,

E: oh=,
do you remember when we played carnaval with rotten egg?
R: .. (H) no,
I don't remember.
E: how disgusting, ((LIT. 'what a disgust')
Rodrigo,

(Viejos Amigos:278-283)

8.5.3 Summary of 'Other NPs'

As this section has shown, NPs frequently occur in the syntactic role 'Other' in these data, and here again we find referential and non-referential NPs behaving very differently. Referential NPs in the syntactic role 'Other' are likely to perform functions related to information flow, such as introducing new referents, negotiating referents and

answering questions. Non-referential NPs in the syntactic role ‘Other’ can perform these functions, but more often have an interactional purpose, such as regulating the flow of the conversation, showing stance, attitude and agreement.

8.6 Given Full NPs

As was seen in Chapter 5, given activation status and topic continuity do not always lead to lesser coding of an NP. Recall that in the conversational data, 27% (201/748) of given NPs are expressed as full NPs, and over two thirds of those occur within 2 clauses of the previous mention. In this section, seven examples of such NPs are discussed in detail and factors beyond information flow are explored as they affect the realization of these NPs.

In (152) (see also (143) above), Julio and Rodrigo are discussing how Rodrigo transcribes the conversations. There are two lists in this excerpt—the first one consisting of three different subject nouns (Julio, Roddy, Junior) but the exact same verb phrase (*dice esto*), and the second one with the same subject noun (Junior) but different verb phrases. Parallel structures are common in lists, as this is one cue to the listener that a speaker is creating a list. Therefore, the three consecutive realizations of Junior as a full NP can be understood as a speaker strategy that tells his interlocutor that he is creating a list about a person, rather than tracking the referent itself.

(152) Full given NPs in a list

J: .. o sea,
tú pones Julio dice esto,
Roddy dice esto.
.. Junior dice esto.
... **Junior** es vago.
R: más o menos.
... [ah=].
J: [**Junior** no] quiere nada de la vida,
mentira.

J: .. so,
 you put Julio says this,
 Roddy says this.
 .. Junior says this.
 ... **Junior** is lazy.
 R: more or less.
 ... [ah=].
 J: [**Junior** doesn't] want anything out of life,
 just kidding.

(Jugo:92-100)

Lists are not the only interactional construction that influence NP realization. NPs that are repeated for interactional purposes, as *un rico cafecito* in (153) below, show the same realization. This repetition occurs after a pause and immediately precedes a shift in topic, suggesting that the repetition is functioning to avoid a lull in the conversation, or perhaps an acceptance of a contribution (cf. Clark & Schaefer 1989).

(153) Full NP in repetition.

M: vamos a servirnos un rico café [ahora] que nos hace frío.
 G: [XX]
un rico cafecito.
 M: .. **un rico cafecito.**
 G: y el Richi no viene?
 M: pueda que ø ya venga.

 M: let's have a nice cup of coffee [now] that it's cold.
 G: [XX]
a nice little cup of coffee.
 M: **a nice little cup of coffee.**
 G: and Richi isn't coming?
 M: it's still possible that (he) will come.

(Café:386-391)

Full given NPs can also occur if the speaker feels that the NP was not activated the first time it was mentioned, as occurs when speakers' turns overlap with each other. In (154), Rodrigo and Adrián are jointly constructing a list of their childhood friends from the neighborhood. Their turns frequently overlap in this excerpt precisely where friends are mentioned for the first time. As a result, several names (Caloy, Lenin, and *un orejón* ('a big-eared kid')) are repeated as they may not have been activated the first time

they were mentioned. Note that it is the same speaker that produces the NP twice. The NP *César* was also produced in overlap, but here Rodrigo repeats the NP, acknowledging that he did in fact hear the name and it was activated as part of the list.

(154) Full given NP due to overlap

R: ... pero quién más vivía aquí?
.. habíamos --
éramos bastantes,
de chiquitos.
A: el [Caloy].
R: [es que el Lenin] también vivía aquí.
A: **el Caloy.**
R: .. **el Lenin,**
nos llevábamos con el Lenin.
...(2.0) el Carlos.
.. vivía el Edgar,
el Suco.
...(2.0) (H) también [salía]--
A: [el César].
R: **el César.**
y salía el Santiago,
hermano de la Lisset,
también [jugaba] con nosotros.
A: [ah].
.. también --
...(2.0) también= .. salían a jugar --
sabía jugar el Mario,
el Ricardo,
R: .. [claro].
A: [el hermano] del Lenin --
[2del Darío2].
R: [2y había un orejón2],
y --
.. **un orejón había,**

R: ... but who else lived here?
.. we were --
there were a lot of us,
little kids.
A: [Caloy].
R: [it's that Lenin] also lived here.
A: **el Caloy.**
R: .. **Lenin,**
we hung out with Lenin.
...(2.0) Carlos.
.. Edgar lived here,
Suco.
...(2.0) (H) also [there was]--
A: [César].
R: **el César.**
and Santiago would come out
Lisset's brother,

also [played] with us.
A: [ah].
.. also --
...(2.0) also .. would play --
Mario would play,
Ricardo,
R: .. [of course].
A: [the brother] of Lenin --
[2of Darío2].
R: [2and there was a big-eared kid2],
and --
.. **there was a big-eared kid,**

(Fumar:340-369)

It has been suggested by (Givón 1983, Hochberg 1986) that subjects may be expressed in order to avoid ambiguity that arises when there are two semantically compatible nouns in the context that could potentially compete for the subject role. The full NP *Ceci* in the last line of (155) does appear to fulfill this function. In this example, there are two third person feminine animate NPs (Alexandra and Ceci) which could both potentially be the subject of the verb *había dicho* ('had told'). In this case, pronominal or unexpressed realization could result in confusion, so the choice of a full NP is clearly beneficial for referent tracking in this instance. In addition, note that the immediately preceding mentions of Ceci were as indirect object (*le, a la Ceci*) and object of a preposition (*con la Ceci*); work by Paredes Silva (1993) on Brazilian Portuguese has indicated that change in syntactic role may influence subject expression. Note that the change in setting may also influence the NP form.

(155) Given NP in different syntactic roles

M: ... (2.0) (TSK) la Alexandra debería hablar con la Ceci.
P: ... (H) sabe que sí le ha dicho,
creo,
alguna vez ella a la Ceci,
verá.
.. (H) pero **la Ceci** le había dicho que ya no hay como,
M: ... (2.0) (TSK) Alexandra should talk with Ceci.

P: ... (H) you know she has said something to her,
 I think,
 one time she (did) to Ceci,
 see.
 .. (H) but **Ceci** told her that there is no way,

(Café:707-712)

While some subjects may be expressed in order to resolve ambiguity, this is not always the underlying motivation behind an expressed given NP, as evidenced in the following examples. In (156), the NP *el Suco* occurs as a full NP upon first mention and again as a full NP two clauses later. In this example, as in the previous, the NP does change syntactic role, but it appears that the change of conversational action (i.e., listing all the people who have left to talking about missing one of those people) is more influential in the choice to express the NP as a full NP. This instance of realization of *el Suco* as a full NP could also be partly attributed to potential inference with the semantically compatible NP *el Edgar*, who also left, and is therefore a candidate for being missed.

(156) Full given NP shift in conversational action

R: .. y=,
 los demás=s,
 .. ya no hay muchos,
 pues,
 sólo el Dario y yo quedamos de los .. que vivíamos en la
 célula.
 E: .. de ley.
 [y qué es del Edgar]?
 R: [porque el Edgar ya] se fue=,
 el Suco se fue=,
 .. todos se [fueron].
 E: [oh=],
 yo le extraño al **Su=co**.
 .. ya quisiera que venga.

R: .. and,
 everyone else,
 .. there aren't many any more,
 well,
 just me and Darío remain of those .. that lived on the block.

E: .. for sure.
 [and what about Edgar]?

R: [because Edgar already] left,
 Suco left,
 .. everyone [left].

E: [oh=],
 I miss **Suco**.
 .. I wish he could come back.

(Viejos Amigos:154-166)

Example (157) shows a different pattern than examples (152) through (156), in which the NP was repeated in precisely the same form as the previous mention. Rather than a repetition of a full NP, this example shows the introduction of a referent with a full NP (*tu ex pelada* ‘your ex-girlfriend’), followed by five unexpressed subject mentions and two mentions as an object, then a return to a full, but different, NP, *la man* (‘the chick’). The reason for the switch is not entirely clear, but it likely involves subjectivity/expression of stance. On the one hand, it could be argued that the speaker expressed the subject because the verb form *hubiese llamado* (‘would have called’) is ambiguous for person, so with no expressed subject it could be either first person singular or third person singular. However, given the context, with a first person singular object (i.e., *me hubiese llamado* ‘would have called **me**’), the case for ambiguity is weak. It is doubtful that the speaker would feel it necessary to further resolve the potential for ambiguity by expressing the subject, *la man*. In addition, if the case were to disambiguate the subject, the speaker easily could have selected another lesser form of coding, i.e., the personal pronoun *ella* (‘she’). A more plausible explanation arises when we consider the fact that the speaker uses a different full NP, *la man* (‘the chick’), than the first mention (*tu ex pelada* ‘your ex-girlfriend’). It appears that by using a full NP Adrián wants to

create some distance between himself and his ex-girlfriend and indicate his feelings about her to Rodrigo and therefore refers to her with a less personal term than ex-girlfriend. Lexical forms carry more information than lesser forms of coding, and pronoun or unexpressed mention would not be able to convey the stance that the full NP does in this example.

(157) Full NP after unexpressed mention

- R: ... (3.0) que pasó con tu= --
con= --
tu ex pelada que le dijiste que venga,
ya no ø viene?
- A: .. no,
ø no vino ya.
mejor.
... [le] --
- R: [pero],
le llamaste a decir que no ø venga?
- A: no.
- R: .. ø no vino?
- A: .. no.
.. es que ya salí de mañana.
... porque de ley **la man** me hubiese llamado de mañana acá a la
casa y ø me hubiese dicho ya voy.
... entonces ø ya no me encontró.
- R: ... (3.0) what happened with your --
with --
your ex-girlfriend that you told to come over,
(she's) not coming?
- A: .. no,
(she) didn't come.
better.
... [she] --
- R: [but],
did you call her to tell her not to come?
- A: no.
- R: .. (she) didn't come?
- A: .. no.
.. it's that I left this morning.
... because for sure **the chick** would have called this morning
and (she) would have told me she's coming.
... so (she) didn't find me.

(Fumar:527-541)

The NP *la silla* ('the chair') in (158) presents an interesting case. At this point in the conversation, the wheelchair has already been the topic of conversation for quite

some time. In this excerpt, it first occurs expressed as an object pronoun (*le* in *le utilicé* ‘I used it’), then as a full NP as an object of a preposition (*con la silla de ruedas* ‘with the wheelchair’), immediately followed by an unexpressed mention as a subject (\emptyset *me ha sido muy útil* ‘(it) has been very useful to me’). After a change in syntactic role, we may have predicted that *silla de ruedas* would be expressed in \emptyset *me ha sido muy útil* ‘(it) has been very useful to me’). Even more curious, the very next occurrence of *silla*, in the following clause, is expressed, although here there is continuity of syntactic role between the two mentions. It should also be noted that the intonation of the IU where *silla de ruedas* appears as an object of a preposition and an unexpressed subject is interesting—a more typical pattern would be to have *con la silla de ruedas* appear in its own IU, as an increment to the previous IU, followed by the clause \emptyset *me ha sido muy útil* on the following line, as the beginning of a new action, a characterization of the wheelchair. There is, however, a plausible interactional explanation. Just prior to the IU in question (*con la silla de ruedas* \emptyset *me ha sido muy útil*), Beatriz’s turn was interrupted by her brother. Beatriz then reaches a possible CRTP in the IU following the overlap (*también* \emptyset *asistí* ‘(I) also attended’), yet does not want to give up her turn at talk, so she speeds up her speech in the next IU (a technique known as ‘rush through’, see Schegloff (1982, 1987)), producing what would normally be two IUs under one intonation contour. Perhaps she realized that the resulting IU sounds awkward, with the same referent (wheelchair) occurring in two syntactic roles and later attempts to repair this by expressing *silla* as a full NP in the IU *la silla es muy agradable* ‘the chair is very nice’, then completes the list with another unexpressed mention. The resulting IUs: *también* \emptyset *asistí. con la silla de ruedas* \emptyset *me ha sido muy útil*, ‘(I) also attended. with the

wheelchair (it) has been very useful,' present a combination of two clauses that is similar to what has been recognized in the literature as syntactic amalgams (Lambrecht 1988). It should also be noted that the syntax of this three-part list deviates from the norm described above, where NPs in a list tend to exhibit the same expression. In this list, however, perhaps due to the previous interactional problems, we see the first part realized with an unexpressed subject (*ø me ha sido muy útil*), an expressed subject in the second (*la silla es muy agradable*) part, and a return to unexpressed in the third part (*ø es muy buena* '(it) is very good').

(158) Full given NP after change in syntactic role

- B: sí,
 sí.
 ø le utilicé para el velorio de --
 (H) de= mi mami,
 el tra- --
 el traslado de mi mami.
 .. (H) también fui a la misa de --
- G: ... [cla=ro,
 a la misa de mamá].
- B: [a la misa de un año de fallecimiento] de mis padres,
 también ø asistí.
 con la silla de ruedas ø me ha sido muy útil,
 ... (H) mhm.
la silla es muy agradable,
 ø es muy buena.
- B: yes,
 yes.
 (I) used it for the wake of -
 (H) of my mom,
 the fu- --
 the funeral of my mom.
 .. (H) I also went to the mass for -
- G: ... [of course,
 to mom's mass].
- B: [the mass for the one year anniversary of the death] of my
 parents,
 (I) also attended.
 with the wheelchair (it) has been very useful,
 ... (H) mhm.
the chair is very nice,
 (it) is very good.

(Hermanos:172-186)

To summarize, in this section we have examined seven NPs that occur as full NPs in contexts that theories of information flow would predict lesser coding. These examples exhibit patterns that we see repeatedly in the discourse, and it is only by examining each token in its natural context that these patterns are unveiled. It is evident from these examples that factors beyond activation status and distance from previous mention contribute to the choice of encoding an NP. Speakers choose to code given NPs as full NPs to accomplish different interactional goals—indicating to their interlocutor that they are constructing a list, repetition, managing turn-taking problems, marking an NP to indicate attitude toward or psychological distance from that referent, and to clarify a referent. Although it is not possible to quantify the scope of each of these interactional factors that contribute to NP realization, as each instance must be examined in the context in which it occurs, it is important and worthwhile to note that overall, interactional factors contribute to the realization of a significant number of NPs in conversation.

8.7 Summary

It is clear from this chapter that the functions of NPs in conversation go beyond issues related to information flow. NPs are involved in a broad array of social actions—managing turn taking, expressing stance and agreement, providing emphasis and delivering speech acts among others. When examined in the social and interactional context of conversation, patterns of use emerge. NPs in increments, which are not integrated into a main clause, tend to fall into syntactic roles least associated with the clause core—‘Other’ and prepositional phrases. There is also an observed tendency for NPs in the ‘Other’ syntactic role, especially non-referential NPs, to perform highly interactional functions. NPs in lists can show patterns of realization that are contrary to

the realization predicted by information flow, which signal to the speaker that a list is being constructed. The realization of given NPs is largely dictated by information flow, but deviations from this pattern, or ‘marked’ forms, are often interactionally motivated.

The findings presented in this chapter show that interactional factors clearly do contribute to the grammar (form and distribution) of NPs. In some contexts, there is no competition between interactional and information flow pressures, and the grammar of an NP is compatible with both. In other circumstances, the information flow and interactional pressures appear to compete, and speakers choose to use forms that reflect the interactional function of the NP, rather than the statistically unmarked patterns dictated by information flow.

9 Conclusions

In seeking out the answers to the three research questions proposed in Chapter 2, a number of interesting findings were reported in addition to the findings that address the original research questions. This comprehensive account of 3rd person NPs in Spanish convincingly shows that grammatical structure arises from discourse and is dependent on the immediate cognitive and interactional demands. In this section I will first summarize the major findings of this investigation and secondly will touch on areas of future research.

9.1 Summary of major findings

9.1.1 Grammar and competing motivations

The suggestion that speakers may at times struggle with competing motivations was first referred to here in Chapter 2. Throughout this investigation we saw that a variety of factors can influence the form of a NP. There were cognitive concerns related to information status, referent tracking, and high information flow pressures; interactional concerns such as managing turn-taking and expressing stance; and there was even evidence of discourse structural concerns, such as opening and closing topics.

It is only natural that over the course of a conversation there are times when two or more of these factors come into play together. If the factors do not coincide in their linguistic coding of the NP, then a competition results between the different ways of coding a particular NP. Speakers then must then choose which factor exhibits stronger influence in the context and code the NP accordingly.

Information status is always a factor in the coding of referential NPs, and identifiability or shared knowledge also proves to be important, as evidenced in the differences in definite and indefinite markers on new referents between the two genres under consideration. Theories of anaphoric coding that link referent realization with activation status (e.g., Givón 1981) do account for a large portion of the data—90% of new referents are coded as full NPs in the conversations and 73% of given referents are coded in lesser forms (pronoun and unexpressed mention). From this we can gather that information status is a very influential factor in the coding of the majority of referential NPs, and therefore results in the least marked pattern of coding. There is, however, a significant number of referential NPs that are coded in ways not predicted by theories of information flow, and must be accounted for under different frameworks. As Chapter 8 showed, many times it is evident that local interactional concerns do override the effects of information flow on the coding of a NP. It was also shown that there are times in which the coding of the NP reflects both the interactional and information flow context of the discourse, so while there are two or more motivations for coding the NP in a particular way, they are not in competition.

What we, as researchers, should interpret from this is the knowledge that speakers are incredibly adept at dealing with a number of different influential forces on grammar at any one time. They rely on their past conversational experience in order to choose the grammar that best reflects what they are trying to say. Grammar is an incredibly rich and flexible resource that speakers can adapt to their immediate cognitive and interactional needs.

9.1.2 *Referentiality*

One of the methodological aims of this study was to operationalize coding for referentiality. In doing so it was found that referentiality is best evidenced in the co-occurrence of several linguistic features and is not exclusively correlated with just one trait. It was shown that a number of characteristics that had previously been suggested to be indicative of non-referentiality are highly correlated with non-referentiality (bare NP, full NPs, and oblique syntactic role), but that none of these can be considered characteristic of all non-referential NPs. As Chapter 7 showed, one of the reasons that non-referential NPs cannot be accounted for by a single criterion is because they are not a single uniform class of NPs, rather there are many different types of non-referential NPs. Let's consider the case of morphological marking. Recall from Chapter 5 that 92% of all bare NPs are non-referential, but only 19% (narratives) and 43% (conversations) of non-referential NPs are bare. So while it is true that if a NP is bare it is very likely to be non-referential, it is not true that if a NP is non-referential, it is very likely to be bare. Non-referential NPs were shown to be a distinct category from both given and new referential NPs in many ways—realization of non-referential NPs is in no way constrained by previous mention, morphological marking (definite, indefinite or bare) patterns are different from those seen for referential NPs, and distribution by syntactic role is very different.

The findings of this study attest to the importance of distinguishing between referential and non-referential NPs in the quantitative study of language. In this study we saw how referentiality comes into play in the realization of PAS, as well as how different discourse functions are met by referential and non-referential NPs. Referentiality has been linked to transitivity (cf. Hopper & Thompson 1980)—clauses with referential NPs

are higher in transitivity than those with non-referential NPs. This is important when considering PAS, as these data revealed. In conversation especially, the constraint of ‘one lexical core argument per clause’ could be modified to ‘one lexical *referential* core argument per clause’, as seen in the relative frequency of clauses with lexical A and O, where the O proved to be non-referential. This also comes into play when measuring the overall transitivity of a genre—the lower transitivity of conversations is evidenced in the lower occurrence of referential NPs, especially A and O, and the high frequency of V-O Compounds, which Hopper and Thompson classify as intransitive predicates.

We also saw how speakers use referential and non-referential NPs for different discourse functions. Consider the analysis of the syntactic role ‘Other’. It was shown that in this role, referential NPs are more related to information flow needs (e.g., referent negotiation) and non-referential NPs are more likely to perform interactional functions (increments, (dis)agreement, stance-taking, attitude, etc.). .

9.1.3 Genre

This investigation has much to contribute to our understanding of genre differences. The Pear Film narratives were shown to be higher in information flow pressures and higher in transitivity in the conversations by several measures—a higher concentration of referential NPs, proportionately more NPs in the A and O roles, and the greater rate of persistence of NPs.

The higher information flow pressures of such narratives have grammatical consequences, including a higher rate of NPs modified by relative clauses, more objects of prepositions, and a stricter adherence to the constraints of PAS. The narratives show a more ergative structure, as reflected in the alignment of the S and O roles in introducing new information, whereas S and A are more closely aligned in the conversations.

Different discourse needs between the genres also results in different rates of use of grammatical markers. For example, this is very clear in the distribution of articles. More unidentifiable referents in the narratives leads to more indefinite first mentions. More identifiable referents in the conversation, due to the higher degree of shared knowledge between the participants, are reflected in the higher rates of definite first mention. One example of this is seen in first names, which are common in the conversations and do not appear in the Pear Story narrative, leading to a significant number of NPs being introduced with a definite article in the conversations (as first names are nearly always marked with the definite article when referential). More proper nouns (excluding first names) in conversations lead to more bare referential nouns than in the narratives.⁸⁶

The differential distribution of referential NPs and core arguments between the genres can partially be understood by the wider range of functions performed by non-referential NPs in the conversations, including vocatives, interrogative pronouns, non-specific 3rd person plural subjects, indicating closed topics and regulating turn-taking. There were also differences in the presentative lexicons of the two genres, with *haber* showing a presentative function in both genres, but copulas and perception verbs also showed a strong presentative function in the narratives, and the verb *tener* was strongly correlated with a presentative function in the conversations.

It was also seen that the syntactic role ‘Other’ is much more productive in conversations than in the narratives. It plays a significant role in the introduction of new referents in the conversations in addition to fulfilling a wide range of interactional functions.

⁸⁶ Future work might want to consider treating proper names separate from lexical mentions.

I would like to point out the important methodological and theoretical implications that arise from these findings. These data show that ‘grammar’ across genres can look very different. For instance, if one were to compare the rates of NPs that occur with and without determiners in these data, one could arrive at the conclusion that articles are not used as often in conversations as they are in narratives. Close examination of the *functions* of determiners explains why we observe the different rates: specifically the higher rate of non-referential NPs (which are much more likely to be bare than referential NPs) produces a higher rate of bare NPs in the conversational data. The researcher must be very careful not to confound rates of occurrence of grammatical forms with their linguistic conditioning; and it should be noted that no differences in the latter were observed in these data. Additionally, the genres themselves must be carefully defined. Categories such as ‘narrative’ and ‘conversation’ are helpful tools, but even within these categories, there are subtypes. The narratives used in this research are a particular kind of narrative—one that was designed by linguists to represent high information flow pressures—and the results presented within this monologue may not be applicable to other types of narratives, such as narratives of personal experiences. Similarly, conversation is a broad label and the type of conversations included in this study (informal conversations between friends and family members) may not reflect the same discourse needs present in other types of conversations. As more and more corpus-based research is conducted, researchers must be aware of the ways in which genre differences are grammatically manifested and must take care when comparing results of linguistic investigations conducted with different corpora and different genres.

9.2 Avenues of future research

This study has also drawn attention to areas of investigation that call for further research. The findings in Chapter 6 are suggestive of a possible priming effect in 3rd person subject expression, yet a much more comprehensive study which focuses solely on subjects would be necessary to either prove or disprove the existence of priming among third person subject realization.

This work has also paved the way for more studies of interactional functions of NPs. A larger sample of NPs, including those which do not display a more marked grammar could be examined in order come to a greater understanding of the factors that influence NP form and distribution. The study of NPs in the narrative genre could also be broadened to include the study of how factors beyond those related to information flow affect the form and distribution of NPs in narratives. Phenomena such as lists are certainly not exclusive to conversations, and could be examined across genres. In addition, more attention to the role that NPs play in discourse structure could also prove to be fruitful in both the conversations and narratives.

List of Appendices

Appendix 1: List of transcription symbols

Appendix 1

Transcription symbols (Du Bois et al. 1993)

(H)	inhalation
(Hx)	exhalation
@	laughter (each '@' represents one 'pulse' of laughter)
--	truncated IU
-	truncated word
.	final intonation
,	continuing intonation
?	rising intonation
..	short pause
...	long pause
...()	measured pause
[]	speech overlap
=	lengthened syllable
<X X>	unclear speech

References

- Ariel, Mira. 1990. Accessing noun-phrase antecedents. New York, NY: Routledge, Chapman & Hall.
- Ashby, William J. 1995. French presentational structures. In John Amastae, Grant Goodall, M. Montalbetti and M. Phinney (eds.). *Contemporary research in Romance linguistics*, 91-104. Amsterdam: Benjamins.
- Ashby, William & Paola Bentivoglio. 1993. Preferred argument structure in spoken French and Spanish. *Language Variation and Change* 5:61-76.
- 1997. Strategies for introducing new referents into discourse: A comparative analysis of French and Spanish presentational structures. In Robert M. Hammond and Marguerita G. MacDonald (eds.). *Linguistic studies in honor of Bohdan Saciuk*. West
- Bentivoglio, Paola. 1993. Full NPs in spoken Spanish: A discourse profile. In William J. Ashby, Marianne Mithun, Giorgio Perissinotto & Eduardo Raposo (eds.), *Linguistic perspectives on the Romance languages*, 212-224. Amsterdam / Philadelphia: John Benjamins.
- Biber, Douglas. 1988. Variation across Speech and Writing. Cambridge: Cambridge University Press.
- 1995. *Dimensions of register variation: A cross-linguistic comparison*. Cambridge: Cambridge University Press.

- Biber, Douglas, Mark Davies, James K. Jones and Nicole Tracy-Ventura. 2006. Spoken and written register variation in Spanish: A multi-dimensional analysis. *Corpora* 1(1): 1-37.
- Blackwell, Sarah E. 2003. *Implicatures in Discourse: The Case of Spanish NP Anaphora*. Amsterdam/Philadelphia: John Benjamins.
- Bosque, Ignacio. 1989. *Las categorías gramaticales*. Madrid: Síntesis.
- Burgos, Felix Manuel. 2007. Co-construcción en español: ¿cooperación por medio de una interrupción? *Forma y Funcion* 20: 13-39.
- Butt, John and Carmen Benjamin. 2004. *A new reference grammar of modern Spanish*. 4th edition. Chicago: McGraw Hill.
- Bybee, Joan. 2007. *Frequency of use and the organization of language*. Oxford: Oxford University Press.
- 2010. *Language, usage and cognition*. Cambridge: Cambridge University.
- Bybee, Joan and Paul Hopper (eds.) 2001. *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.
- Cameron, Richard. 1994. Switch reference, verb class and priming in a variable syntax. *Papers from the Regional Meeting of the Chicago Linguistic Society: Parasession on variation in linguistic theory* 30(2): 27-45.
- Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects and topics. In Charles Li (ed.) *Symposium on subject and topic*. New York: Academic Press. 25-55.
- 1980. (ed.) *The pear stories: Cognitive, cultural and linguistic aspects of narrative production*. Norwood, NJ: Ablex.
- 1982. Integration and involvement in speaking, writing, and oral literature. In Deborah Tannen (ed), *Spoken and written language: Exploring orality and literacy* (Advances in discourse processes), 35-53. Norwood, NJ: Ablex Publishing Corporation.
- 1985. Linguistic differences produced by differences between speaking and writing. In David R. Olson, Nancy Torrance and Angela Hildyard (eds), *Literacy, language and learning: The nature and consequences of reading and writing*, 105-123. Cambridge: Cambridge University Press.
- 1987. Cognitive constraints on information flow. In Russell Tomlin (ed.), *Coherence and grounding in discourse*. 21-51. Amsterdam/Philadelphia: John Benjamins.
- 1994. *Discourse, consciousness and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Clark, Herbert & Schaefer, Edward. 1989. Contributing to Discourse. *Cognitive Science* 13.259-294.
- Clancy, Patricia. 1980. Referential choice in English and Japanese discourse. In Wallace Chafe (ed.), *The pear stories: Cognitive, cultural and linguistic aspects of narrative production*, 127-202. Norwood, NJ: Ablex.
- Conrad, Susan & Biber, Douglas. 2001. *Variation in English: Multi-Dimensional studies*. London: Longman.
- Couper-Kuhlen, Elizabeth & Ono, Tsuyoshi. 2007. 'Incrementing' in conversation: A comparison of practices in English, German and Japanese. *Pragmatics* 17(4):513-552.

- Croft, William. 1991. Syntactic categories and grammatical relations: The cognitive organization of information. Chicago: University of Chicago Press.
- . 1995. Intonation units and grammatical structure. *Linguistics* (33):839-882.
- . 2007. Intonation units and grammatical structure in Wardaman and English. *Australian Journal of Linguistics*. 27.1-39.
- De Mello, George. 1992. El artículo definido con nombre propio de persona en el español contemporáneo. *Studia Neophilologica* 64.221-34.
- Dixon, Robert. 1979. Ergativity. *Language* (55): 59-138.
- Du Bois, John W. 1980. Beyond definiteness: The trace of identity in discourse. In Wallace Chafe (ed.). *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. 203-274. Norwood, NJ: Ablex.
- . 1985. Competing motivations. In John Haiman (ed.). *Iconicity in syntax*. 343-366. Amsterdam: Benjamins.
- . 1987. The discourse basis of ergativity. *Language* 63:4.805-855.
- . 2003. Discourse and grammar. In Michael Tomasello (ed.), *The new psychology of language: Cognitive and functional approaches to language structure*, vol. 2, 47-88. Mahwah, NJ: Lawrence Erlbaum Associates.
- Du Bois, John, Lorraine Kumpf and William Ashby. 2003. (eds.) *Preferred Argument Structure: Grammar as architecture for function*. Amsterdam/Philadelphia: John Benjamins.
- Du Bois, John, Stephan Scheutze-Coburn, Susanna Cumming and Danae Paolino. 1993. Outline of discourse transcription. In Jane Edwards and Martin Lampert (eds.), *Talking data: Transcription and coding in discourse*, 45-89. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Du Bois, John and Sandra Thompson. 1991. Dimensions of a theory of information flow. Manuscript.
- Dumont, Jenny. 2006. Full NPs as Subjects. In Nuria Sagarra and Almeida Jacqueline Toribio (eds.). *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, 286-296. Somerville, MA: Cascadilla Proceedings Project.
- Duranti, Alessandro and Elinor Ochs. 1979. Left-dislocation in Italian conversation. In T. Givón (ed.) *Syntax and semantics*, vol. 12, 377-416. New York: Academic Press.
- Durie, Mark. 2003. New light on information pressure: Information conduits, “escape valves”, and role alignment stretching. In John W. Du Bois, Lorraine E. Kumpf, and William J. Ashby (eds.), *Preferred Argument Structure*, 159-196. Amsterdam: Benjamins.
- Dutra, Rosália. 1987. The hybrid S-category in Brazilian Portuguese: Some implications for word order. *Studies in Language* 11(1): 163-180.
- Ewing, Michael. 2005. *Grammar and inference in conversation: Identifying clause structure in spoken Javanese*. Amsterdam: John Benjamins.
- Field, Margaret. 2007. Increments in Navajo conversation. *Pragmatics* 17(4):637-646.
- Flores Ferrán, Nydia. 2002. *Subject personal pronouns in Spanish narratives of Puerto Ricans in New York City: A sociolinguistic perspective*. Munich: Lincom Europa.
- Ford, Cecilia E., Barbara A. Fox and Sandra A. Thompson. 2002. Constituency and the grammar of turn increments. In Cecilia E. Ford, Barbara A. Fox, and Sandra A. Thompson (eds.) *The language of turn and sequence*, 14-38.

- Ford, Cecilia, and Barbara Fox. 1996. Interactional motivations for reference formulation: *He had. This guy had, a beautiful, thirty-two O:lds*. In Barbara A. Fox (ed.) *Studies in Anaphora*. Amsterdam/Philadelphia: John Benjamins.
- Ford, Cecilia and Sandra A. Thompson. 1996. Interactional units in conversation: Syntactic, intonational, and pragmatic resources for turn management. In Elinor Ochs, Emanuel Schegloff, and Sandra A. Thompson (eds.) *Interaction and Grammar*, 134-184. Cambridge: Cambridge University Press.
- Fox, Barbara A. 1987. *Discourse structure and anaphora: Written and conversational English*. Cambridge: Cambridge University Press.
- 1995. The category 'S' in English conversation. In Werner Abraham, T. Givón & Sandra A. Thompson (eds.), *Discourse, grammar and typology: Papers in honor of John W. M. Verhaar*, 153-178. Amsterdam: John Benjamins.
- Fox, Barbara and Sandra Thompson. 1990. A discourse explanation of the grammar of relative clauses in English Conversation. *Language* 66(2): 297-316.
- Garachana Camarero, Mar. Forthcoming. La creación y generalización del artículo indefinido. In Concepción Company Company (ed.), *Sintaxis histórica de la lengua española. Segunda parte: La frase nominal* (2 vols.). México: UNAM/FCE.
- Givón, Talmy. 1978. Definiteness and referentiality. In J. Greenberg (ed.), *Universals of human language: Volume 4, Syntax*, 291-330. Stanford: Stanford University Press.
- 1981. Logic vs. pragmatics, with natural language as the referee. *Journal of Pragmatics* 6:81-133.
- 1983. *Topic continuity in discourse*. Amsterdam: Benjamins.
- 1995. *Functionalism and grammar*. Amsterdam: Benjamins.
- Goodwin, Charles. 1981 *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press.
- 2007 Interactive Footing. In *Reporting Talk: Reported Speech in Interaction*, edited by Elizabeth Holt and Rebecca Clift. Cambridge: Cambridge University Press. pp. 16-46.
- Gundel, Jeanette, Nancy Hedberg and Ron Zacharski. 1993. *Cognitive Status and the Form of Referring Expressions in Discourse*. *Language* 69:274-307.
- Helasvuo, Marja-Liisa. 2001a. Emerging syntax for interaction: Noun phrases and clauses as a syntactic resource for interaction. In Elizabeth Couper-Kuhlen and Margret Selting (eds.) *Studies in interactional linguistics*. 25-50. Amsterdam/Philadelphia: John Benjamins.
- 2001b. Syntax in the making: The emergence of syntactic units in Finnish conversational discourse. *Studies in Discourse and Grammar*, John Benjamins, Amsterdam.
- Hochberg, Judith. 1986. Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language* 62(3):609-621.
- Hopper, Paul. 1987. Emergent grammar. *BLS* 13, 139-157.
- Hopper, Paul J., & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56.2: 251-299.
- 1984. The discourse basis for lexical categories in Universal Grammar. *Language* 60(4):703-752.
- Jefferson, Gail. 1990. List-construction as a task and a resource. In George Psathas, (ed.). *Interaction Competence*. Washington, D.C.: University Press of America: 63-92.

- Kärkkäinen, Elise. 1996. Preferred argument structure and subject role in American English conversational discourse. *Journal of Pragmatics* 25(5):675-701.
- Kim, Kyu-hyun. 2007. Sequential organization of post-predicate elements in Korean conversation: Pursuing uptake and modulating action. *Pragmatics* 17(4):573-603.
- Kumpf, Lorraine. 1992. Preferred Argument Structure and second language acquisition. *Studies in Language*: 16:2.
- Labov, William. To appear. Oral narratives of personal experience. In Patrick Hogan (ed.), *Cambridge Encyclopedia of the Language Sciences*. Cambridge: Cambridge University Press.
- Lambrecht, Knud. 1988. There was a farmer had a dog: Syntactic amalgams revisited. *Proceedings of the Berkeley Linguistics Society* 14:319-39.
- Lapidus, Naomi and Ricardo Otheguy. 2005. Contact induced change? Over nonspecific *ellos* in Spanish in New York. In Lotfi Sayahi and M. Westmoreland (eds.), *Selected proceedings of the second workshop on Spanish sociolinguistics*. Somerville, MA: Cascadilla Proceedings Project. 67-75.
- Lerner, Gene. 1987. *Collaborative turn sequences: Sentence construction and social action*. Unpublished doctoral dissertation. University of California, Irvine:CA.
- 1991. On the syntax of sentences-in-progress. *Language in Society* 20:441-458.
- Levinson, Stephan. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Matsumoto, Kazuko. 2003. *Intonation Units in Japanese conversation: Syntactic, informational, and functional structures*. Amsterdam: Benjamins.
- Miller, George A. 1956. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review* 63:81-97.
- O'Dowd, Elizabeth. 1990. Discourse pressure, genre, and grammatical alignment - after Du Bois. *Studies in Language* 14, 365-403.
- Ono, Tsuyoshi & Couper-Kuhlen, Elizabeth. 2007. Increments in cross-linguistic perspective: Introductory remarks. *Pragmatics* 17(4):513-552.
- Ono, Tsuyoshi & Thompson, Sandra. 1994. Unattached NPs in English Conversation. In *Proceedings of the 20th annual meeting of the Berkeley Linguistics Society*, 402-419. Berkeley: Berkeley Linguistics Society.
- Ortiz Ciscomani, Rosa María. Forthcoming. La creación y generalización del artículo definido. In Concepción Company Company (ed.), *Sintaxis histórica de la lengua española. Segunda parte: La frase nominal* (2 vols.). México: UNAM/FCE.
- Paredes Silva, Vera Lucia. 1993. Subject omission and functional compensation: Evidence from written Brazilian Portuguese. *Language Variation and Change* 5(1): 35-49.
- Prince, Ellen. 1981. Toward a typology of given-new information. In Peter Cole (ed.) *Radical pragmatics*. 223-55. New York: Academic Press.
- 1992. The ZPG letter: Subjects, definiteness and information-status. In William Mann and Sandra Thompson (eds.), *Discourse description: Diverse linguistic analyses of a fund-raising text*. 295-325. Amsterdam: John Benjamins.
- Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4):696-735.
- Sánchez-Ayala, Ivo. 2003. Constructions as Resources for Interaction: Lists in English and Spanish Conversation. *Discourse Studies* 5: 323-349.
- 2005. *The Spanish presentative verbal lexicon*. Unpublished manuscript.

- Schegloff, Emanuel. 1982. Discourse as an interactional achievement: Some uses of “uh huh” and other things that come between sentences. In Deborah Tannen (ed.), *Analyzing Discourse: Text and Talk*, 71-93. Georgetown: Georgetown University Press.
- 1987. Recycled turn beginnings: A precise repair mechanism in conversation’s turn-taking organisation. In Graham Button and John R. Lee (eds.) *Talk and Social Organization*, 70-85. Clevedon, England: Multilingual Matters.
- 1996. Turn organization: one intersection of grammar and interaction. In Elinor Ochs, Emanuel Schegloff, and Sandra A. Thompson (eds.) *Interaction and Grammar*, 52-133. Cambridge: Cambridge University Press.
- Scheibman, Joanne. 2001. Local patterns of subjectivity in person and verb type in American English conversation In J. Bybee and P. Hopper *Frequency and the emergence of linguistic structure*, 61-89. Amsterdam: John Benjamins.
- Silva-Corvalán, Carmen. 1994. *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon Press.
- Tannen, Deborah. 1987. Repetition in conversation: Toward a poetics of talk. *Language* 63:574-605.
- Tao, Hongyin. 1992. NP intonation units and referent identification. In Proceedings of the 18th annual meeting of the Berkeley Linguistics Society, 237-247. Berkeley: Berkeley Linguistics Society.
- Thompson, Sandra A. 1984. Subordination in formal and informal discourse. *Georgetown University Round Table on Languages and Linguistics*, 85-94.
- 1997. Discourse motivations for the core-oblique distinction as a language universal. In Akio Kamio (ed.), *Directions in functional linguistics*, 59-82. Amsterdam.
- 2002. 'Object complements' and conversation: Towards a realistic account. *Studies in Language* 26(1) 125-164.
- Thompson, Sandra A. and Paul Hopper. 2001. Transitivity, clause structure and argument structure: Evidence from conversation. In J. Bybee and P. Hopper (eds.) *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamins, 27-60.
- Tomlin, Russell. 1987. Linguistic reflections of cognitive events. In Russell Tomlin (ed.) *Coherence and Grounding in Discourse*, 455-480. Benjamins: Amsterdam.
- Toribio, Almeida Jacqueline. 1993. Parametric variation in the licensing of nominals. PhD Dissertation, Cornell University, Ithaca, New York.
- 1996. Dialectal variation in the licensing of null referential and expletive subjects. In Claudia Parodi, Carlos Quicoli, Mario Saltarelli and María Luisa Zubizarreta (eds.) *Aspects of Romance linguistics: Selected papers from the Linguistic Symposium on Romance Languages XXIV March 10-13, 1994*, 409-432. Washington D.C.: Georgetown University Press.
- 2000. Setting Parametric Limits on Dialectal Variation in Spanish. *Lingua* 110(5): 315-341.
- Torres Cacoullos, Rena and Jessi Aaron. 2003. Bare English-origin nouns in Spanish: Rates, constraints, and discourse functions. *Language Variation and Change*, (15)289-238.
- Travis, Catherine. 2005. The yo-yo effect: Priming in subject expression in Colombian Spanish. In Randall Gess and Edward J. Rubin (eds.), *Theoretical and experimental*

- approaches to Romance linguistics: Selected papers from the 34th Linguistic Symposium on Romance Languages, 2004*, 329-349. Amsterdam / Philadelphia: John Benjamins.
- 2007. Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19(2): 101-135.
- Wilkes-Gibbs, Deanna. 1986. *Collaborative processes of language use in conversation*. Unpublished doctoral dissertation, Stanford University, CA.