

ROMANCE LANGUAGES AND MODERN LINGUISTIC THEORY

**AMSTERDAM STUDIES IN THE THEORY AND
HISTORY OF LINGUISTIC SCIENCE**

General Editor
E.F. KONRAD KOERNER
(University of Ottawa)

Series IV - CURRENT ISSUES IN LINGUISTIC THEORY

Advisory Editorial Board

Henning Andersen (Los Angeles); Raimo Anttila (Los Angeles)
Thomas V. Gamkrelidze (Tbilisi); John E. Joseph (College Park, Md.)
Hans-Heinrich Lieb (Berlin); Ernst Pulgram (Ann Arbor, Mich.)
E. Wyn Roberts (Vancouver, B.C.); Danny Steinberg (Tokyo)

Volume 91

Paul Hirschbühler and Konrad Koerner (eds)

Romance Languages and Modern Linguistic Theory

ROMANCE LANGUAGES AND MODERN LINGUISTIC THEORY

PAPERS FROM THE 20TH LINGUISTIC SYMPOSIUM
ON ROMANCE LANGUAGES (LSRL XX)
OTTAWA, 10-14 APRIL 1990

Edited by

PAUL HIRSCHBÜHLER and KONRAD KOERNER
University of Ottawa, Ottawa, Canada

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM/PHILADELPHIA

1992

Library of Congress Cataloging-in-Publication Data

Romance languages and modern linguistic theory : papers from the 20th Linguistic Symposium on Romance Languages (LSRL XX), Ottawa, 10-14 April 1990 / edited by Paul Hirschbühler and Konrad Koerner.

p. cm. -- (Amsterdam studies in the theory and history of linguistic science. Series IV, Current issues in linguistic theory, ISSN 0304-0763; v. 91)

Includes indexes.

1. Romance languages--Congresses. 2. Linguistics--Congresses. I. Hirschbühler, Paul. II. Koerner, E.F.K. III. Linguistic Symposium on Romance Languages (20th ; 1990 ; Ottawa, Ont.)

PC11.R64 1992

440--dc20

92-25813

ISBN 90 272 3591 0 (Eur.)/1-55619-148-0 (US)(alk. paper)

CIP

© Copyright 1992 - John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 75577 · 1070 AN Amsterdam · The Netherlands
John Benjamins North America · 821 Bethlehem Pike · Philadelphia, PA 19118 · USA

FOREWORD

Habent libelli sua fata, and the present volume is no exception to this venerable adage. However, rather than recounting the genesis — and the pain — involved in the preparation of the camera-ready text of this 420-page book, I would like to say a few words about the origin of the papers here united as well as seize the opportunity to say a few words about the publishing history of previous LSRL meetings.

As on earlier occasions, this volume of papers from the 20th Linguistic Symposium on Romance Languages, held at the bilingual (English/French) University of Ottawa, in Canada's capital, in April 1990, reflects the state of North American Romance linguistics carried out within a broadly defined generative framework. As is obvious from the length of several articles, some of the papers included here remain close to the size and scope of the original presentation at the conference, whereas others constitute significant elaborations of the original papers, thereby in effect producing research articles of some importance. This might not have been the original intent of the conference organizers, but the delay in the production of the proceedings and ongoing research commitments made it desirable for several authors to revise their contributions considerably. In the end, it appears that the result was a better volume than it might have been otherwise.

As general editor of the series in which the present proceedings from a LSRL meeting appears, I would like to say something about the volumes of papers from such symposia which have thus far been published therein, and what my intent had been from the beginning. Traditionally, if one ignores university presses that tend to be well cushioned by the very fact that their actual overhead costs (equipment, salaries of employees, office space, maintenance, etc.) are covered by the general budget of the institution, regular publishers must make enough money to remain in business. Editors of a book series, if they understand their job well (and are not in a conflict-of-interest situation as sometimes happens), must have the health of the discipline at heart, which means in effect to defend the interests of the consumer, i.e., students and scholars. I am not thinking so much about the relatively high price these days of a book in linguistics (and in most other academic disciplines) about which we hear complaints quite regularly — people tend to forget that the cost of liv-

ing generally has rising dramatically over the years and it seems unreasonable to expect that books should remain at a 1975 level — but about the service to researchers by bringing out volumes of work in particular areas in one and the same place, so that also those who had no opportunity to participate in a given meeting can subsequently follow the trends of the field in question without any difficulty. This is, I thought back in 1982, when I attended an LSRL meeting for the first time, best accomplished by having the proceedings published in the same series on a regular basis. Before that time the papers from such symposia appeared in a variety of places (if at all), usually with a publisher near the conference venue. As a result, several of the pre-1982 proceedings have become difficult to track down and consult. Even after the 1982 meeting, I have been unable to persuade the conference organizers to submit the papers for publication in “Current Issues in Linguistic Theory” (CILT), but it was not for lack of trying. For instance I traveled to Los Angeles to attend LSRL XIV in order to meet with the organizers and discuss technical matters regarding the preparation of the papers for publication. The result is in the public domain. I am hoping, however, in the interest of continuity and the reasons stated above, that from LSRX XVII (1987) onwards the bulk of papers presented at these annual meetings will appear regularly in CILT — like those from Urbana-Champaign, Illinois, 1988 (CILT 69), Columbus, Ohio, 1989 (CILT 74), and the present volume.

Even though responsibility for the contents and form of the final product rests with the contributors and the editors, help has been received, quite generously at times, from graduate students of the Department of Linguistics, University of Ottawa, in the key-boarding of several papers not on disk, the transfer from an IBM to a Macintosh system of a number of submissions, implementing corrections, and the production of the indexes (prepared on the basis of documents originally developed by Douglas A. Kibbee of the University of Illinois for the LSRL XVIII proceedings): Pierre Carboneau, Kofi Saah, Jean Veall, and Badia Zahouni. Several ‘trees’ were produced by the co-editor of the present volume and our colleague John T. Jensen. Technical assistance was cheerfully given by Daming Xu. Financial assistance, first for the holding of the 1990 conference, then for the preparation of the manuscripts for publication, was received from the Social Sciences and Humanities Research Council of Canada, the School of Graduate Studies and the Department of Linguistics, University of Ottawa. They all deserve our gratitude.

Ottawa, Ontario, 5 February 1992

Konrad Koerner

CONTENTS

Foreword	v
Alexandra Alvarez	
Creole Interference in Venezuelan Spanish: The absence of <i>ser/estar</i> .	1
Paola Bentivoglio	
Linguistic Correlations between Subjects of One-argument Verbs and Subjects of More-than-one-argument Verbs in Spoken Spanish	11
Denis Bouchard	
Psych Constructions and Linking to Conceptual Structures	25
J. Clancy Clements	
Semantics of Control, Tense Sequencing, and Disjoint Reference . . .	45
Terry B. Cox	
Morphological Evolution in Molese, a Southeast Italian Dialect: Simpli- fication or decay?	57
Donka F. Farkas	
On the Semantics of Subjunctive Complements	69
Lorie Heggie	
On Reference in Copular Contexts	105
Julia Herschensohn	
On the Economy of Romance Nonlexical Datives	123
Paula Kempchinsky	
The Spanish Possessive Dative Construction: θ -role assignment and proper government	135
Jurgen Klausenburger	
Three Problems in Italian Inflectional Morphology	151
France Martineau	
The Evolution of Complements of French Causative and Perception Verbs	161
Pascual José Masullo	
Antipassive Constructions in Spanish	175

Jorge E. Porras	
Reflexification in Palenquero	195
Lisa Reed	
On Clitic Case Alternations in French Causatives	205
Johan Rooryck	
On the Distinction between Raising and Control	225
Mario Saltarelli	
The Subject of Psych-Verbs and Case Theory	251
Bernard Tranel	
On Suppletion and French Liaison	269
Esthela Treviño	
Subjects in Spanish Causative Constructions	309
Daniel Valois & Fernande Dupuis	
On the Status of (Verbal) Traces in French: The case of stylistic inversion	325
Dieter Wanner	
Subjects in Old Spanish: Conflicts between typology, syntax, and dynamics	339
Marc Wilmet	
Sur les articles <i>le/les</i> génériques ou l'énigme du cosmonaute	373
Karen Zagona	
Tense, V-Movement, and Wh-in-situ	389
Index of Authors	403
Index of Subjects & Terms	409
Index of Languages & Language Families	415

CREOLE INTERFERENCE IN VENEZUELAN SPANISH THE ABSENCE OF *SER/ESTAR**

ALEXANDRA ALVAREZ
Universidad Central de Venezuela, Caracas

0. *Introduction*

This paper will present the results of investigations carried out in Venezuelan Spanish regarding the presence vs. absence (henceforth: \emptyset *ser/estar*) of forms of the verbs *ser/estar*, a case of linguistic variation in the Spanish spoken in Caracas. I refer to sentences such as the following:

- (1) a. ... *el allà no explota, no dice nada, ni llama la atención ni nada.*
El \emptyset tranquilo como si no le importara. 1AM554
“There he doesn’t burst, he says nothing, he doesn’t call your attention or anything. He [is] calm as if nothing would bother him.”
- b. ... *el allà no explota, no dice nada, ni llama la atención ni nada.*
El está tranquilo como si no le importara.
- (2) a. *A mi no me gusta meterme en problemas. \emptyset un tipo de trabajo, echo vaina por ahí, mi novia y tal, y mi rancho* 1BH541
“I don’t like to get into trouble. I [am] a working man, I have fun, I have my girlfriend, and my hut.”
- b. *Soy un tipo de trabajo, echo vaina por ahí, mi novia y tal, y mi rancho.*

My hypothesis is that (1a) and (2a) are structurally very similar to (1b) and (2b) but lack a form of the verbs *ser/estar* (henceforth also called *copula*).

* I would like to acknowledge my debt to Ralph Fasold for his teaching and constant encouragement; to Paola Bentivoglio for her reading of the manuscript and her always valuable suggestions; to the Instituto de Filología “Andrés Bello” of the Universidad Central de Venezuela, for allowing the use of materials of the Estudio Sociolingüístico de Caracas. My thanks go also to the Consejo de Desarrollo Científico y Humanístico for the grant in support of this study.

There are no major differences in the referential meaning of both versions of these sentences.

The data have been extracted from 48 tape recordings taken from the Sociolinguistic Corpus of Caracas (IFAB), of women and men of three socio-economic groups (high, middle, and low), and of two age groups which I have considered 'young' (15-35) and 'old' (36-60).

Part 1 describes the methodology used in this study. Part 2 analyses the feature in Venezuelan Spanish. Part 3 proposes explanations concerning its origin. In part 4, I will offer some conclusions.

1. *Methodology*

In the frame of this study, I will assume that the 'missing' *ser/estar* forms function either as:

- i. a copulative verb with a following noun phrase, or predicate adjective;
- ii. an existential with a following locative; or
- iii. an auxiliary, with a following gerund.

This classification is, in general terms, based on the studies of Labov (1969) and Poplack & Sankoff (1987) on Vernacular Black English. The structures \emptyset + *negative* and \emptyset + *gonna*, are, of course, not possible in Spanish. Therefore, *gonna* and *negatives* will only be taken into account when speaking about English-based creoles.

In Venezuelan Spanish, the \emptyset construction has the following structures:

i) *Copulative use:*

1. \emptyset + *noun phrase*
 - (3) *A mí no me gusta meterme en problemas. \emptyset un tipo de trabajo, echo vaina por ahí, mi novia y tal, y mi rancho* IBH541
 "I don't like to get into trouble. I [am] a working man, I have fun, I have my girlfriend, and my hut."
2. \emptyset + *predicate adjective*
 - (4) "... *el allá no explota, no dice nada, ni llama la atención ni nada* *El \emptyset tranquilo como si no le importara.* IAM554
 "There he doesn't burst ... he says nothing, he doesn't call your attention or anything. He [is] calm as if nothing would bother him."

ii) Existential use:3. \emptyset + *locative*

- (5) *Bueno, la primaria la hice en ... un cole ... en tres colegios, pero yo dije ... este ... que empecé desde segundo ahí, desde segundo, desde segundo grado, porque el primero, el preparatorio, el kinder y ... pre ... prekinder y todo eso \emptyset por los Caobos, que nosotros vivíamos por allá.* IBM546

“Well, elementary school, I studied in three schools, but I said ... I started there since I was in second grade, in second, in second, because the first grade ... kindergarten, preparatory, pre-kindergarten and all that stuff [was] in Los Caobos, ‘cause we lived there.”

iii) Use as an auxiliary:4. \emptyset + *gerund* (Only *estar* cases).

- (6) ... ahorita tú ves a los muchachos vale que ... ‘*Que pava, cómo estás?*’ y broma, un besito y \emptyset pellizcándole el cachete y broma. IBH538

“Now you see the kids, man. ‘Hi baby, how are you?’ and stuff. A kiss and they pinch (are pinching) their cheek and stuff.”

For the analysis, I took into account only those structures which provided an adequate environment for the absence of *ser/estar*: only *main clauses* with one of the preceding contexts:

- (7) (i) *a noun phrase*:: Juan, papá, la comida internacional, yo, nosotros, él, uno todo, lo peor de todo
 (ii) *a conjunction*:: y, o
 (iii) *a sentence adverb*: quizás, entonces, allá, después, siempre, también, así
 (iv) *a prepositional phrase*, with a locative function: a esta hora, en el mercado de Chacaíto

2. Analysis

As already mentioned, the occurrence of \emptyset *ser/estar* was studied in the corpus with relation to three socioeconomic groups, two age groups and two sexes. Of the total of 48 informants, 39 of them, that is 81.25%, use zero copula. The total use of zero copula vs. copulative verbs can be seen in (8).

- (8) Table 1: *Ser/estar* presence vs. absence

presence	absence	total
1.037	108	1.145
91%	9%	100%

When correlating the variables: age, sex, and socioeconomic group (henceforth: SEG), as in (9) \emptyset *ser/estar* appears as a low-class indicator in the older generation. In fact, it is least used by the high group, slightly more by the middle group and most by the lower SEG. In the younger generation, however, there is an increase within the men's group, which might show a change in attitude towards *zero copula* (Labov 1966).

(9) Table 2: *Absence of copula according to age, sex, and SEG*

	young				old			
	men		women		men		women	
	abs	%	abs	%	abs	%	abs	%
high	22	12	15	10	1	1	6	3
middle	40	18	11	6	4	2	9	4
low	26	15	27	12	15	8	26	17
Total	88	15	53	28	20	11	41	24

As to the factors following the variable, I discovered a relationship as can be seen in (10).

(10) Table 3: *Relationship among constraints in ser/estar variation*

		copula	absence
Auxiliary	(Gerund)	82.04%	17.96%
Copulative	(NP)	88.79%	11.21%
Copulative	(PA)	94.01%	5.99%
Existential	(Locative)	95.39%	4.61%

The constraints favoring the absence of *ser/estar* are from large to small, namely: *gerund* < *nominal phrase* < *predicate adjective* < *locative*.

(These results will be discussed in section 3.2. below.)

3. *Possible explanations: \emptyset S/E, a creole feature?*

According to Spanish grammars, copula absence is not rare in *all* standard environments. Real Academia Española (1973:365) and Seco (1969:193-194) describe lack of the copulative verbs in the following cases:

- (15) a) Interrogative sentence:
Tú, amigo suyo? (RAE:365)
 b) Exclamative sentences:
Quien más honesto y más valiente que el famoso Amadís de Gaula!
 and

c) Proverbs:

Mal de muchos, consuelo de tontos; Perro labrador, nunca mordedor.

These examples, which I consider literary devices, do not seem, however, to be related to the \emptyset *S/E* examples which I will discuss in what follows. As I see it, the existence of \emptyset *ser/estar* in Venezuelan Spanish could possibly be explained in two ways:

1. The use of \emptyset *ser/estar* is an extension of the standard absence of copula allowed in exclamative and interrogative sentences — cited above — to declarative sentences. This means that the use of *zero copula* in Venezuelan Spanish is a standard Spanish feature. The change would be due exclusively to internal forces in the Spanish language.

2. The absence of *ser/estar* is not a feature of Spanish origin. Sociohistorical reasons related to former African slavery in the Caribbean allow the presumption that a creole spoken among slaves may have influenced the Spanish of the region. In recent studies, some other low status varieties of Spanish, such as Cuban ‘popular Spanish’ (Perl 1989a) and Brazilian ‘popular Portuguese’ (Guy 1989) have been considered as having a creole ancestor.

In this line of thought, \emptyset *ser/estar* would be due to contact with a former creole substrate. Its use in Venezuelan Spanish and other varieties in the region might be explained, at least partially, as the result of external forces of linguistic change. In other words, \emptyset *copula* can be understood as a case of syntactic interference.

3.1 *Zero copula in Spanish-based creoles*

Copula absence has been observed in Spanish-based creoles, especially in Palenquero by Granda (1978), in the Portuguese trade pidgin by Naro (1978), in Habla Bozal Antillana by Otheguy (1973). Lipski (1989) records the elision of *ser* and *estar* in the speech of Panamá Congos, as well as in Teatro-Bufo of the 19th century as characteristic of the speech of blacks. The same is observed by Perl (1989a, 1989b) in Habla Bozal Cubana.

Examples (12), (13) and (14) are from Lydia Cabrera’s *El Monte*, cited by Perl (1989a):

- (12) *Palo duro guayacán*
 [El] guayacán [es] [un] palo duro
 Guayacán [is] a hard type of wood.
- (13) *Pavo Real ta bucán palo*
 The peacock [is] looking for wood.

- (14) *mi caballito ta estropiao*
My small horse [is] hurt.

In examples (13) and (14) *ta* is a present-tense marker, found in most Caribbean creoles.

3.2 Zero copula in English-based creoles

The feature *zero copula* has also been observed in English-based creoles. Nichols (1981) found it in Gullah, and Baugh (1983) noticed parallel constructions in Gullah and in Jamaican English. With regard to English, copula absence has been a controversial topic. Labov's (1969) article on Vernacular Black English (henceforth: VBE) fomented numerous studies favoring or countering the creole origin hypothesis of this variety. Recently, the discussion concerning copula absence has been centered on the ranking of the factors, which is said to weaken the creole-origin hypothesis of VBE because it does not follow the so-called *creole-order*, with high deletability of the copula before a predicate adjective (Poplack & Sankoff 1987).

The Caracas corpus shows higher percentages of deletion with following noun phrase and gerund, and smaller percentages in the use of predicate adjective and existential/locative. This ranking also does not follow the so-called *creole-order* but rather the ranking found by Labov (1969) in VBE. One interesting result in the Caracas corpus is the strong increase in the use of \emptyset *ser/estar* + *gerund*.

These results could suggest: a) that there is no creole interference in Venezuelan Spanish, or b) that we deal here with the very early stages of a decreolization process. There are, to my knowledge, no analyses of the contextual factors of copula absence in Spanish-based creoles. Studies in this direction will have to be awaited to reach conclusions. In any case, if a creole base should be posited for Venezuelan Spanish, this variety would be nearer to the acrolect than any of the varieties it has been compared to, and it may therefore follow the standard Spanish structure more closely than other cases studied.

3.3 New insights: \emptyset S/E, a mesolectal feature?

There is evidence that zero copula is not a feature of the *basilect* of creole languages, but rather of mesolectal varieties. (Bickerton 1975, Fasold 1990). This is also contemplated by Perl (1989a, b) with regard to Habla Bozal Cubana.

Regarding Guyanese Creole, Fasold suggests that “*zero copula* is not part of the Guyanese creole grammar itself, but is introduced into the mesolects due

to a clash between the existing basilect rules and the incoming acrolect system” (p. c.). This idea is interesting for the Venezuelan case where slavery endured until the 19th century. It is possible that a basilectal creole form may have had the structure

Pro + INFLECTION + V

which we now find in most Spanish-based creoles, as stated by Perl (1989). The case of today’s Papiamentu, with a gerund as in *e ta estudiando* (he is studying), or *i ta kantando ele* (I am singing to her), would correspond to an early *mesolectal* structure.

The Venezuelan Spanish form *yo comiendo* may be considered a later, *mid-mesolectal* form. Finally, we have the *acrolectal* form:

(Pro) + ser/estar + NP/PA/gerund/LOC

as in *(yo) estoy comiendo*, which corresponds to the standard. The diachronic pattern I am suggesting is as stated in:

- (15) i) *mi ta comé*
 Pro INFL verb
 ii) *e ta kantando*
 Pro INFL gerund
 iii) *yo Ø S/E comiendo*
 iv) *(yo) estoy comiendo*

One argument favoring this claim is that the particle *ta* functions in Spanish based creoles both as a copulative, as in (12) and as an auxiliary as in (13). Maurer (1989) observes, in Palenquero, an alternation in the present-tense between *ta* and *[Ø]*. The appearance of one or the other form depends on the type of verb used.

An interesting point to consider is the fact that the corpus shows sentences with the ‘wrong’ copula, such as in examples (16)a and (17)a:

- (16) a. ... *todo el mundo era así viéndolo*. 2MM527
 ... everybody was watching it.
 b. *todo el mundo estaba así viéndolo*
- (17) a. ... *todo el año es vagando* 1AH561
 I do nothing during the year.
 b. *todo el año estuve vagando*

Note that (16b) and (17b) are the standard Spanish versions. These cases cannot be considered simple performance errors, since they are often heard in Caracas. They are rather examples of linguistic insecurity, I would argue, which could favor the creole origin hypothesis. In fact, they may point to some former 'erasure' of semantic limits between *ser* and *estar*, as in the case documented by Fernandez Marrero (1989) in *Habla Bozal Cubana* for *sar*, as shown in (18) and (19):

(18) *Mañana sa Corpus Christi.*
Tomorrow is Corpus Christi.

(19) *Samo negra pecandora.*
[We] are black sinners.

My hypothesis needs, of course, further study. Also to be examined is the possibility of a reanalysis of the forms, showing presence or absence of *S/E* in Venezuelan Spanish. It seems, in fact, that \emptyset *S/E* is being used evaluatively, while *S/E* is being used descriptively.

5. *Conclusions*

I have studied presence vs. absence of *ser/estar* in Venezuelan Spanish in its copulative, existential and, auxiliary use. The corpus consists in main sentences preceded by an NP, a conjunction, or a sentence-adverb. The variable was considered to be comparable to 'zero copula' or 'copula absence' of creole languages.

There are two possible explanations for the presence of this feature in Venezuelan Spanish: first, that it is due to internal developments of the language. The second, which is favored in this paper, is that it may be a feature of creole origin. There is evidence that the feature *zero-copula* is present in varieties of Spanish-based creoles. The ranking of the factors following \emptyset *S/E* in Venezuelan Spanish, which could be considered a counterargument to this hypothesis, has not yet been studied in Spanish-based creoles, and further evidence has to be awaited to reach any safe conclusion.

Zero copula seems to be not a basilectal but a mesolectal feature, borrowing Fasold's words, "due to a clash between the existing basilect rules and the incoming acrolect system" (Fasold, p. c.).

The overall socioeconomic pattern of greater use of zero copula in the low socioeconomic group supports the idea of a creole continuum. The sociolinguistic variation identifies this variable as a stigmatized feature, especially

among speakers of the older generation, which seems to constitute a point in favor of this claim. The greater use of the variable by young people leads to the question whether it is a case of refocussing towards the hypothesized previous creole (Lepage 1980). Recent research in Palenquero (Schwegler 1989) too suggests the presence of a tendency to reaffirm roots in order to stress local identity.

REFERENCES

- Baugh, John. 1983. *Black Street Speech: Its history, structure, and survival*. Austin, Tex.: Univ. of Texas Press.
- Bickerton, Derek. 1975. *Dynamics of a Creole System*. Cambridge: Cambridge Univ. Press.
- Fasold, Ralph. 1981. "The Relation between Black and White Speech in the South". *American Speech* 56:3.163-189.
- , 1990. *The Sociolinguistics of Language*. Oxford: Basil Blackwell.
- Fernández Marrero, Juan Jorge. 1989. "Ein historiographisches Dokument zur 'Lengua Bozal' in Kuba: *El Catecismo de los negros Bozales* (1795)". Perl ed. 1989.53-60.
- Granda, Germán de. 1978. *Estudios lingüísticos hispánicos, afrohispanicos y criollos*. Madrid: Gredos.
- Guy, Gregory. 1989. "On the Nature and Origins of Popular Brazilian Portuguese". *Estudios sobre español de América y lingüística afro-americana*, 227-245. Bogotá: Instituto Caro y Cuervo.
- Labov, William. 1969. "Contraction, Deletion and Inherent Variability of the English Copula". *Language* 45:4.715-762.
- LePage, Robert B. 1980. "Theoretical Aspects of Sociolinguistic Studies in Pidgin and Creole Languages". *Theoretical Orientations in Creole Studies* ed. by Albert Valdman & Arnold Highfield, 331-367. New York: Academic Press.
- Lipski, John. 1989. "El lenguaje de los congos panameños: ¿Vestigios de un criollo afrohispanico?". *Actas del VII Congreso de la Asociación de Filología y Lingüística de América Latina (ALFAL)*, vol.II, 63-88. Santo Domingo: Univ. Nacional de Pedro Henríquez Ureña.
- Maurer, Philippe. 1989. "Der Vergleich der Tempussysteme des Papiamentu und des Palenquero: Argument für die Monogenese?". Perl ed.1989.33-50.
- Naro, Anthony 1978. "A Study on The Origins of Pidginization". *Language* 54:2.314-347.
- Otheguy, Ricardo. 1973. "The Spanish Caribbean: A Creole perspective". *New Ways of Analysing Variaton in English* ed. by Charles-James N. Bailey & Roger W. Shuy, 323-339. Washington, D.C.: Georgetown Univ. Press.

- Perl, Matthias. 1989a "Algunos resultados de la comparación de fenómenos morfosintácticos del 'habla bozal', de la 'linguagem dos musseques', del 'palenquero' y de lenguas criollas de base portuguesa". *Estudios sobre español de América y lingüística afroamericana*, 369-380. Bogotá: Instituto Caro y Cuervo.
- , 1989b. "Habla Bozal" — eine Spanish-basierte Kreolsprache?". Perl ed. 1989.17-31.
- , ed. 1989. *Beiträge zur Afrolusitanistik und Kreolistik*. Bochum: Norbert Brockmeyer.
- Poplack, Shana & David Sankoff. 1987. "The Philadelphia Story in the Spanish Caribbean". *American Speech* 62:4.291-314.
- Real Academia Española. 1973. *Esbozo de una nueva gramática de la lengua española*. Madrid: Espasa-Calpe.
- Rickford, John. 1977. "The Question of Prior Creolization in Black English". *Pidgin and Creole Linguistics* ed. by Albert Valdman, 190-221. Bloomington: Indiana Univ. Press..
- Seco, Rafael. 1988. *Manual de Gramática española*. Bogotá: Aguilar.
- Schwegler, Armin. 1989. "El Palenque de San Basilio (Colombia): Persistencia africana y problemas de (auto-)identificación de elementos lingüísticos subsaháricos". To appear in *Persistencia Africana en el Carib*. San Juan, P.R.: Univ. de Puerto Rico, Rio Piedras.
- Silva Corvalán, Carmen. 1987. *Sociolingüística: Teoría y análisis*. Madrid: Alhambra.

**LINGUISTIC CORRELATIONS BETWEEN
SUBJECTS OF ONE-ARGUMENT VERBS AND
SUBJECTS OF MORE-THAN-ONE-ARGUMENT VERBS
IN SPOKEN SPANISH***

PAOLA BENTIVOGLIO
Universidad Central de Venezuela, Caracas

0. *Introduction*

The object of this paper is to empirically validate Du Bois (1985, 1987) hypothesis about two tendencies apparently present in every natural language which combined characterize the notion of 'Preferred Argument Structure' (PAS).

The data supporting Du Bois' statements are drawn from Sacapultec, an ergative Mayan language. Dutra (1987) showed that Du Bois' findings also hold for Brazilian Portuguese. In order to verify whether the notion of PAS is valid and consequently applicable to another Romance language, I conducted an analysis based on Caracas spoken Spanish, whose results constitute the core of this paper.

In the first part I will summarize the basic theoretical notions on which the research is based as well as its main purposes; in the second, I will sketch the methodology; in the third part, the outcome of the analysis will be discussed; finally, in the fourth, some conclusions will be offered.

* The data on which this paper is based derive from the Project "Estudio sociolingüístico del habla de Caracas" of 1987, carried out under the coordination of Mercedes Sedano and the author; it was sponsored by the Consejo de Desarrollo Científico y Humanístico of the Universidad Central de Venezuela (grants # H-07.16/86 and H-08.33.1766/88). The data have been processed by means of the Child Language Analysis (CLAN) program, which was generously offered to us by Dr. Catherine Snow of Harvard University, and efficiently implemented by Alfonso Mosquera. Evelyn Castro cooperated with the coding of the data according to the CLAN program; Sandra A. Thompson greatly contributed to the stylistic improvement of the final version of this paper. To all of them I wish to express my gratitude. It seems unnecessary to say that I am totally responsible for all errors.

1. *Antecedents and basic theoretical notions*

The notion of PAS is based on the direct arguments of a verb, that is, the subject and direct object. Following Dixon (1979) I will call *S* the subject of one-argument verbs, *A* the subject of more-than-one-argument verbs, and *O* the direct object. Du Bois affirms that two tendencies, both related to the presence or absence of full noun phrases (*NP*) in the same clause, are responsible for the PAS of natural languages:

- i) according to the first, called the One Lexical Argument Constraint (*OLAC*),¹ clauses tend to contain no more than one full noun phrase as direct arguments; clauses with fewer than one full *NP* (i.e., pronouns or zeros) occur frequently, but two full *NPs* are “extremely rare” (Du Bois 1985:348);
- ii) according to the second one, called the Non-lexical Argument Constraint (*NAC*), the single full *NP* tends to be either the subject (*S*) of one-argument-verb clauses, or the object (*O*), but not the *A* subject of more-than-one-argument-verb clauses.

Du Bois (1985:349) adds that

- (1) [...] the shape that this preferred argument structure takes is statable in terms of specific limits on the number of direct argument noun phrases in a clause and on the syntactic roles in which these noun phrases are likely to occur.

Summarizing, even though more than one *NP* (fulfilling either *S*, *A* or *O* roles) could be present in a clause, a high percentage of clauses only have at most one full *NP*, that tends to be an *S* or an *O*, but rarely an *A*. These tendencies have proven to hold in a number of structurally different languages, both ergative and non-ergative, such as English, French, German, Brazilian Portuguese, Modern Hebrew, Japanese, Papago, Quechua, and Rama.²

Du Bois' statements, complemented by Dutra's fruitful application to Portuguese, convinced me that the same analysis could be successfully applied to Spanish.

The main purpose of the present paper is to demonstrate that in spoken Spanish discourse the notion of ‘preferred argument structure’, as proposed by Du Bois (1985, 1987), is valid, hence it should be taken into account when

¹ It is necessary to clarify that Du Bois (1985:348) uses the term constraint to indicate “a strong statistical tendency in text: tokens, rather than a categorical — or even variable — rule in grammar”.

² The list of languages and the related references appear in Du Bois (1987).

analyzing the syntactic, semantic and pragmatic dimensions of the language. Because of space constraints, this paper will mainly concentrate on the syntactic dimension of Spanish PAS, leaving a more in-depth analysis of the semantic and pragmatic dimensions for future research.

2. *Methodology*

The corpus analyzed consists of the first one hundred (100) main declarative clauses³ extracted from the recorded 'casual' speech of fourteen Caracas-born speakers (age group 14 to 30) belonging to five socioeconomic levels. In the present study, however, due to the reduced sample taken into account, sociolinguistic variables will not be analyzed.

The fourteen hundred (1,400) clauses (which will be referred to henceforth as the *corpus*), were coded according to the following linguistic variables: i) the subject type (*S* or *A*); ii) the subject form: zero (\emptyset), pronominal (*P*), or nominal (*N*); iii) the object form: clitic (*CLI*), pronominal (*P*), or nominal (*N*); iv) the grammatical person of the subject referent: *1st*, *2nd* or *3rd*; v) the referent mention of subjects and objects: *first vs. previous mention*; and vi) the *position* of nominal subjects, viz. their verbs.

Note that for the purpose of the analysis I did not take into account those clauses which show some of the following characteristics: i) sentential subjects and objects; ii) verbs such as *gustar*, *importar*, etc., which second argument is a dative and not an accusative; iii) cognitive verbs (*creer*, *pensar*) verbs of saying (*decir*, *hablar*), impersonal verbs (*hay que*, *llover*, *tronar*, and the like), and the verb *agarrar* in lexicalized expressions (for example, *agarrábamos y prendíamos candela a todo* "we started burning everything"); iv) the form *se* that can be interpreted as an impersonal subject (for example, *se vive bien en Caracas* "one lives well in Caracas"); and v) all types of clefts (see Sedano 1990 for details).⁴ Clauses which are identical, i.e., a literal repetition, of antecedent ones, were also excluded from the present analysis.

Let us consider examples of *S* and *A* subjects in (2) and (3), respectively. In (2) *S* appears as zero in (a), as a pronoun in (b), and as a full *NP* in (c):

³ I have limited the analysis to main declarative clauses, in order to examine a homogeneous set of data.

⁴ Cleft constructions were excluded because of a number of properties according to which they differ, pragmatically and grammatically, from other types of constructions.

- (2) a. /S-Ø/
 ...Ø fui un pésimo estudiante,... (C114015.87)⁵
 [I] was a very bad student
 "...I have been a very bad student..."
- b. /S-P/
 ...yo antes vivía en Altamira,... (C113012.87)
 I before lived in Altamira
 "...before I was living in Altamira..."
- c. /S-N/
 ...toda mi familia se-mudó a Caurimare,... (C111004.87)
 all my family moved to Caurimare
 "...all my family moved to Caurimare..."

Examples in (3) illustrate A subjects in combination with an object, which appears, successively, as a clitic (*O-CLI*), as a pronoun (*O-P*) or as a noun (*O-N*). The A subject is zero in (a-c), a pronoun in (d-f), and, finally, a noun in (g-i).

- (3) a. /A-Ø + O-CLI/
 ...bueno, Ø me pegaban de vez en cuando... (C113012.87)
 well [they] me bit from-time-to-time
 "...well, they used to beat me from time to time..."
- b. /A-Ø + O-P/
 ...y entonces Ø aprovechamos todo... (C112005.87)
 and then [we] used everything
 "...and then we used everything ..."
- c. /A-Ø + O-N/
 ...y Ø tenían un tremendo carro... (C111012.87)
 and [they] had a terrific car
 "...and they used to have a terrific car ..."
- d. /A-P + O-CLI/
 ...él me pasaba buscando por mi casa. (C111001.87)
 he me passed-by looking for my home
 "...he used to come by my place and pick me up."

⁵ All examples are identified by the code of the speaker, which appears between parentheses at the end of each example.

- e. /A-P + O-P/
 ...y entonces yo agarré eso como guachafita... (C112005.87)
 and then I took that as fun
 "...and then I took that as fun ..."
- f. /A-P + O-N/
 ...y . uno [...] nota la diferencia... (C113012.87)
 and one notes the difference
 "...and one notes the difference..."
- g. /A-N + O-CLI/
 ...los números siempre me han tenido bajo... (C112005.87)
 the figures always me have kept low
 "...figures have always beaten me up..."
- h. /A-N + O-P/
 ...los muchachos y uno apaleaban al otro... (C114015.87)
 the boys and me hit the other
 "...the boys and oneself used to beat the other..."
- i. /A-N + O-N/
 ...muchos colegios [...] tienen esa disciplina bien desarrollada...
 many schools have that discipline well developed
 "... many schools [...] have that discipline well developed..."
 (C11004.87)

3. The analysis

3.1 The One Lexical Argument Constraint (OLAC)

The results of the analysis are shown in tables 1, 2 (a and b) and 3. In table 1 we appreciate that out of 1,400 main declarative clauses approximately two thirds of them have *S* subjects and only one third has *A* subjects. Table 2a depicts the form of *S* and *A* subjects, and Table 3 that of *O*.

<i>S</i>	%	<i>A</i>	%
920	66	480	34
total:		1,400	

Table 1: *Distribution of verb direct-arguments into S & A*

	<i>S</i>	%	<i>A</i>	%	Total	%*
\emptyset	458	58	331	42	789	56
<i>P</i>	227	69	102	31	329	24
<i>NP</i>	235	83	47	17	282	20
Total*	920		480		1,400	

Table 2a: Distribution of *S* & *A* subjects according to form (\emptyset , *P*, *N*)

* Percentages in this column do not represent the sum of percentages in the corresponding rows

	<i>No-NP</i>	%	<i>NP</i>	%	Total
<i>S</i>	685	74	235	26	920
<i>A</i>	433	90	47	10	480
Total	1,118		282		1,400

Table 2b: Distribution of *S* & *A* subjects into NPs & no-NPs

<i>CLI</i>	%	<i>P</i>	%	<i>N</i>	%	Total
111	23	26	5	343	72	480

Table 3: Distribution of *O* according to form (*CLI*, *P*, *N*)

The following comments are at order:

i) *Frequency of S*: as shown in Table 1, the number of *S* subjects almost doubles that of *A* subjects. It could seem a trivial observation; it is not, however, since the frequency of *S* subjects may well be related to word order phenomena;

ii) *S & A favor \emptyset and P forms*: this is clearly shown in Table 2b. The tendency, however, is greater for *A* (90%) than for *S* subjects (74%). In both instances, full *NPs* represent a significantly smaller percentage as compared to the total amount of zeros and pronouns: 10% of *A* and 26% of *S* subjects. This is an important result for future analyses;

iii) *O favors nominal forms*: in Table 3 it may be appreciated that, contrary to what happens for *S* and *A*, objects appear as *N* in the 72% of the cases, whereas *CLI* and *P* only represent the remaining 28%.

In order to answer the crucial question about the distribution of lexical arguments across the roles of *A*, *S*, and *O* — namely, is this distribution random or ‘skewed in some way’? (Du Bois 1987:821) — we may observe Table 4, where the role distribution of the corpus is presented:

<i>S</i>	%	<i>A</i>	%	<i>O</i>	%	Total
235	37.5	47	7.5	343	55	625

Table 4: *Distribution of lexical mentions among grammatical roles*

Table 4 shows very clearly that both *S* and *O* contain each a substantial proportion of lexical mentions, whereas *A* has a significantly smaller quantity of them (only 7.5%). These findings coincide with Du Bois’ for Sacapultec: 48% of lexical mentions in the *S* role, 46% in *O* and 6% in *A*. The greatest similarity appears in the *A* role: Spanish 7.5 vs. Sacapultec 6%.⁶

In order to verify whether the *OLAC* constraint is also valid for Spanish, the results appearing in tables 1-4 do not provide sufficient evidence, as they only show the distribution of lexical mentions and their roles. This would be enough if one would take into account only one-argument clauses, but is not sufficient when considering also more-than-one-argument clauses (cf. Table 2b). Among these latter we may note that 47 lexical mentions are present as *A*, and 343 as *O*. It is consequently necessary to know whether or not *A* and *O* lexical mentions coincide in the same clause. The outcomings of this analysis are shown in Table 5:

⁶ Most of the differences among the several corpora analyzed (Sacapultec, Brazilian Portuguese, and Spanish) may be due to the different methodology used in the data collecting process. Sacapultec and Brazilian Portuguese data consist of elicited narratives, following Chafe (1980). The Caracas corpus consists of semicasual conversations of one speaker in the presence of two interviewers. The main difference between the first two corpora, in one hand, and the Caracas one, in the other, resides in the difference of genres: mainly narrative in the first corpora (hence predominance of 3rd person), and a mixture of narrative with description and evaluations in the last one (hence predominance of 1st person).

	<i>O-CLI</i>	%	<i>O-P</i>	%	<i>O-N</i>	%	Total	%*
<i>A-Ø</i>	81	24	17	5	233	71	331	69
<i>A-P</i>	12	12	7	7	83	81	102	21
<i>A-NP</i>	18	38	2	4	27	7	47	10
Total	111		26		343		480	100

Table 5: *Distribution of A & O, according to their form*

* Percentages in this column do not represent the sum of percentages in the corresponding rows.

The cross-section between *A-N* and *O-N* — see the third row of the third column in Table 5 — proves that only 27 (or 7%) out of the 480 more-than-one-argument clauses examined have a nominal subject as well as a nominal object. This means that in a 1,400 — clause corpus two lexical mentions in the core arguments of the same clause only occur in the extremely low percentage of 5.62% (27/480). This result clearly validates for spoken Spanish the existence of the constraint proposed by Du Bois (1985), whose words seem very appropriate to end this section (p.349):

- (4) Sentences with two full noun phrases are extremely rare in Sacapultec discourse, and *perhaps in the purely spoken discourse of most languages* (emphasis added: PB), as preliminary research by myself and others suggests.

3.2 *The Non-lexical Argument Constraint (NAC).*

It is now necessary to validate the existence of the second tendency mentioned above, according to which the single lexical mention in a clause tends to be *S*, if the verb has one argument, or *O*, if the verb has more-than-one argument, but almost never *A*.

Let us consider again the percentages of Table 4 and compare them, in Table 6, with those obtained by Du Bois for Sacapultec and by Dutra for Brazilian Portuguese:

	Sacapultec %	Spanish %	Br. Portuguese %
<i>S</i>	48	37.5	39
<i>A</i>	6	7.5	6
<i>O</i>	46	55	53

Table 6: *Comparison of the percentages of lexical mentions distribution in three languages*

Examining Table 6, we observe striking similarities between the results obtained by Du Bois and Dutra, in one hand, and those from this study, in the other. Especially important, from a Romance linguistics viewpoint, is the comparison between the percentages obtained in the Brazilian Portuguese and Caracas analyses, respectively: 55% vs. 53% of lexical mentions in the role of *O*, 37.5% vs. 39 in the role of *S*, and 7.5 vs. 6% in the role of *A*.

The preceding observations permit us to affirm that the Non-lexical Argument Constraint is present in Caracas spoken Spanish. In fact, full NPs tend to occur as objects of transitive clauses and subjects of intransitive ones, but seldom as subjects of transitive clauses.

3.3 *The Preferred Argument Structure of Spanish*

The analysis described in the previous sections (cf. 3.1 and 3.2) validates, without any doubt, both tendencies on which the notion of PAS is based, thus demonstrating that the notion of ‘preferred argument structure’ — as set forth by Du Bois (1985, 1987) — is valid for spoken Spanish, and may consequently be used in the linguistic analyses of this language. This finding may be stated as in (5):

- (5) In spoken Spanish discourse, there is a preferred argument structure with respect to the verb’s direct arguments. This structure consists of the presence of only one NP in each clause; furthermore, this unique NP fulfills the role of subject in one-argument verb clauses, and the role of object — rather than that of subject — in more-than-one-argument-verb clauses.

4. *PAS and information flow*

4.1 *The given A constraint*

Du Bois (1985) justifies the existence of *PAS* on the basis of information flow (in Sacapultec and other languages as well). He states (p.350) that new mentions occur with relative freedom in the *S* and *O* roles, but almost never in the *A* role. The reduced presence of full *NPs* accomplishing the *A* role is due to the fact that this role is seldom used to convey 'new information'. Du Bois calls this limitation of *A* subjects 'the given *A* constraint'.

In order to validate the existence of the given *A* constraint for Spanish, it is necessary to examine only those forms that are likely to code 'new information'. This means excluding from the analysis all 1st and 2nd person tokens, as they refer to the speaker and the hearer and are, therefore, not relevant to the present discussion, since they can never be new.

	<i>S</i>	%	<i>Su-A</i>	%	Total	%*
<i>1st person</i>	398	62	247	38	645	46
<i>2nd person</i>	19	63	11	37	30	2
<i>3rd person</i>	503	69	222	31	725	52
Total	920		480		1,400	

Table 7: Distribution of *S* & *A*, according to the referent person.

* Percentages in this column do not represent the sum of percentages in the corresponding rows.

In table 7 we may appreciate that nearly half (48%) of all subjects refer to the speaker or the hearer, and 52% to other participants. In order to concentrate our attention on the given *A* constraint, out of the 725 occurrences of third person reference zero and pronominal forms must be detracted, as they are

used to code participants already mentioned in the discourse or 'known'.⁷ The results are presented in table 8:

	<i>S</i>	%	Su- <i>A</i>	%	Total	%*
\emptyset	164	54	139	46	303	42
<i>P</i>	108	74	39	26	140	19
<i>N</i>	231	83	44	17	282	39
Total	503	69	222	31	725	

Table 8: *Distribution of 3rd person subjects, according to their form*

* Percentages in this column do not represent the sum of percentages in the corresponding rows.

Table 8 shows that the great majority of *Ns* (83%) appears in the *S* role, and only 17% in the *A* role. This result confirms the validity of the given *A* constraint for spoken Spanish. In order to explain this tendency, I will try to establish the characteristics differentiating the *S* from the *A* ful *NPs*, as they cannot be distinguished on the basis of their syntactic function (both *S* and *A* are subjects). Let us examine, in table 9, these *NPs* according to the pragmatic function of first (*FM*) vs. previous mention (*PM*):⁸

⁷ For the purposes of the present paper, the term 'given' applies to referents that can be easily processed by the hearer because of previous mentions or thanks to the 'knowledge' shared by the speaker and the hearer (Givón 1984).

⁸ These terms, and the related classification, have been used in Bentivoglio & Weber (1986).

	<i>FM</i>	%	<i>PM</i>	%	Total	%*
<i>S</i>	141	60	94	40	235	83
<i>A</i>	19	40	28	60	47	17
Total	160	57	122	43	282	

Table 9: *Distribution of S & A subjects, according to FM/PM*

* Percentages in this column do not represent the sum of percentages in the corresponding rows.

Table 9 shows that *S* subjects code both kinds of participants, *FM* and *PM*, even though the *M* are more frequent than the *PM* subjects. The results for *A*-subjects are exactly the opposite: in fact, more continuous, i.e., already mentioned, participants tend to appear in this role. First mentions are less frequent. This finding confirms Du Bois' statement on the correlation between *A*-subjects and the flow of information.

4.2 Word order

What can an analysis which explains ergativity in a Mayan language contribute to the problem of word order in Spanish? It seems that, within the limits of this paper, it can be affirmed that separating syntactic subjects into *S* and *A* allows a more precise characterization of their use, as shown in table 10:

	<i>Su</i> ⁹ - <i>V</i>	%	<i>V-Su</i>	%	Total	%*
<i>S</i>	100	43	135	57	235	83
<i>A</i>	37	79	10	21	97	17
Total	137	49	145	51	282	

Table 10: *Distribution of S & A subjects, according to verb position*

* Percentages in this column do not represent the sum of percentages in the corresponding rows.

⁹ Note that the abbreviation *Su* is used here in order to avoid confusion with *S* role.

We may observe in table 10 that the proposed distinction of subjects into *S* and *A* enhances another of their differences: *S* subjects prefer the postverbal position, whereas *A* subjects tend to appear preverbally. With the purpose of correlating the pragmatic properties of *S* and *A* subjects, the results of tables 9 and 10 are collapsed into table 11:

<i>S</i>				<i>A</i>			
	<i>FM</i> %	<i>PM</i> %	Total	<i>FM</i> %	<i>PM</i> %*	Total	
<i>Su-V</i>	42 42	58 58	100	14 38	23 62	37	
<i>V-Su</i>	99 73	36 27	135	5 50	5 50	10	
Total	141 60	94 40	235	19 40	28 60	47	

Table 11: *Distribution of S & A subjects, according to FM/PM and verb position*

* Percentages in this column do not represent the sum of percentages in the corresponding rows.

Table 11 reveals that there are no great differences, when first- and previously mentioned *S* and *A* occur in preverbal position. The postverbal position is significant only for first-mentioned *S* subjects which tend to occupy it 73% of the cases, whereas *A* subjects show no preferences at all (50% for both *FM* and *PM*). The most interesting result of this table is the clear tendency of first-mentioned *S* to appear in postverbal position. Further research is needed in order to verify whether and how *S* subjects differ from *A* subjects, in terms of information flow and word order (see Bentivoglio 1990).

5. Conclusion

The analysis conducted on a sample of Caracas speech has proved that *PAS* is valid for spoken Spanish, as Du Bois' three tendencies, attested in several natural languages, are also present in spoken Spanish:

i) only one, and no more than one, *NP* tends to be present in the direct arguments of a clause, thus confirming the One Lexical Argument Constraint proposed by Du Bois;

ii) if an *NP* is present, it tends to fulfill the role of subject in a one-argument-verb clause, or that of object in a more-than-one-argument-verb clause, but very rarely appears in the role of an *A*-subject; this confirms Du Bois' Non lexical Argument Constraint;

iii) a first-mentioned *NP* tends to occur as a *S* subject or as an object, but almost never as an *A* subject; the results of the analysis validate Du Bois' given *A* Constraint.

As a corollary to the mentioned constraints, the analysis has also shown that a first-mentioned subject referent tends to occur in an *S* role, whereas a previously mentioned subject tends to appear as an *A*. Furthermore, *A* subjects occur preverbally more frequently than *S* subjects.

From a more general viewpoint, the present study has revealed the convenience of analyzing a non-ergative language as Spanish according to the different categories of subject — *S* and *A* — found to be useful in the description of ergative languages. This paper constitutes the very first step towards a more in-depth analysis, which is the object of my ongoing research.

REFERENCES

- Bentivoglio, Paola. 1990. *La estructura argumental preferida del español hablado: los sujetos del tipo "S"*. Paper read at the IX Congreso Internacional de la ALFAL. Campinas, S.P., Brazil.
- & Elizabeth G. Weber. 1986. "A Functional Approach to Subject Word Order in Spoken Spanish". *Studies in Romance Linguistics* ed. by Osvaldo Jaeggli & Carmen Silva-Corvalán, 23-40. Dordrecht: Foris.
- Chafe, Wallace L. 1980. "The Development of Consciousness in the Production of a Narrative". *The Bear Stories: Cognitive, cultural, and linguistic aspects of narrative production* ed. by W. L. Chafe, 9-50. Norwood, N.J.: Ablex.
- Dixon, Robert M. W. 1979. "Ergativity". *Language* 55.59-138.
- Du Bois, John W. 1985. "Competing Motivations". *Iconicity in Syntax* ed. by John Haiman, 343-365. Amsterdam & Philadelphia: John Benjamins.
- , 1987. "The Discourse Basis of Ergativity". *Language* 63.805-855.
- Dutra, Rosalia. 1987. "The Hybrid *S* Category in Brazilian Portuguese: Some implications for word order". *Studies in Language* 11.163-180.
- Givón, Talmy. 1984. *Syntax: A functional-typological introduction*. Amsterdam & Philadelphia: John Benjamins.
- Sedano, Mercedes. 1990. *Hendidas y otras construcciones con ser en el habla de Caracas*. Caracas: Univ. Central de Venezuela.

PSYCH CONSTRUCTIONS AND LINKING TO CONCEPTUAL STRUCTURES*

DENIS BOUCHARD
Université du Québec à Montréal

0. *Introduction*

Following Jackendoff (1983, 1990), we assume that the correspondence rules that relate Conceptual Structures (CS) to Syntactic Structures (SS) are rather direct, so that every syntactic node will be licensed by its correspondence with a CS node.¹ This restricts the grammar in two major aspects:

- 1 a. Restriction on X-bar theory:
There can be no superfluous nonbranching nodes
- b. Restriction on movements:
The two positions involved in a movement must both be licensed by their correspondence with a CS node.

Among other things, restriction (1b) precludes movement from an object position to an 'open' subject position because the 'open' subject position is not properly licensed.² We now turn to Psych-V constructions because there is an

* Thanks are due to Nicolas Ruwet for numerous discussions on the subject, to my students at UQAM and at Paris VIII, to the audiences at Paris VIII, UMASS and at the LSRL in Ottawa. Errors and omissions are mine. I benefitted from support by the Social Sciences and Humanities Council of Canada.

¹ This is stronger than Jackendoff's proposal which assumes such a correspondence only for maximal projections. The details of our proposal are to be found in Bouchard (in preparation), with the additional assumption that there be only one CS tier, so that the structural organization of the SS is essentially reflecting the organization of the CS. Among other aspects of this more direct correspondence, the linking of argument positions to syntactic GF positions is not indexical: rather, the relative relations between arguments in CS are carried over to SS, so that linking is straightforward.

² The notion of 'open' syntactic position seems to highly deserve its scare quotation marks since it is not properly definable within X-bar theory, even more so if one adopts the idea that C-selection derives from S-selection. Independent phrase structure rules don't exist in X-bar theory: a node is not licensed in PSRs in general, but rather it is licensed in each particu-

analysis of these constructions by Belletti & Rizzi (1988; henceforth: B&R) which has received a lot of attention and which derives one class of Psych-V constructions by a promotion of an object to an 'open' subject position.

1. *The problem*

Recall what the facts are. There are two basic classes of Psych-Vs, the Fear class (*mépriser* in French) and the Frighten class (*dégoûter* in French), illustrated by (2) and (3) respectively.

2. John fears Mary/ your decision/ the storm.
3. Mary/ your decision/ the storm frightens John.

It was noted very early on (cf. Jackendoff 1972, Postal 1971 and references therein) that the selectional restrictions on the subject of Fear verbs and those on the object of Frighten verbs were essentially the same, and similarly, that the selectional restrictions on the object of Fear verbs and those on the subject of Frighten verbs were also essentially the same. Jackendoff (1972) identifies the subject of Frighten Verbs as a Theme and its object as a Goal, and Ruwet (1972) notes that we get the opposite in (2). In short, there seems to be a crossing of theta roles. This was correctly perceived as problematic since it was assumed that GF positions in Deep Structure should correspond in a fairly regular way to semantic roles (cf. Fillmore 1968, for an early proposal to this effect; for a more recent proposal, cf. Baker's 1988 Unified Theta Assignment Hypothesis). In order not to have radically different Deep Structures for sentences like (2) and (3), as their Surface Structures suggested, a number of linguists proposed early on to derive (3) from a DS close to (2) (Rosenbaum 1967, Chapin 1967, Lakoff 1970, Chomsky 1970, Postal 1971). We illustrate this by using Postal's PSYCH-MVT as in (4).

particular structure in which it appears by selectional properties of a given lexical head (including semantically justifiable functional nodes like T, D, C). A rule like S--> NP-TENSE-(M)-VP no longer exists to license an open subject position: if a node has the subject relation with another element at DS, its presence must be licensed by the semantic (selectional) relation that exists between the two. So for example, if the subject position is SPEC of IP, then an XP is licensed in that position if XP has a 'proper' semantic relation with I'. But then the notion of 'open' subject position appears contradictory: XP is licensed in DS if it can enter into a semantic relation, which presumably forces it to have features. But how then could it be unfilled, open? This is reminiscent of reasons for abandoning NP-postposing: *John* could not move to D in (i) because the object position of *by* is licensed at DS by a semantic relation established between the element in that position and *by*, which is impossible here since D cannot enter into semantic relations.

- i. John stopped Bill by D.

4. DS John frightens Mary/ your decision/ the storm ==>
 SS Mary/ your decision/ the storm frightens John.

Frighten Verbs were marked [+PSYCH-MVT], whereas Fear Verbs were [-PSYCH-MVT]. This was in the spirit of Transformational Grammar from its inception where regular correlations of this type were said to be accounted for in a more economical way if analysed transformationally, and explained.

Of course, such an analysis is no longer possible in a GB framework because it violates the theta criterion and because marking of lexical items for specific T-rules is not possible since there are no such rules anymore, all movement rules, for example, being now replaced by move alpha. But the motivation behind the analysis still holds. So B&R, assuming that theta roles must be assigned in a unified way, propose to derive (3) in a manner compatible with present assumptions as in (5).

- 5 DS [[frightens Mary/ your decision/ the storm VP] JohnVP] ==>
 SS Mary/ your decision/ the storm[[frightens t VP] JohnVP]

In (5), using their terminology, the Theme is assigned in Direct Object position, just as it is in (2), so that this theta role can be said to be uniformly assigned to the Direct Object position. Since it is impossible given current assumptions to mark Frighten Verbs as [+RAISE-OBJECT-TO-SUBJECT], they will be marked for a trigger for the movement, namely that Frighten Verbs are deficient Case assigners and hence the NP object is forced to raise. Note that this does only half of the job: the role Experiencer, although it is now assigned higher than the Theme in both (2) and (5), is not assigned uniformly to the same position: it is assigned to the subject position in (2), but to a VP adjoined position in (5).

The analysis is supposed to account for some of the peculiar properties of the Frighten class of verbs. For example, binding of an object reflexive by the subject is odd in English, bad in Romance, because an improper chain is formed with the local binding of the trace of the subject by the Experiencer reflexive.

6. a?*They frighten themselves.
 DS [[frighten they VP] themselves VP] ==>
 b**Jean se préoccupe.*

On the other hand, backward anaphora is fine because the reflexive is properly bound at DS.

7. ?Stories about herself generally please Mary.
 DS [[generally please stories about herself VP] Mary VP]

This analysis is not compatible with the general constraints that we are assuming because it crucially makes reference of an open subject position. We therefore have to look for an alternative.

2. *An alternative analysis*

While exploring the avenues of the analysis that we turn to directly, we quickly realized that it was preferable not only on these very general grounds, but also that, upon closer study and taking more data into account, the analysis of the binding facts was seriously faulted. Moreover, there are two problems with the basic assumptions of the analysis.

2.1 *Two inadequacies of the standard analysis*

The first problem is that the thematic analysis generally assumed is incorrect: there is no crossing of theta roles, so that the motivation for these analyses stemming from an attempt to assign thematic roles uniformly at Deep structure is not substantiated. Ruwet (1972) notes that the selection restrictions on the subject of *dégoûter* and those on the object of *mépriser* are not quite the same. It seems to be just about impossible to find true minimal pairs, as Ruwet (1972) amply illustrated by showing that minor differences exist between just about every potential minimal pair. We cannot repeat all his examples here since the material would cover several dozens of pages and is already easily available in any event, but it should be pointed out that, even for the verbs used as typical examples of Class 1 and Class 2 Psych verbs in French and English, sentences can be found where the selectional restrictions are not exactly the same and crossing does not hold ((8.a-b) from Ruwet 1972):

8. a. *La couleur vertel/que Paul ait pu dire une chose pareille dégoûte Pierre.*
 b. **Pierre méprise la couleur vertel/que Paul ait pu dire une chose pareille.*
9. a. The brown spots on Ronald's skin frighten Nancy.
 b. #Nancy fears the brown spots on Ronald's skin.

Kenny (1963) noted that there are differences in interpretation between what appear to be minimal pairs:³

10. a. Bill was very angry at the article in the Times (object= TARGET)
- b. The article in the Times angered Bill greatly (subject= CAUSE)

In (a), *the article*: is necessarily evaluated by *Bill*; but in (b), it need only be causally connected to the emotion described by the predicate and borne by the Experiencer.

The second problem for analyses that treat Frighten verbs as special is that Psych constructions with similar syntactic and aspectual properties are very productive, contrary to what the analysis predicts, with verbs switching fairly freely from a 'normal' use to a Psych use, with all the syntactic and aspectual properties of Frighten verbs in the latter case. In his discussion of Postal's transformational account, Ruwet (1972) noted that there is a third class of Psych Verbs that have to be taken into account: it is the *Frappé* (strike) verbs as in (11).

11. a. *Paul a frappé/ébloui/empoisonné Marie par son discours.*
- b. Paul strikes Mary as intelligent.

Ruwet shows that this class is very productive in French. He lists dozens of verbs of the class and adds that it seems that almost any causative action verb can become a Psych verb. We believe that productivity is even more general: Psych verbs are but a subcase of a very productive class of Psych constructions. For a vast class of verbs, if one of their argument positions is filled by a *psy-chose** (a psychological object, found only in mental space, like an emotion), then the construction is Psych. This is possible with just about any Change-of-state verb with two or more arguments. Psych verbs are always Psych because the *psy-chose* argument is incorporated in the verb (fright, fear, disgust, love, hate, etc.): but that is all there is special about them. As far as most other properties are concerned, Psych verbs are like ordinary verbs. Psych verbs have no special Deep structure. They have no special thematic structure or thematic roles.

³ Cf. Pesetsky (1988) who assigns different theta roles to *the article* in these two sentences to account for this difference

* [The author has coined this hybrid compound, which plays on the French *chose* "thing" (and has nothing to do with *psychosis*), to express his particular argument. — Ed.]

To get an idea of the extreme productivity of the construction, consider the following sets of examples (some of which are due to Ruwet 1972). The examples in (12-16) are instances of Psych constructions without incorporation of the psy-chose argument, whereas in examples (17-19), the psy-chose argument is incorporated in the verb, so that the verb is always Psych.

12. THEME psy-chose without incorporation:
 - a. *Jean fait peur à Marie* [Exp. = Goal]
 - b. *Paul a peur de Marie* [Exp. = Place]
 - c. *Jean donne du soucis à Marie* [Exp. = Goal]
 - d. *Il y a en Pierre un profond dégoût/mépris de l'argent.* [Exp. = Place]
 - e. *Paul voue une haine féroce à Virginie.* [Exp. = Agent and Source]
 - f. *Ce désir de vengeance me vient de Paul.* [Exp. = Goal]
13. PLACE psy-chose without incorporation:
 - a. *Cela a mis Marie en colère* [Exp. = Theme]
 - b. *Paul est tombé de désespoir en désespoir.* [Exp. = Theme]
14. GOAL psy-chose without incorporation:
 - a. *Paul a poussé Marie à la haine/à la colère/au désespoir* [Exp. = Theme]
 - b. *Le vieux moine a poussé Paul jusqu'au dégoût de soi.* [Exp. = Theme]
 - c. *Paul a donné libre cours à sa colère* [Exp. = Agent]
15. SOURCE psy-chose without incorporation:
 - a. *Marie n'est pas revenue de l'horreur de ce spectacle.* [Exp. = Theme]
16. INSTIGATOR ('Agent') psy-chose without incorporation (whole class of action verbs like *frapper* (to strike)):
 - a. *Jean a frappé/ébloui/empoisonné Marie par son discours.* [Exp.=Theme]
 - b. *Ces nuages me disent qu'il va pleuvoir.* [Exp. = Goal]
 - c. *Une peur bleue le gagna/ s'empara de lui* [Exp. = Theme]
 - d. *Paul a fait tomber Marie à la renverse par son intelligence.* [Exp. = Theme]⁴
 - e. *La rage l'étouffe.* [EXP= Theme]
 - f. *La jalousie le mine.* [EXP= Theme]
17. THEME psy-chose with incorporation (numerous):
 - a. *Pierre méprise l'argent/les idées de Paul.* [Exp. =Agent (and Source/Goal?)]
 - b. *L'argent /cette idée dégoûte Pierre .* [Exp. = Place]

⁴ Note that the psy-chose *Paul* and the Experiencer *Marie* are not coarguments in (16.d).

- c. *Merckx a profité/bénéficié de l'abandon d'Ocana.* [Exp. = Agent+Goal]
- d. *L'abandon d'Ocana a profité/bénéficié à Merckx* [EXP= Goal/Place]
- e. *Jean aime Marie.* [Exp. = Source (or Goal?)]
- 18. PLACE psy-chose with incorporation:
 - a. *Cet article/ Marie a enragé Paul.* [Exp= Theme]
- 19. GOAL and SOURCE psy-chose with incorporation: these seem to be impossible.⁵

This great productivity of Psych constructions raises essentially the same problem for B&R as the one observed by Ruwet (1972) with respect to Postal's (1971) analysis: these analyses require a proliferation of homonymous verbs in the language, so that language seems to suffer from almost generalized homonymy. For example, a verb like *frapper* would have to have two lexical entries, depending on whether it is used 'normally' or with a Psych interpretation. That means that in an analysis that attributes the peculiar syntactic properties of Psych verbs to a particular Deep structure, then the two sentences in (20) and (21) would have radically different Deep structures, with the two arguments occupying different syntactic positions.

- 20 *Marie frappe Paul (avec un marteau)*
DS [IP Marie [VP frappe Paul]]
- 21 *Marie frappe Paul (par son intelligence)*
DS [IP __ [VP [VP frappe Marie] Paul]]

This is needed because the *Frapper* class of Psych verbs do exhibit the peculiar properties for binding of reflexives, among others.

- 22. a. **Jean se frappe par son intelligence.*
b? *John strikes himself as pompous.

Moreover, as Ruwet (1972:232) noted, the homonymy here is not at all perceived as similar to the one in *voler* (to fly) and *voler* (to rob). The two uses of *frapper* and the like are not pure coincidence: speakers perceive a systematic semantic correspondence between the two uses.⁶

⁵ This is probably due to a general condition against incorporation of these arguments which would account for the contrast in possible verbs like *butter* vs **bread*. Cf. Jackendoff (1983) and references therein.

⁶ Ruwet (1972:231) expresses this intuition very clearly in his discussion of the *Frapper* class of verbs:

If we consider additional data about Psych constructions that cover more than the two limited classes of Psych verbs with incorporation like Fear verbs and Frighten verbs, we are forced to conclude that an account of the peculiar syntactic phenomena found in Psych constructions which would be based on special theta roles or special Deep structures is not tenable.

2.2 *What is a Psych construction?*

So what is a Psych construction and how do we account for the peculiar properties of some of those constructions? As the data in (12-18) show, if the theta roles assigned in Psych constructions are constant, something like EXPERIENCER and TRIGGER⁷, a systematic alignment of these theta roles with GFs is impossible: the EXPERIENCER and the TRIGGER can be in just about any syntactic position. A similar attempt to align the EXPERIENCER and TRIGGER roles with the 'old' thematic roles like AGENT, THEME, etc. in the cases where the constructions can alternate between a Psych use and a normal use will also fail, as a cursory look at the data will quickly indicate. Moreover, if we also consider the psy-chose argument, which is the emotion that the TRIGGER induces in the EXPERIENCER, we see that this element can also appear in any syntactic position and with any 'old' theta role.

Psych constructions can exhibit a variety of thematic relations and they are indistinguishable from other constructions in this respect. What sets Psych constructions apart from the others is that one of their arguments is a psy-chose, and that there is another argument which is 'affected' by the psy-chose. We assume that, when a verb expresses a physical contact between two objects, that contact induces a change of state in one or the other of these objects, hence that it is affected. In order to affect it, the psy-chose is therefore somehow put in contact with an entity capable of hosting the emotion or feeling that

Il est évident qu'on a affaire ici à un phénomène productif, très général, et qu'il y a un rapport systématique entre les classes de verbes A et B; une grammaire du français qui n'en tiendrait pas compte serait inadéquate. De plus, dans la majorité des cas, tout sujet parlant natif du français a l'intuition d'un rapport sémantique entre les deux verbes homonymes. En gros, et d'une manière impressionniste, ce rapport peut s'exprimer de la manière suivante. Dans les exemples (a) comme dans les exemples (b), les verbes décrivent un processus dont le NP sujet désigne la cause, et ce processus affecte, ou a un effet, d'une manière ou d'une autre, sur l'être ou l'objet désigné par le NP objet. La différence est que, dans les exemples (a), l'effet en question est d'ordre purement physique, alors que, dans les exemples (b), il s'agit d'un effet psychologique, mental.

⁷ The TRIGGER is the thing that triggers the emotion. We use TRIGGER instead of THEME because that element does not have the properties of a Theme as defined in Gruber (1965), where Theme is the element being situated on a Path or in a Place.

the psy-chose refers to (we will refer to this entity as a human from now on, although this may certainly vary from one society or individual to another).

The difference between a Psych construction and a 'normal' construction is that, the psy-chose being a mental object, the event where the psy-chose establishes contact with a human takes place at the level of mental space rather than at the level of physical space. The fact that elements expressing basic spatial relations extend to other domains in this way is well known: for example, Jackendoff (1983) notes that functions in other fields like time, possession, temperature use essentially spatial functions.⁸

If we assume that a psy-chose must affect a human and that it does this by making contact with it in mental space, then there are two possible spatial relations that could underlie a Psych construction:

— PATH: the psy-chose and the human could be at the same point on a PATH: the initial point if one or the other is Source, the final point if one or the other is Goal.⁹

— PLACE: the psy-chose and the human could be at the same Place.

We thus expect two kinds of Psych constructions, those based on a Path and those based on a Place. Moreover, a further distinction can be made on the basis of incorporation: if the psy-chose argument is incorporated, then we get an essentially Psych V, in the sense that it will always be in a Psych construction because of the very nature of the incorporated argument (this is the class usually referred to as Psych verbs).¹⁰ Another distinction could also be made: since the psy-chose and the human have to make contact, one of them has to be the Theme in Gruber's sense (the element being situated spatially) whereas the other will be a point on a Path or a Place. But which element is what need not always be the same. Thus, the psy-chose *peur* seems to be the Theme and the human *Marie* the Goal or Place in (23), but we seem to get the opposite in (24), with the psy-chose *colère* being the Place and the human *Marie* being the Theme.

23. *Pierre fait peur à Marie.* [P CAUSE PEUR GO TO/BE AT MARIE]

24. *Pierre met Marie en colère.* [P CAUSE MARIE BE AT COLERE]

⁸ See also Dowty (1979), Fauconnier (1984), among many others, on logico-mental space.

⁹ It is even possible for the contact to be made at an intermediate point, like the verb *effleurer* (to graze) which can be used in a Psych construction as in (i).

i. *Quelques soupçons avaient effleuré Jean.* "Some misgivings had crossed Jean's mind".

¹⁰ Roger Higgins (p.c.) points out that the hypothesis that so-called Psych verbs are instances of incorporation of a psy-chose argument finds support in the fact that these verbs are essentially denominals in Indo-European.

With these three factors — incorporation or not, Place or Path, psy-chose or human as Theme — we get eight potential classes of Psych constructions.

25. potential classes of Psych Constructions

+incorporated	PLACE	Theme= psy-chose	Psych 1
		Theme= human	Psych 2
	PATH	Theme= psy-chose	Psych 3
		Theme= human	Psych 4
-incorporated	PLACE	Theme= psy-chose	Psych 5
		Theme= human	Psych 6
	PATH	Theme= psy-chose	Psych 7
		Theme= human	Psych 8

The fact that there can be so many different types of Psych constructions and that they should be so productive comes as no surprise if one assumes that, as far as their thematic organization is concerned, they are basically like any 'normal' physical spatial relation.

We can now ask ourselves how this relates to our previous observations where we had four pre-theoretical classes of Psych constructions as shown schematically in (26).

26. a. Class 1: Fear, *Mépriser* EXPERIENCER V TRIGGER
 b. Class 2: Frighten, *dégoûter* TRIGGER V EXPERIENCER
 c. Class 3: Strike, *Frapper* TRIGGER V EXPERIENCER
 d. Class 4: all other nonincorporated constructions

The easiest turns out to be Class 4 since it is the most transparent: we see that there is no incorporation, the psy-chose is the argument with lexical properties that identify it as such, like anger, fear, joy, etc., and the surface prepositions help a lot in determining what spatial relation is at hand — Path or Place — and in what direction the relation is established. For example, in (24) above, we have a Path with a human Theme and the endpoint of the Path is the psy-chose *colère*.

Class 1 verbs are the result of incorporating a psy-chose Theme and are based on a Path, with the direct object as one endpoint of the Path.

Class 2 verbs are the result of incorporating a psy-chose and are based on a Place, with the psy-chose and the direct object as the two elements in the Place relation and the subject as a CAUSE.

Class 3 is not a case of incorporation of a psy-chose since it is not always Psych in nature unlike classes 1 and 2; the human argument (the experiencer) is

the direct object. But where is the psy-chose? There is no element in sentences like (27) which has the inherent lexical properties of a psy-chose like anger, fear or disgust.

27. a. Mary strikes me as intelligent.
b. *Marie me frappe (par son intelligence).*

It is the external argument of the predicate CAUSE in the CS of these verbs which is the psy-chose: an Instigator can be interpreted as either an individual or the properties of an individual. When interpreted as properties, the Instigator is then a psy-chose since it is in the realm of mental space (Grimshaw 1990, Partee & Rooth 1983, Kenny 1963).¹¹

We assume that an action verb like *frapper* has a CS representation roughly as in (28).

28. a. *Jean a frappé Paul.*
b. CS: x CAUSE [y [PATH GO A z]], where x=*Jean*, y=x
and z=*Paul*.¹²

¹¹ Partee & Rooth (1983) are concerned with problems of type ambiguity in conjunction. The problem is that different types of elements can be conjoined and that this is a problem for their interpretation. For example, consider a conjunction of verbs. If two extensional verbs are conjoined, then the interpretation is that the two verbs have the same direct object term phrase, so that there is just one same fish in (i) and just the same three women in (ii).

- i. John caught and ate a fish.
ii. John hugged and kissed three women

If two intensional verbs are conjoined, they are interpreted as having different direct object term phrases (cf. (iii)) and the same holds if an intensional and an extensional verb are conjoined (cf. (iv)).

- iii. John wants and needs two secretaries.
iv. John needed and bought a new coat.

Conjoining NPs creates a similar problem. Thus (v) is three way ambiguous:

- v. The department is looking for a phonologist or a phonetician
a de re: specific person, who is either a phonologist or a phonetician
b de dicto: they would be satisfied if they found one or another
c second de dicto: they have a particular kind of person in mind, but I don't know

which kind (equivalent to The department is looking for a phonologist or looking for a phonetician). — Basing their work on Montague (1974) and Keenan & Faltz (1978) among others, where an individual is a set of term denotations, Partee & Rooth propose that an extensional phrase can always be 'lifted' to a higher type intensional phrase. To put it in slightly different terms, a NP can be interpreted as an individual or as the properties of an individual.

¹² The Source is not in the CS representation here because it is obtained by inference.