



2006

## Full NPs as Subjects

Jenny Dumont  
*Gettysburg College*

Follow this and additional works at: <https://cupola.gettysburg.edu/spanfac>

 Part of the [Spanish and Portuguese Language and Literature Commons](#)

**Share feedback about the accessibility of this item.**

---

Dumont, Jenny. 2006. Full NPs as Subjects. In Selected Proceedings of the 9th Hispanic Linguistics Symposium, ed. Nuria Sagarra and Almeida Jacqueline Toribio, 286-296. Somerville, MA: Cascadilla Proceedings Project. [www.lingref.com](http://www.lingref.com), document #1386.

This is the publisher's version of the work. This publication appears in Gettysburg College's institutional repository by permission of the copyright owner for personal use, not for redistribution. Cupola permanent link: <https://cupola.gettysburg.edu/spanfac/8>

This open access conference proceeding is brought to you by The Cupola: Scholarship at Gettysburg College. It has been accepted for inclusion by an authorized administrator of The Cupola. For more information, please contact [cupola@gettysburg.edu](mailto:cupola@gettysburg.edu).

---

# Full NPs as Subjects

## **Abstract**

The purpose of this study is to investigate and identify factors that influence 3rd person (singular and plural) subject expression in the spoken Spanish of New Mexico, with particular emphasis on the role of full NPs as subjects. Subject expression has been the topic of many linguistic studies, especially with respect to personal pronouns (Bayley & Pease Alvarez 1997; Bentivoglio 1987; Cameron 1994, 1995; Cameron & Flores 2003; Flores-Ferrán 2002; Silva-Corvalán 1994; Travis 2005, To appear). What distinguishes this investigation from other studies of subject expression is that it focuses on the third person, which is a more heterogeneous category than first or second person in that it encompasses many distinct types of subjects, both semantically (animate and inanimate, general and specific, etc.), and formally (full NP, pronoun and unexpressed). Additionally, the present study is concerned with the role of full NPs and particularly how full NPs contrast with other forms of third person subjects, namely pronouns and unexpressed subjects. [*excerpt*]

## **Keywords**

Spanish, New Mexico, linguistics

## **Disciplines**

Spanish and Portuguese Language and Literature

## **Comments**

This paper was presented at the 9th Hispanic Linguistics Symposium in Somerville, Massachusetts in 2006.

# Selected Proceedings of the 9th Hispanic Linguistics Symposium

edited by Nuria Sagarra  
and Almeida Jacqueline Toribio

Cascadilla Proceedings Project Somerville, MA 2006

## Copyright information

Selected Proceedings of the 9th Hispanic Linguistics Symposium  
© 2006 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 1-57473-413-X library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.  
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

## Ordering information

Orders for the library binding edition are handled by Cascadilla Press.  
To place an order, go to [www.lingref.com](http://www.lingref.com) or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA  
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: [sales@cascadilla.com](mailto:sales@cascadilla.com)

## Web access and citation information

This entire proceedings can also be viewed on the web at [www.lingref.com](http://www.lingref.com). Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Dumont, Jenny. 2006. Full NPs as Subjects. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, ed. Nuria Sagarra and Almeida Jacqueline Toribio, 286-296. Somerville, MA: Cascadilla Proceedings Project.

or:

Dumont, Jenny. 2006. Full NPs as Subjects. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, ed. Nuria Sagarra and Almeida Jacqueline Toribio, 286-296. Somerville, MA: Cascadilla Proceedings Project. [www.lingref.com](http://www.lingref.com), document #1386.

# Full NPs as Subjects

Jenny Dumont  
University of New Mexico

## 1. Introduction

The purpose of this study is to investigate and identify factors that influence 3<sup>rd</sup> person (singular and plural) subject expression in the spoken Spanish of New Mexico, with particular emphasis on the role of full NPs as subjects. Subject expression has been the topic of many linguistic studies, especially with respect to personal pronouns (Bayley & Pease Alvarez 1997; Bentivoglio 1987; Cameron 1994, 1995; Cameron & Flores 2003; Flores-Ferrán 2002; Silva-Corvalán 1994; Travis 2005, To appear). What distinguishes this investigation from other studies of subject expression is that it focuses on the third person, which is a more heterogeneous category than first or second person in that it encompasses many distinct types of subjects, both semantically (animate and inanimate, general and specific, etc.), and formally (full NP, pronoun and unexpressed). Additionally, the present study is concerned with the role of full NPs and particularly how full NPs contrast with other forms of third person subjects, namely pronouns and unexpressed subjects.

Each potential form of expression is exemplified in (1)-(3). For each example the relevant subject is in bold, verbs are underlined, and unexpressed subjects are given in parentheses in the English translation.

(1) Full NP

*los pegó un carro<sup>1</sup> (10-8A)<sup>2</sup>*  
'**a car** hit us.'

(2) Pronoun

*ella sabe en inglés también. (20-5A3)*  
'**she** knows in English too.'

(3) Unexpressed

*.. Cosechaba trigo y molía trigo y, .. molían el trigo, y hacían harina pa' m- -- pa' .. tortillas. (4-5B2)*  
'**(One)** would harvest the wheat and **(one)** would grind the wheat and, **(they)** would grind the wheat, and **(they)** would make flour for m- -- for ..tortillas.'

The role of full NPs as subjects has previously been described as mainly for introducing new information into discourse (Silva-Corvalán 1994:143), yet the results presented in this study describe additional uses of full NPs as subjects. In particular, it was found that the same full NP can and does occur in two sequential clauses, and two underlying speaker motivations were identified. On the one hand, full NPs are found in question/answer sequences, indicating a repetition effect between speaker and interlocutor. On the other hand, it is found that speakers use full NPs when making contrasts. These findings indicate that theories of information flow alone are not sufficient to account for all the uses of full NPs as subjects, and they provide insight into both the pragmatic and the interactional effects on language patterns and the organization of grammar.

---

\* I am grateful to Catherine Travis for her valuable comments on this paper. I would also like to thank Rena Torres Cacoullos and three anonymous reviewers for their helpful suggestions.

<sup>1</sup> los = nos.

<sup>2</sup> Letters and numbers in parentheses denote the interview number, tape and side.

## 2. Data

The data on which this study is based were taken from the materials of the New Mexico- Colorado Spanish Survey (NMCROSS) a project directed by Garland D. Bills and Neddy A. Vigil at the University of New Mexico and funded in part by the National Endowment for the Humanities (cf. Bills and Vigil 1999). The interviews belong to a larger corpus of interviews conducted between 1992 and 1995 by graduate students at the University of New Mexico. There are several recurrent topics of conversation in the interviews, including family life, language use, and childhood events.

Transcripts of 5 speakers were used for this analysis, representing approximately 2 hours and 45 minutes of speech, and yielding 1028 clauses containing verbs in the 3<sup>rd</sup> person. Two of the interviewees were men and three were women. All speakers were 45 years of age or older at the time of the interviews, have a high level of proficiency in Spanish, and varying degrees of English proficiency. Although the varieties of Spanish spoken in the Southwest have been in contact with English for several centuries and there is evidence of lexical borrowing, the degree of influence from English on the syntax is unknown. Travis (To appear) reports no significant differences among the linguistic constraints of first person subject expression in Colombian Spanish and New Mexican Spanish. However, studies of subject expression in other languages such as Finnish, Hebrew (Levy and Vainnika 1999/2000; Vainnika and Levy 1999) indicate that first and second person pronouns have different constraints on subject expression from third person pronouns and full NPs. Therefore is not known if the results of this study are specific to the dialect under investigation. The findings presented in this study may be specific to the genre under investigation, one-on-one interviews.

Table 1 contains the distribution of 3<sup>rd</sup> person subjects for each form of expression. Unexpressed subjects account for the majority of the clauses (58%), followed by full NPs (22%), and pronouns (20%). As the focus of this paper is the role of the full noun phrase as it differs from the other forms of expression, pronominal and unexpressed subjects were combined for the statistical analysis discussed in Section 4.

Table 1: Distribution of 3<sup>rd</sup> person subjects

Full NP	Pronoun	Unexpressed	Total
22% (224)	20% (206)	58% (598)	100% (1028)

## 3. Methodology and coding

### 3.1 Sociolinguistic variationist method

This research was carried out using the sociolinguistic variationist method (Labov 1966/1982, 1971, 1984), which assumes that there are patterns to the inherent variability of spoken language, and that by means of careful quantitative analysis these patterns can be identified and described by the linguist. The data were coded for the following factors: distance from previous mention, realization of previous mention, information status, specificity, grammatical number, semantic class of verb, animacy, referentiality (cf. Du Bois 1980), syntactic role, and potential for ambiguity of morphological form. The following subsections describe the coding of the factors found to have a significant effect on the expression of subjects as full NPs.

### 3.2 Distance and previous realization

Givón (1983) was perhaps the first to state the importance of topic continuity in discourse, and to propose counting clauses as a means of operationally measuring referential distance. Subsequent studies of subject expression have pointed to the role of topic continuity, as well as discontinuity, as relevant to form of expression. Change of subject, or switch reference, has consistently been found to be a significant factor contributing to subject expression in Spanish (Bayley & Pease Alvarez 1997; Bentivoglio 1987; Cameron 1995; Flores-Ferrán, 2002; Silva-Corvalán 1994; Travis 2005, To appear). On the other hand, priming effects, or the perseveration of a form of subject expression whereby the

same form of expression is repeated across subsequent clauses, have also been observed (Cameron and Flores-Ferrán 2003; Flores-Ferrán 2002; Travis 2005).

The data for this study were coded for both the distance from the previous mention of coreferential subject and the realization of the previous mention. Although only third person subjects produced by the interviewees were extracted for coding, the speech of the interviewer was included in the counts of distance from previous mention and realization of previous mention. Including the speech of the interviewer allows for a better understanding of the interactional dimension of referent realization, as seen in the repetition effect in question/answer sequences.

Referents which had not previously been mentioned in the discourse, and therefore had no previous coreferential mention, were given a separate code. Likewise, referents which had been introduced into the discourse in other syntactic roles (object, oblique, etc.) but had not been subjects in the previous coreferential mention were given a separate code. Another code was designated for those referents for which previous mention was not applicable, as for general statements, and non-specific *ellos* and *uno*.

For the remaining referents, distance was initially measured up to 10 intervening clauses. Although the 1028 clauses forming the basis of the analysis are strictly from the interviewees, the speech of the interviewer was included in the counting of the intervening clauses.

The coding of distance from previous mention is illustrated in examples (4)- (8).

(4) First mention of a referent

*Ahí iba yo sola, porque el marido en veces estaba trabajando. y en ve- luego, me acompañaba un muchito<sup>3</sup>.* (10-8A)

‘I would go there alone, because sometimes my husband was working. and later, **a little boy** would accompany me.’

(5) First mention as subject

A: *La -- .. Cómo la conservaban digamos para el invierno?*

N: *Pues hacían subterráneos. .. Alguna gente que tenía subterráneo como ahora. Hacían pozos. Un pozo. Y luego le ponían techo. Y -- .. y lo ponía ahí. Y ahí duraba. Las cebollas, las papas. Las papas también las pueden tener ahí en el subterráneo. y ahí duraban por el invierno.*

A: *Eh.*

→ N: *Y los subterráneos eran más frescos. En el verano eran muy frescos. Y en el invierno eran calientes. Y ahí duraba.* (10-3A2)

‘A: The --How would they store it, let’s say, for winter?’

N: Well they made cellars. ..Some people had cellars like they do now. They made holes. A hole. And then they would put a roof on it. And-- and they would put it there. And it would keep there. The onions, the potatoes. The potatoes they can also have them there in the cellar. And there they would keep for the winter.

A: *Eh.*

N: And **the cellars** were cooler. In the summer they were very cool. And in the winter they were warm. And there it would keep.’

(6) No intervening clauses

S: *Te va a agarrar el perro.*

D: *Te va a agarrar el perro.*

S: *Tiene rabia el perro.* (102-3A2)

‘S: The dog is going to get you.’

D: ‘**The dog** is going to get you.’

S: ‘**The dog** has rabies.’

<sup>3</sup> muchito= muchachito.

## (7) One intervening clause

y ya iba a gradar. ... Y ... para mí era un-- una cosa tan feliz pa' conocerlo. .. porque era muy brillante el muchacho. (88-1A3)

'and (he) was already going to graduate. And for me (it) was such a happy thing to meet him... because **the boy** was very brilliant.'

## (8) Three intervening clauses

*Diferentes enfermedades. curaba ella, con diferentes remedios. pero yo no me recuerdo qué clase de remedios serían, pues, estaba yo -- todavía estaba mediana. ... Pero ésa curaba mucho.<sup>4</sup> (4-2A2)*

'Different sicknesses. She would cure with different remedies. but I don't remember what type of remedies they were. well, I was -- (I) was still young. ...But she (lit. 'that one') cured a lot.'

The distribution of the data is given in Table 2. Full NPs are clearly preferred for the first mention of a subject, with 80% of first mentions encoded as full NPs, and the remaining 20% are pronominal or unexpressed. Full NPs account for nearly one third of first mentions as subjects, at 32%, but the majority (68%) are expressed pronominally or unexpressed. Where there are no intervening clauses between coreferential mentions, subjects are expressed with a full NP in 16% of the cases, and are pronominal or unexpressed in the remaining 84%. The overall tendency toward expression with a full NP for 1-4 intervening clauses is somewhat less, and therefore these tokens were combined for the statistical analysis. Due to the low number of tokens occurring 5-9 clauses after the previous mention (n=13), these were excluded from the statistical analysis. Those occurring ten or more clauses away were not included in the analysis either, although there is a clear tendency for such referents to be reintroduced as a full NP (74%), which indicates that they pattern similarly to referents being introduced for the first time. For certain referents distance from last mention is not applicable, such as general assessments with no clear subject, impersonal *uno* and third person plural impersonal statements.

Table 2: Referential Distance

Distance	Full NP (n)	Pronominal unexpressed	or Total
First mention	80% (114)	20% (28)	142
First mention as subject	32% (30)	68% (64)	94
0	16% (38)	84% (207)	245
1	10% (7)	90% (63)	70
2	18% (5)	82% (23)	28
3	11% (2)	89% (16)	18
4	0% (0)	100% (10)	10
5	43% (3)	57% (4)	7
6	0% (0)	100% (3)	3
7	50% (1)	50% (1)	2
8	0% (0)	100% (1)	1
9	-	-	-
10+	74% (17)	26% (6)	23
Not applicable	2% (7)	98% (378)	385
<b>Totals</b>	22% (224)	78% (804)	100% (1028)

As mentioned earlier, the previous realization of a coreferential subject has been found to contribute to the form of expression of a given referent. Therefore, the data were coded for three possible forms of previous realization: full NP, pronoun, unexpressed, and a separate code was given

<sup>4</sup> Repaired/truncated clauses (indicated in the transcript with two dashes [--]) were not included in the clause count.

for those referents for which previous realization is not applicable, including the first mention of subjects, general statements. General statements are those which do not have a clear NP as a subject, rather they tend to be assessments and refer to an entire proposition or situation. This coding is exemplified in (9)- (12) below.

- (9) Previous realization as full NP

*Don Julio Restrepo curaba de brujería. Ése era de Santa Cruz.* (4-2A2)  
 ‘Don Julio Restrepo cured (people) of witchcraft. **He** was from Santa Cruz.’

- (10) Previous realization as pronoun

y él vino y la curó y sanó. (4-2A2)  
 ‘and **he came** and (**he**) cured her and (she) got better.’

- (11) Previous realization unexpressed

*Tengo una= -- .. bisnieta es. No es nieta. Es bisnieta.  
 .. Y esa muchachita entiende todo en español.* (20-5A3)  
 ‘I have a= -- .. (she) is a great-granddaughter. (**She**) is not a granddaughter. (**She**) is a great-granddaughter. .. And **that girl understands** everything in Spanish.’

- (12) General statement

*Porque si yo no le entendía a él lo que me estaba diciendo o -- o -- o -- O él a mí, .. pues era difícil. Hm, no podíamos tener ninguna conversación.* (20-5A3)  
 ‘Because if I didn’t understand him, what he was saying or -- or -- or -- or him me, .. well (**it**) was difficult. Hm, we couldn’t have any conversation.’

### 3.3 Information Status and referentiality

Several studies have highlighted the role of information flow in the distribution of full NPs (Bentivoglio 1993; Ocampo 1990, 1995). Cameron (1994) found the specificity of the referent (specific *tú* vs. non-specific *tú*) to be a significant factor for second person subject expression. In this study the specificity of the referent was coded in conjunction with information status. Preliminary analysis showed that all new referents, general and specific, tended to be encoded as full NPs, therefore the two categories were combined for the statistical analysis presented below.

- (13) New information, general referent

*después emb- comenzaron a entrar máquinas pa' hacer el ice cream.* (10-3A2)  
 ‘after st- **machines for making ice cream** started coming around.’

- (14) New information, specific referent

*en El Vado estaba un restaurante.* (10-8A)  
 ‘in El Vado there was a restaurant.’

- (15) Given information, general referent

C: *Y recuerdas algunos ... cuentos de brujería?*  
 V: .. *No. Otras cosas no -- Antes sí, que había brujas pero --  
 ... no eran -- ... O sea, sí, eran de aquí pero no -- podían  
 probarles que eran brujas.* (4-2A2)  
 ‘C: And do you remember any.. stories of witchcraft?  
 V: ..No. Other things no – Before yes, there were witches but –  
 ...they weren’t -- ...Well, yes, (**they**) were from here but not – they couldn’t prove that  
 (**they**) were witches.’

- (16) Given information, specific referent

N: *Y los hijos de esta muchacha, me tratan como grandma.*

*los de la Juliana, ésta que te digo yo.*

A: *Ajá.*

N: *ese me trujo también un -- un canasto ansina también de todo.* (10-8A)

‘N: And the kids of this girl, they treat me like a grandma.

Juliana’s, the one that I’m telling you about.

A: Uhuh.

N: **That one** also brought me a -- a basket also with everything.’

### 3.4 Number

All tokens were coded for grammatical number (singular or plural).

### 3.5 Semantic class of verb

Previous studies have found the semantic class of the verb to be an important factor in subject expression (Bentivoglio 1987; Enríquez 1984; Silva-Corvalán 1994; Travis 2005). Specifically, cognitive and communication verbs have been found to favor expressed first person subjects, perhaps due to the epistemic value or subjectivity associated with these verbs (Scheibman 2001). All verbs in this study were coded for the following semantic classes.

Verbs of communication (*decir*<sup>5</sup>) and verbs of cognition (*saber, pensar, creer*, etc.) were at first coded separately, and then later combined for the statistical analyses when they were shown to pattern similarly. Linguistically, it is not surprising that these two semantic classes would show similar patterns in the third person, as they are both reporting similar types of information about a third person, be it what the third person thinks, knows or says.

A third class was designated for verbs of movement (*ir, venir, salir, llegar*, etc.). A fourth class of verbs, copulas, were given a separate code, and all remaining verbs were coded as ‘other’. The last class was excluded from the statistical analysis.

## 4. Results

### 4.1 Analysis

Of the 1028 clauses analyzed, 224 clauses (22% of the data) contained full NPs as subjects. The data were analyzed with the program GoldVarb (Rand & Sankoff 1990), which indicates which factor groups have a significant effect on the dependent variable, in this case form of expression, as well as the magnitude of effect (strength of effect) of each factor group. A probability weight of 0.50 indicates that there is no effect of the factor on the variable, whereas a weight of greater than 0.50 indicates that the variable is favored, and less than 0.50 suggests that the variable is disfavored. The closer the probability weight to 1 or to 0, the stronger the effect. The magnitude of effect is measured by the range, which is the difference between the highest probability weight and the lowest probability weight within a factor group. The groups with the greatest ranges show a stronger effect, those with smaller ranges indicate a less strong effect.

Five factor groups were selected as significant: number, specificity/information status, distance from previous mention, realization of previous mention, and semantic class of verb. The results are displayed in Table 3.

<sup>5</sup> Although there are other verbs of communication, *decir* was the only one present in this data.

Table 3: Variable rule analysis; factor groups selected as significant to the expression of a subject with a full NP <sup>6</sup> ( $p < .01$ )

	weight	% (n)	% of data <sup>7</sup>
<b>Prev. Realization</b>			
NP	.79	25 % (115)	11%
unexpressed	.55	7% (351)	34%
pronominal	.13	<1% (121)	12%
<i>Range</i>	.66		
<b>Distance</b>			
1 <sup>st</sup> mention	.87	80% (142)	14%
1 <sup>st</sup> mention as subject	.38	31% (94)	9%
0 intervening clauses	.38	15% (245)	24%
1-4 intervening clauses	.31	11% (126)	12%
<i>Range</i>	.56		
<b>Information flow</b>			
New (general & specific)	.82	89% (88)	9%
Given/general	.67	33% (116)	11%
Given/specific	.35	14% (350)	34%
<i>Range</i>	.47		
<b>Verb Class</b>			
Motion	.78	34% (70)	7%
Copula	.47	28% (257)	25%
Cog./Comm.	.38	11% (111)	11%
<i>Range</i>	.40		
<b>Number</b>			
Singular	.61	28% (624)	61%
Plural	.33	11% (404)	39%
<i>Range</i>	.28		

Input= .139, Log likelihood= -311.808, Chi-square/cell= 1.25

#### 4.2 Previous realization and distance

The factor group with the largest effect is that of previous realization, with previous realization as full NP favoring expression as a full NP (probability weights of .79 for previous realization as full NPs, .55 for previous realization unexpressed, and .13 for previous realization pronominal). Distance from previous mention also had a strong effect on the form of expression, with 1<sup>st</sup> mention clearly favoring expression as a full NP (probability weight .87), and all other factors disfavoring expression as a full NP. Although it is not surprising that new referents tend to be expressed as full NPs, it is however interesting that full NPs appear to be slightly more likely to occur when there are no intervening clauses (probability weight .38) as opposed to when there are 1-4 intervening clauses (probability .31). In other words, although one might expect that increased intervening material between subsequent mentions of a subject would increase the probability of expression as a full NP, these data indicate that a full NP may be more likely to follow another full NP with no intervening clauses.

The relationship between distance and previous realization can be seen in Table 4. As seen in the table, 72% (21/29) of full NPs which were previously realized as full NPs immediately followed coreferential mentions of the same full NP. Less often were full NPs used when coreferential mentions were expressed 1 or more clauses prior. The table also shows that once a referent has been expressed pronominally it is unlikely to be expressed again as a full NP, with only 1 token of this kind,

<sup>6</sup> Factors included in the analysis but not selected as significant are: animacy, referentiality, syntactic role (subject of transitive or intransitive clause), and potential for ambiguity of morphological form.

<sup>7</sup> Some percentages do not total 100% due to the exclusion of factors within factor groups.

appearing more than 5 clauses away from the prior realization. When a subject has not been expressed in the previous mention, if there are no intervening clauses it is most likely to be expressed with a full NP, slightly less likely to be expressed with a full NP after 1-4 clauses, and least likely to be expressed with a full NP after 5 or more clauses.

Table 4: Full NPs according to distance and realization

	Previous Realization		
	Full NP	Pronoun	Unexpressed
<b>0 Inter. Clauses</b>	21	0	13
<b>1-4 Inter. Clauses</b>	4	0	8
<b>5-10 Inter. Clauses</b>	4	1	2
<b>Total</b>	29	1	23

The occurrence of the same NP in two sequential clauses indicates a repetition effect, and when the tokens in question were examined, two types of repetition were noted: question/answer sequences, in which the speaker repeats an NP produced previously by the interviewer, and repetition in contrast, in which the speaker repeats an NP in two consecutive clauses. Both types of repetition are exemplified below.

(17) Repetition in a question/answer sequence

- A: *Qué bien. Y este -- Su papá, dónde nació?*  
 R: .. *Mi papá también nació allá en Questa. En el lugar ése. Sí.*  
 A: *Y su mamá también?*  
 R: *Y mi mamá también.*  
 A: *Así que los dos eran de allí.*  
 R: *Ahá. Sí. Los dos.*  
 A: *Mhm.*  
 R: *Mhm.*  
 A: *Y el --y su esposo dónde nació?*  
 R: .. *Mi esposo nació en Ranchos de Taos.* (20-1A1)  
 ‘A: Great. And so -- **Your father**, where was (he) born?  
 R: ..**My father** also was born there in Questa. In that place. Yes.  
 A: And your mother too?  
 R: And my mother too.  
 A: So they were both from there.  
 R: Uhuh. Yes. Both.  
 A: Mhm.  
 R: Mhm.  
 A: And the-- and **your husband** where was (he) born?  
 R: **My husband** was born in Ranchos de Taos.’

(18) Repetition in contrast

- la familia respetaba a los padres.*  
*Hoy en día no respeta la familia.*  
 .. *a muchos de los padres.*  
*Los padres las madres se van a trabajar y,*  
 .. *y los hijitos los dejan solos.*  
 .. *y los hijitos se quedan sueltos,*  
 .. *y antes no=. antes casi no --*  
*no -- no trabajaba la mujere.*  
*La mujer cuidaba de la familia.* (4-5B2)  
 ‘**Families** respected parents.’

Today **families** don't have respect.  
 ..for many parents.  
 The fathers the mothers go to work and,  
 ..and leave the kids alone.  
 ..and the kids are left loose,  
 .. and before no=. before hardly ever--'  
 ever -- **women** didn't work.  
**Women** took care of the family.'

In example 17, the interviewee (R) is seen to respond to the interviewer's questions with a striking similarity in syntactic form, including the use of full NPs. Note that this repetition of NPs is evident throughout the exchange, even when the NP is not a syntactic subject (*su mamá/mi mamá; los dos/los dos*).

In example 18, the repetition of the NP has a more clear discourse function- that of contrast. In this example, the speaker is contrasting families of the past with families of the present, and the repetition of the NP can be seen as fulfilling a pragmatic function. Additional contrast is seen in the role of the woman: what a woman did do versus what she did not do.

Given the examples above, full NPs clearly serve two functions beyond the mere introduction of new referents into the discourse.

#### 4.3 Information flow

Clearly, the preferred form of expression for new referents is with a full NP, with a probability weight of .82. The results also indicate that there is difference between the two types of given referents: general and specific. There is a stronger tendency of using a full NP to express given/general referents (probability weight .67) than for given/specific referents (.35). This may be attributed to the nature of general referents, which often cannot be expressed pronominally and are used to make generalizations, as seen in example (19). The emergence of two subclasses of given referents, one which favors expression with a full NP and one which disfavors full NPs, also lends support to the claim that the role of the full NP as a subject is not limited to new referents, and furthermore shows that within given referents expressed by full NPs there are patterns of usage.

##### (19) Given/general referent

C: .. *Y en qué los .. sepultaban?*  
 V: *En -- en cajones de -- de madera hechos.*  
 .. *Los hacían.*  
 [5 intervening clauses]  
 ... *No es como ahora. .. Que ahora con el --*  
*ahora, **el cajón** es más lujoso. (4-2A2)*  
 'C: ..And in what did they .. bury them?  
 V: *In -- in coffins of -- made of wood.*  
 .. *They made them.*  
 [5 intervening clauses]  
 .. *It's not like now. .. Now with the --*  
*now, **the coffin** is more luxurious.'*

#### 4.4 Semantic class

Type of verb also proved to be a significant factor, with motion verbs favoring the use of a full NP (.78 factor weight), copulas slightly disfavoring the use of a full NP (.47) and cognitive/communications also disfavoring full NPs (.38). One explanation for why motion verbs show a strong favoring effect for full NPs may be due to their argument structure. Preferred Argument Structure (Du Bois 1987; Ashby & Bentivoglio 1993) is the tendency for new information to be presented in syntactic roles S and O, and least often in syntactic role A, hence the greater likelihood of arguments in

the S and O syntactic slots being expressed as lexical NPs. Motion verbs, as intransitive verbs, are therefore more likely to occur with lexical subjects than are transitive verbs, because subjects of transitive verbs (syntactic role A) tend not to be coded with lexical NPs. It has also been noted that motion verbs are frequently used for introducing new information (cf. Sánchez-Ayala 2005), hence the increased likelihood that the subject of a motion verb will be encoded as a full NP.

#### 4.5 Grammatical number

Singular subjects are shown to be more often expressed as full NPs than plural subjects (.61 vs. .33). This finding is similar to what has been reported in studies of pronominal subject expression, in which singular subjects tend to be expressed (pronominally) more often than plural subjects (Bentivoglio 1987; Cameron 1994; Bayley 1997; Flores-Ferrán 2002). Additionally, the high percentage of impersonal statements in this corpus (236 tokens of non-specific 3<sup>rd</sup> person plural subjects) may be a contributing factor to the lower expression of plural subjects.

### 5. Discussion

This study of full NPs as subjects in Spanish has demonstrated that the role of the full NP goes beyond merely introducing new referents into discourse, and additional patterns of full NP usage emerged from the analysis. It was shown that there is a greater tendency for general/given referents to be expressed with a full NP than their specific/given counterparts, that motion verbs favor full NPs, and that plural rather than singular subjects are more likely to be expressed with a full NP. Additionally, two speaker based motivations for the repetition of a full NP were identified: an interactional motive (repetition between speakers), as well as marking contrast.

Although traditional theories of information flow can account for some patterns of NP distribution, what they fail to account for are the interactional and pragmatic forces that also shape grammar. Language is best described when examined in its natural setting, in interaction between two or more people, and the patterns of usage are best understood bearing in mind the social nature of language use.

### References

- Ashby, William J. and Bentivoglio, Paola. 1993. Preferred argument structure in spoken French and Spanish. *Language variation and change* 5: 61-76.
- Bayley, Robert and Lucinda Pease Alvarez. 1997. Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change* 9(3): 349-371.
- Bentivoglio, Paola. 1987. *Los sujetos pronominales de primera persona en el habla de Caracas*. Caracas: Universidad Central de Venezuela, Consejo de Desarrollo Científico y Humanístico.
- . 1993. Full NPs in spoken Spanish: A discourse profile. In William Ashby, Marianne Mithun, Giorgio Perissinotto and Eduardo Raposo (eds). *Linguistic perspectives on the romance languages*, 212-224. Amsterdam: John Benjamins.
- Bills, Garland D. and Vigil, Neddy A. 1999. Ashes to ashes: The historical basis for dialect variation in New Mexican Spanish. *Romance Philology*, (53): 46-67.
- Cameron, Richard. 1994. Ambiguous agreement, functional compensation and nonspecific *tú* in the Spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language Variation and Change* (5) 305-334.
- . 1995. The scope and limits of switch reference as a constraint on pronominal subject expression. *Hispanic Linguistics* 6-7: 1-27.
- Cameron, Richard, and Nydia Flores-Ferrán. 2003. Perseveration of subject expression across regional dialects of Spanish. *Spanish in Context* 1(1): 41-65.
- Du Bois, John. 1980. Beyond definiteness: The trace of identify in discourse. In Wallace Chafe (ed.), *The pear stories: Cognitive, cultural and linguistic aspects of narrative production*, 203-274. Norwood, NJ: Ablex.
- . 1987. The discourse basis of ergativity. *Language*, 63:805-855.
- Enríquez, Emilia V. 1984. *El pronombre personal sujeto en la lengua española hablada en Madrid*. Madrid: Consejo Superior de Investigaciones Científicas, Instituto Miguel de Cervantes.
- Flores-Ferrán, Nydia. 2002. *Subject personal pronouns in Spanish narratives of Puerto Ricans in New York City: A sociolinguistic perspective*. Munich: Lincom Europa.

- Givón, Talmy. 1983. "Introduction", in Givón, Talmy (ed.): *Topic continuity in discourse*, 4-41. Amsterdam, Benjamins.
- Labov, William. 1966/1982. *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- . 1971. Some principles of linguistic methodology. *Language in Society*, 1(1), 97-120.
- . 1984. Field methods of the project on linguistic change and variation. In J. Baugh and J. Sherzer (eds), *Langugae in Use: Readings in Sociolinguistics*, 28-54. Englewood Cliffs, NJ: Prentice-Hall.
- Levy, Yonata and Anne Vainikka. 1999/2000. The development of a mixed null subject system: A cross linguistic perspective with data on the acquisition of Hebrew. *Language Acquisition* 8(4): 393-384.
- Ocampo, Francisco. 1990. The pragmatics of word order in constructions with a verb and a subject. *Hispanic Linguistics* 4(1) 87-127.
- . 1995. The word order of two-constituent constructions in spoken Spanish. In P. Downing and M. Noonan (eds.), *Word order in discourse*, 425-447. Amsterdam: John Benjamins.
- Rand, David and David Sankoff. 1990. GoldVarb 2.0. Montreal: University of Montreal.
- Sánchez-Ayala, Ivo. 2005. The Spanish presentative verbal lexicon. Manuscript.
- 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.
- Travis, Catherine. 2005. The yo-yo effect: Priming in subject expression in Colombian Spanish. In Edward J. Rubin and Randall Gess (eds), *Selected papers from the 34<sup>th</sup> Linguistic Symposium on Romance Languages (LSRL), 2004*, 329-349. Amsterdam: John Benjamins.
- . To appear. The effects of interaction on structural priming: Subject expression in Spanish conversation vs. narrative. *Language Variation and Change*.
- Vallduvi, Enric and Maria Vilkkuna. 1998. On rheme and Kontrast. In P. Culicover and L. McNally *The limits of syntax. Syntax and Semantics* 29, 79-108. New York academic press.
- Vainikka, Anne and Yonata Levy. 1999. Empty subjects in Hebrew and Finnish. *Natural Language and Linguistic Theory* 17:613-671.