

5-1-2016

VERBS OF SPEECH AND COGNITION: STRUCTURAL PATTERNS IN NEW MEXICAN SPANISH

Víctor A. Valdivia Ruiz

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

Recommended Citation

Valdivia Ruiz, Víctor A.. "VERBS OF SPEECH AND COGNITION: STRUCTURAL PATTERNS IN NEW MEXICAN SPANISH." (2016). https://digitalrepository.unm.edu/ling_etds/38

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

Víctor Alfredo Valdivia Ruiz

Candidate

Linguistics

Department

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

Approved by the Dissertation Committee:

Melissa Axelrod, PhD, Chairperson

Naomi Shin, PhD

Damián Vergara Wilson, PhD

Catherine Travis, PhD

José Esteban Hernández, PhD

**VERBS OF SPEECH AND COGNITION:
STRUCTURAL PATTERNS IN NEW MEXICAN SPANISH**

by

VÍCTOR ALFREDO VALDIVIA RUIZ

B.A., Hispanic Letters, University of Guadalajara, 2004
M.A. Applied Linguistics, University of Guadalajara, 2008

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

**Doctor of Philosophy
Linguistics**

The University of New Mexico
Albuquerque, New Mexico

May 2016

Dedication

To Jana, for her love and encouragement through the years;

to my parents, Víctor and Patricia,

to my sisters, Adriana y Patricia,

to my brothers, Óscar and Uriel,

and to Sandy,

for all their support.

To Román,

for the joy of being an uncle.

Acknowledgements

I would like to express my deepest gratitude to my advisor, Melissa Axelrod; without her guidance, encouragement, and generosity, I would never have been able to finish my dissertation. I would also like to thank Catherine Travis for guiding my research for the past years. I am very grateful to both of them for the comments and suggestions in the several reviews of this work. My gratitude is also extended to Damián Vergara Wilson, José Esteban Hernández and Naomi Shin; it has been an honor to have you in my dissertation committee.

Special thanks goes to the Department of Spanish and Portuguese, in particular to Dr. María Dolores Gonzales for first inviting me to teach in the Sabine Ulibarrí Spanish as a Heritage Language program, and to Dr. Carmen Julia Holguín Chaparro for her support while teaching upper division courses. Finally, I would to thank Fabiola Parra, Mayra Flores-Soto, Martha Hurd, and Kate Merrill; their friendly words were always an important part of the day.

**VERBS OF SPEECH AND COGNITION:
STRUCTURAL PATTERNS IN NEW MEXICAN SPANISH**

by

VÍCTOR ALFREDO VALDIVIA RUIZ

B.A., Hispanic Letters, University of Guadalajara, 2004
M.A. Applied Linguistics, University of Guadalajara, 2008
PhD, Linguistics, University of New Mexico, 2016

Abstract

From a usage-based approach, this dissertation explores the relationship between structural variability and discourse subjectivity in a series of interviews from the *New Mexico and Southern Colorado Spanish Survey* (Bills and Vigil 2008). Focusing on the use of cognitive and speech verbs, the analysis will demonstrate that a) structural patterns arising from said interaction form a continuum ranging from schematic to fixed constructions; b) the degree of morpho-syntactic variability of these constructions correlates to the speaker's stance towards the information being communicated; and c) in actual interaction, speakers frequently introduce subjective nuances, even if the topic of the discourse may be associated with mere transmission of propositional information. To address the aforementioned goals, a set of features is proposed to determine whether

frequent subject-tense combinations are becoming prefabs or formulaic expressions, or whether they are schematic constructions merely transmitting propositional information. Previous studies have explored structural patterns of subjectivity in conversation, either examining a wide range of verbal classes (Scheibman 2002) or focusing on specific verbs and their relation to other items in the semantic class (Tao 2001; Weber and Bentivoglio 1991), as well as the subjectification of specific verbs (Company Company 2006a; Company Company 2006b; Travis 2005; Travis 2006a; Travis 2006b). This study contributes to the understanding of emergent grammar, subjectivity, and subjectification across languages by addressing the patterning of frequently used semantic classes and verbs in a context of language and sociocultural contact. The study aims to contribute as well to the revalorization of New Mexican Spanish by demonstrating that structural patterns emerging from the discourse of its speakers are as systematic as those arising from the discourse of speakers of more prestigious dialects of the language.

TABLE OF CONTENTS

List of Tables	x
Chapter 1. Emergent Patterns and New Mexican Spanish	1
1.1. Introduction to the Study	1
1.2. Spoken Language and Functional-Cognitive Linguistics	2
1.3. Grammar, Spoken Language and Conversation	5
1.4. New Mexican Spanish: On Overview	6
1.4.1. Retention of Archaisms	8
1.4.2. Independent Developments	9
1.4.3. Influence of English	9
1.4.4. Popular Mexican Spanish and Standard Spanish Influence	10
1.5. Organization of the Study	11
Chapter 2. Theoretical Framework	13
2.1. Organization of the Chapter	13
2.2. Usage-Based Theory	13
2.2.1. Repetition, Chunking and Knowledge of Usage	14
2.2.2. Categorization	15
2.2.3. Exemplars and Networks	16
2.3. Constructions and Other Multiword Units	19
2.3.1. Constructions	19
2.3.2. The Lexicon-Syntax Continuum	21
2.3.3. Idioms	25
2.4. Constructions, Grammar and Conversation	33
2.5. Usage-Based Construction Grammar In Grammaticalization	34
2.5.1. New Constructions Emerge From Already Established Ones	35
2.5.2. Old and Emerging Constructions Compete	36
2.5.3. Grammaticalization of Constructions	36
2.6. The Role of Repetition	38
2.6.1. Type Frequency	38
2.6.2. Token Frequency	40
2.7. Language and Subjectivity	42
2.7.1. Subjectivity and Subjectification	46
2.8. Discourse, Structural Patterns and Subjectivity	50
2.9. Summary	58
Chapter 3. Data and Methodology	59
3.1. Data	59
3.2. Methodology	63
3.3. Coding	64
3.3.1. Subject's Form (Person and Number)	66
3.3.2. Subject's Realization	71
3.3.3. Subject's Position	72
3.3.4. Subject's Referentiality	73

3.3.5.	Polarity	78
3.3.6.	Type of Sentence (Purpose)	79
3.3.7.	Verb's Function and Form	79
3.3.8.	Tense, Aspect and Mood (TAM)	82
3.3.9.	Complement Type	85
3.3.10.	Complement Realization	89
3.3.11.	Connective	93
3.3.12.	Word Order	94
3.3.13.	Indirect Object	95
3.3.14.	Prosodic Distribution of the Complement Regarding the Verb	96
3.3.15.	Prosodic Distribution of the Complement	98
3.3.16.	Transition Continuity of the IU Containing the Verb	99
3.3.17.	Change of Turn	100
3.3.18.	Function	101
Chapter 4.	General Patterns: Verbs of Cognition and Verbs of Speech	103
4.1.	Type and Token Frequency	103
4.2.	Subject	105
4.3.	Tense	106
4.4.	Class, Subject and Tense	108
4.4.1.	Cognitive Verbs	109
4.4.2.	Speech Verbs	110
4.5.	Subject Referentiality	111
4.6.	Type of Complement	112
4.7.	Word Order: Verb-Complement	115
4.8.	Polarity	116
4.9.	Summary	117
Chapter 5.	Cognitive Verbs	120
5.1.	Cognitive Verbs: General Patterns	120
5.1.1.	Cognitive Verbs: Predicate Type	121
5.1.2.	Subject and Tense	121
5.1.3.	Subject and Predicate Type	123
5.2.	Cognitive Verbs: Local Patterns	124
5.2.1.	1sg-Cognitive Constructions	125
5.2.1.1.	(yo) <i>no sé</i> 'I don't know'	125
5.2.1.2.	(yo) <i>creo</i> 'I think / I believe'	137
5.2.1.3.	(yo) <i>me acuerdo</i> & (yo) <i>no me acuerdo</i> 'I (don't) remember'	143
5.2.2.	3sg Cognitive Constructions	150
5.2.2.1.	<i>Quién sabe</i> 'Who knows'	151
5.2.2.2.	<i>X sabe</i> 'X knows'	155
5.2.3.	Other Prefabs and Collocations	156
5.2.3.1.	<i>SABER</i> 'to know how' + Infinitive	157
5.2.3.2.	<i>Se ME hace</i> 'It seems to ME'	159
5.2.3.3.	(<i>Tú</i>) <i>sabes</i> & (<i>usted</i>) <i>sabe</i> 'You know'	161

5.2.3.4. <i>Entiendo</i> 'I understand', <i>¿entiendes?</i> 'Do you understand?' & (X) <i>entiende</i> 'X understands'	165
5.3 Summary	170
Chapter 6. Speech Verbs	172
6.1. Speech Verbs: General Patterns	172
6.1.1. Speech Verbs: Predicate Type	173
6.1.2. Subject and Tense	177
6.2. Speech Verbs: Local Patterns.....	184
6.2.1. 1sg-Speech Constructions	185
6.2.1.1. (yo) (<i>le</i>) <i>digo</i> 'I say / I tell him-her'	185
6.2.1.2. <i>Por eso digo</i> 'that's why I say (so)'	194
6.2.1.3. <i>Ya te digo</i> 'aprox. that's what happened'	196
6.2.1.4. <i>Como (TE) digo</i> 'As I say / tell (YOU)'	197
6.2.2. 3sg-Speech Constructions.....	201
6.2.2.1. (X) <i>dijo</i> 'X said'	202
6.2.2.2. (X) <i>dijo que</i> 'X said that' + [clause]	207
6.2.2.3. (X) (<i>LE</i>) <i>dice</i> 'X says (to him-her)'	210
6.2.3. 3pl-Speech Constructions	214
6.2.3.1. <i>DECIRLE(S)</i> X 'to call it-them X'	214
6.2.3.2. <i>Dicen (que)</i> 'it is said' & <i>decían que</i> 'it was said'	217
6.2.3.3. <i>HABLAR [ESPAÑOL]</i> 'to speak [Spanish] & <i>HABLAR en [ESPAÑOL]</i> 'to speak in Spanish'	219
6.3. Summary	228
Chapter 7. Conclusions.....	230
Appendix A. DuBois et al (1993). List of Symbols and their Meaning	235
References	236

LIST OF TABLES

Table 1. The syntax-lexicon continuum.....	21
Table 2. Consultants' Demographic Information	60
Table 3. Verbal Class and Frequency	103
Table 4. Utterances by verbal class and person	105
Table 5. Utterances by verbal class and tense.....	107
Table 6. Distribution according to verbal class, tense and subject	108
Table 7. Distribution according to verb class and subject referentiality.....	111
Table 8. Verb class and type of complement.....	113
Table 9. Verb-complement order.....	116
Table 10. Distribution according to type of utterance	116
Table 11. Cognitive verbs by type of predicate	121
Table 12. Cognitive Verbs by Subject and Tense.....	122
Table 13. Cognitive Utterances: distribution by subject and type of nucleus	124
Table 14. Verbs of Speech by type of Predicate.....	173
Table 15. Verbs of Speech, distribution for Subject and Tense.....	178
Table 16. Summary of Constructions with Cognitive and Speech Verbs.....	231

Chapter 1. Emergent Patterns and New Mexican Spanish

1.1. Introduction to the Study

In this dissertation I examine the use of cognitive and speech verbs in an oral corpus of New Mexican Spanish, a variety of Spanish often stigmatized due to the presence of what some linguists have described as “archaic” features (Alvar 2000); not to mention syntactic and semantic phenomena resulting from its intense contact with English since early the 19th century. I am looking at these verbal classes in particular in order to explore the relationship between structural variability and discourse subjectivity. Through this examination, I will demonstrate that a) patterns arising from discourse form a continuum ranging from schematic to fixed constructions; b) the degree of morpho-syntactic variability correlates to the speaker’s stance towards the information being communicated; and c) in actual interaction, speakers frequently introduce subjective nuances, even if the topic of the discourse may be associated with the mere transmission of propositional information. Additionally, I will propose a set of features to determine whether frequent subject-tense combinations are losing morpho-syntactic variability and thus developing prefab, idiom or discourse marker status. Previous studies have explored how different verbs work in actual discourse by focusing, for example, on how specific items pattern in relation to the semantic class they belong to (Tao 2001; Weber and Bentivoglio 1991) or on the grammaticalization of specific verbs (Company 2006a; Company 2006b; Travis 2005; Travis 2006a; Travis 2006b). This study adds to that literature by addressing the patterning of frequently used semantic classes and verbs in a context of language and sociocultural contact. Thus, by demonstrating that structural patterns arising from the discourse of speakers experiencing such situations are as

systematic as those emerging from the discourse of speakers of so-called monolingual dialects, this study aims to contribute not only to the body of research addressing the interplay between language use and grammar, but also to the revalorization of New Mexican Spanish and Spanish in the United States.

1.2. Spoken Language and Functional-Cognitive Linguistics

Paying just a little bit of attention to most forms of spoken language, it is easy to notice that speakers often utter “uncompleted” sentences, and produce grammatical “mistakes” either because they need to reformulate an idea already being expressed, because they assume the interlocutor will be able to understand the message given the existence of shared knowledge, or other interactional factors. On the contrary, in written language, particularly in formal and highly planned genres (e.g. academic papers, official speeches and letters, etc.) speakers produce organized and well-structured sentences because there are no interactional elements interfering with the flow of the discourse. Based on such general characteristics, until recent decades, linguistics ignored spontaneous speech, and considered written genre an idealized version of language. As a direct consequence of such categorical distinction, traditional Western linguistics conceptualized grammar as a logical, coherent, and steady set of rules that speakers apply when producing or receiving a statement. Nevertheless, even in formal written language it is possible to find structures that, to some degree, deviate from what is thought as a grammatical model. Examples of these structures in English are ‘There are my shoes’ and ‘Here are my shoes’, in which the nominal phrase following the verb must be considered its subject based on the formal criterion of number agreement (Tomasello 2003a:8), a

situation that contradicts the traditional classification of English as a Subject-Verb language.

The prevalence of written language has influenced not just the conceptualization of what grammar is, but also the methodology followed for decades to study ‘unfamiliar languages’, i.e., those without a corpus of written texts. After eliciting, transcribing, and glossing discourse from native speakers, linguists would look for tokens exemplifying particular grammatical types. The role of discourse was merely to provide linguistic units to substantiate preconceived claims about grammar; utterances that did not fit into grammatical types were discarded as irrelevant (Du Bois 2003). Thus, even in cases in which spoken language was used as the main source of data, only those utterances proving the existence of preconceived and ideal grammatical categories would be considered as object of study.

Contrary to such traditional perspectives, Functional-Cognitive Linguistics considers language to be a cognitive ability that evolves from the need to communicate with others; coherently, it takes spoken language as its primary object of study. Thus, the fact that utterances from spoken language, as well as sentences as those referred to above, do not comply with a perfect model of grammar does not mean they are ungrammatical nor should they be ignored; they simply reveal that language is constantly adapted to overcome constraints imposed by interaction, and, more importantly, that speech “has properties of its own that are different, in some cases very different, from the intuitive model of language that literate, educated people carry around in their heads” (Tomasello 2003a:5). For instance, scholars have demonstrated that, when expressing an event, speakers rarely use full nominal phrases to designate all of the participants (e.g. *Carmen*

got a job in China); instead, they introduce a referent in one unit of speech, and then predicate something about it in another unit, as in ‘Do you remember Carmen? She got a job in China’ (Chafe 1994; Chafe 1998). Another observed phenomenon is that verbs formally classified as transitive ones often occur without any complement in speech (Ford, Fox, and Thompson 2003; Thompson and Hopper 2001).

The increasing interest on spoken language is evident in the diversity of scholars that have undertaken the systematic study of spontaneous (as much as possible) speech not only to examine and describe how speakers use particular linguistic units, e.g., morphemes or constructions (Ashby 1988; Bentivoglio 1992; Bentivoglio 1993), but also to address issues such as grammar and interaction (Sacks, Schegloff, and Jefferson 1974), the correlation between language, cognition and consciousness (Chafe 1979; Chafe 1980; Chafe 1987; Chafe 1993; Chafe 1994), cognitive processes (Clark 1992; Levelt 1989) and conversational organization (Goodwin 1979; Goodwin 1980; Goodwin 1981; Schiffrin 1988)¹. In addition to changing the idea of what grammar is and how it works, studies within the functional-cognitive approach have provided evidence challenging the assumed propositional nature of language; researchers examining spontaneous speech often observe that exchange of information is not always, nor primarily, the focus of interaction; instead, speakers use a good amount of the discourse to talk about their feelings and impressions, as well as their attitude and stance regarding the topic of the conversation (Scheibman 2001; Scheibman 2002; Thompson and Hopper 2001; Verhagen 2005). The present dissertation then, seeks to contribute to these corpora of

¹ For a further discussion, see Ono and Thompson 1995.

research by examining and comparing structural patterns and constructions involving two semantic verbal classes.

1.3. Grammar, Spoken Language and Conversation

Even though the label Functional-Cognitive linguistics includes theoretical approaches that may differ from each other in their particular methodologies, and even in their assumptions about grammar and spoken language, they all include, in one way or another, Langacker's conception of grammar as that "constantly evolving set of cognitive routines that are shaped, maintained, and modified by language use" (1987:57), and Hopper's (1987; 1998) proposal that grammar emerges from communication rather than being an a priori phenomenon, and that, consequently, linguistic units are always provisional and dependent on the uses and contexts in which they had previously occurred.

"The notion of Emergent Grammar is meant to suggest that structure, or regularity, comes out of discourse and is shaped by discourse in an ongoing process. Grammar is, in this view, simply the name for certain categories of observed repetitions in discourse. It is hence not to be understood as a prerequisite for discourse, a prior possession attributable in identical form to both speaker and hearer. Its forms are not fixed templates but emerge out of face-to-face interaction in ways that reflect the individual speakers' past experience of these forms, and their assessment of the present context, including specially their interlocutors, whose experiences and assessments may be quite different." (Hopper 1998:156)

Du Bois (2003) recognizes that there are differences between grammar and discourse causing them to be perceived as mutually exclusive; while the former is considered to general and operating based on restrictive rules, the latter is seen as what speaker do in actual interaction, and thus largely dependent on what speakers are talking about and the way they carry on the conversation. Despite differences, however, Du Bois argues that such dichotomies are more apparent than real; the truth is that grammar and discourse are strongly intertwined. Hence, he makes the following assumptions: "[f]irst,

speakers exploits available grammatical structure to realize their goals in speaking. Second, the aggregate sum of what speakers do in discourse exhibits recurrent patterning beyond what is predicted by rules of grammar. Third, grammatical structure tends to evolve along lines laid down by discourse pattern: Grammars code best what speakers do most” (Du Bois 2003:49). As it is clear from such assumptions, Du Bois agrees with other scholars in that conversation, especially spontaneous one, best embodies the real nature of language because, by being interactional and highly contextualized, it reflects both the speaker’s and the hearer’s cognitive models.

In the same line of thought, Ono and Thompson (1995) argue that spoken language is fundamental for a comprehensive theory of language; furthermore, they consider that, even if existing research has contributed abundantly to such a goal, “we have yet to explore the general question of what conversation can tell us about how syntax works in general, about what syntax *is*, about what a *theory* of syntax must look like” (1995:214). Naturally, addressing such questions is even more urgent in situations of language contact given the implications that using two or more languages in daily basis may have for understanding linguistic phenomena across languages.

1.4. New Mexican Spanish: an Overview

“*¡El español de Nuevo México!* New Mexican Spanish! For lovers of language, lovers of the Spanish language in particular, the name evokes emotions of intrigue, mystery and romance.” (Bills and Vigil 2008:1)

Broadly defined, New Mexican Spanish is the variety of said language spoken in Northern New Mexico and Southern Colorado (Cobos 1983; Alvar 2000; Bills and Vigil 1999; Bills and Vigil 2008; Vigil and Bills 2004; Sanz and Villa 2011). Also referred to as Traditional Spanish, it developed its own features that distinguish it from other dialects of the language spoken in the Southwest (e.g. Border Spanish, Texas Spanish, etc.) and

elsewhere. Two frequent myths about New Mexican Spanish are the claim that is the Spanish of sixteen-century Spain, and that it is a homogeneous dialect. The first idea arises from the use of so-called archaisms (e.g. *cuasi* ‘almost’, *vide* ‘I saw’²) as well as from a strong sense of cultural and historical identification with Spain among many New Mexicans. Even though it is true that Spanish first arrived to with the expedition of Spanish conquistador Francisco Vázquez de Coronado in 1540; but it is true as well that Spanish language and culture did not have a real impact in the area until 1598, when Juan de Oñate, born in New Spain from Basque-Spanish colonialist, and some five hundred persons, many of them born in New Spain or elsewhere in the Americas, established the first permanent Spanish colony north of the present Santa Fe. From that original settlement, colonists dispersed along the upper Río Grande until the Pueblo Revolt of 1680 force settlers and colonizers to escape to what is now the El Paso-Juárez area. By the end of that century, however, the colony had been reestablished with a Spanish-speaking population consisting of prior colonists and new immigrants, totaling a thousand individuals. Not only were most settlers not of European Spanish, but their Spanish was already a hybrid of the varied dialects from all over the Iberian Peninsula (see (Lipski 1994). The second idea relates to the fact that Spanish speakers of the Southwest had few chances to be in contact with other Spanish speaking communities for about three centuries, and it reinforces the myth of the Castilian origin of New Mexican Spanish is often reinforced by implying it did not experience as many changes as other dialects did. The isolation, however, was never absolute, but there was a continuous stream of

² In Modern Spanish, the normative forms for these words are *casi* and *vi*. It is worth mentioning, though, that many of the archaisms found in New Mexican Spanish occur in some sociolects in other Spanish-speaking communities.

immigrants mainly from Mexico, even if weak depending on the time as well as economic and political factors. Thus, even if early newcomers might have assimilated, culturally and linguistically, into their new communities, the ties with other dialects of Spanish have been always present, and have only increased due to new waves of immigrants in the last century.

The reality of New Mexican Spanish, then, is more complex than what myths seek to make us believe; its distinctive features result from the confluence of historical and social factors, and they can be described in terms of four main characteristics: a) retention of archaisms, b) independent developments, c) influence of English, d) contact with Mexican Spanish and standard Spanish (Bills and Vigil 2008:30–31).

1.4.1. Retention of Archaisms

Due to the relatively isolation of the area for various centuries (see above), many archaisms are still alive in New Mexican Spanish. Such archaisms are mainly phonological (e.g. *mesmo* for *mismo* ‘same, self’), morphological (e.g. *trujieron* for *trajeron* ‘they brought’, etc.), and correspond not only to words from Old Peninsular Spanish, but also to words coined to name entities new to the colonizers. In the latter case, a frequent solution was to create compound based on what was known for them, as in *gallina de la tierra* (literally ‘chicken of the land’) for turkey. Archaisms include also phonological speech mannerisms (Hills 1906; Espinosa 1909; Espinosa 1911), such as aspiration of ‘f, h, s’ (*fogata* ‘fire’ [fogata] > [hogata]; *humo* ‘smog’ [umo] > [humo]), and diphthongization of some stressed words (*maíz* > *maiɪz*). Nevertheless, these phenomena arguably occur just in some sporadic areas of the linguistic community (Cobos 1983). Rather sporadic in present times as the result of contact with other dialects of Spanish

(1.4.4 below), the latter type of archaisms illustrate how language is shaped by actual communicative events. On the one hand, presence of archaisms is seen as evidence for the purity of New Mexican Spanish; on the other hand, it causes stigmatization because it often is considered a characteristic of less educated dialects.

1.4.2. Independent Developments

Another phenomena resulting from inconstant contact between New Mexican Spanish and other dialects was the development and retention of innovative changes. Thus, although the word *murciélagos* ‘bat’ already existed in the language, speakers found it more useful to extend the meaning of another existing word to create the descriptive compound *ratón volador* ‘lit. ‘flying mouse’. Another linguistic development characteristic of New Mexican Spanish is the ‘paragogic –e’³ with certain words such as *doctore* for *doctor* or *ele* for *él* ‘he’⁴.

1.4.3. Influence of English

Even though the influence of English is not exclusive of New Mexican Spanish, the fact that ties between this dialect and the Mexican, let alone Peninsular, one were often weak due to geography and politics has created the impression, among speakers and non-speakers of the dialect, that the impact is particularly pervasive. For New Mexican speakers, English loans and code-switching erode the basis of Traditional Spanish thus constituting a threat; for speakers of more prestigious dialects, they represent the incompetence to properly speak either Spanish, which would cause its corruption.

³ Although –e is typically the added vowel, scholar claims that the phenomenon is realized as well with –i.

⁴ Some scholars describe the phenomena as occurring with words ending in –r, –l or –n. Nevertheless, for the vowel to be added, it is necessary that the word is uttered at the end of an intonation unit. Moreover, the occurrence with those particular ending consonants seems to be related to type and token frequency of lexical items rather than to merely phonetic factors (Valdivia 2010a).

Despite such pessimist and negative ideologies, adoption and adaptation of English words is just another proof that speakers shape language to accommodate to the circumstances in which they are immersed. Thus, given the establishment of United States authorities as the result of the Guadalupe-Hidalgo Treaty, Spanish speakers created the terminology they would need to deal with their new political and social reality; for instance, *cherife* for Sheriff, and *mayor* [ma'jor] for mayor.

1.4.4. Popular Mexican Spanish and Standard Spanish Influence

Despite the mythical straight link of New Mexican Spanish to Castilian Spanish, the truth is that the Spanish of Mexico has played a main role in history of the dialect (see above). The influence of the contact, either sporadic or continuous depending on the time, is observable in the propagation of lexical items such as *papalote* 'kite', *cuates* 'twins, friends', and *chicote* 'riding crop, whip', which were adapted from Nahuatl words before coming to New Mexico.

Influence from Mexican Spanish comes not only from its popular variety, but also from the Standard one, namely, the speech associated with the educated segment of the population. Naturally, the notion of cultivated norm carries ideological perspectives and attitudes regarding the value, social and economic, of certain varieties of the language, which, in the case of New Mexican Spanish, has resulted in the gradual substitution of traditional items for normative (i.e. prestigious) ones; as attested by the use of *vi* 'I saw' instead of *vide*; *traje* 'I brought' instead of *truje*, and *pavo* 'turkey' instead of *torque*, *gallina de la tierra* or even *guajolote*⁵.

⁵ The nahuatlism still used in Mexico, particularly to refer to the live bird, or to the meat in traditional dishes (e.g. *mole de guajolote*). In other contexts, *pavo* seems to be the preferred word.

The study of New Mexican Spanish include not only changes in lexicon and morphology, and phonological phenomena (Bills 1996; Bills 1997; Bills and Vigil 2000; Vigil and Bills 2004; Torres Cacoullos and Vigil 2003; Torres Cacoullos and Vigil 2004; Torres Cacoullos and Aaron 2003a; Torres Cacoullos and Aaron 2003b; Brown 2005), but also issues in sociolinguistics (Bernal-Enríquez 2000; Gonzales 2005; Wilson and Martínez 2011; Wilson 2012), morpho-syntax (Edmunds 2005; Travis 2007; Torres Cacoullos and Travis 2010; Torres Cacoullos and Travis 2011; Villa 2010; Dumont 2006; Wilson 2013), and grammaticalization and subjectification (Cacoullos 2000; Torres Cacoullos and Aaron 2003a; Jenkins 2003; Wilson and Dumont 2015)⁶. The present research seeks to contribute to such literature, in general, and to the understanding emergent grammar and subjectification across languages, in particular, by examining the development of constructions and discourse function in the speech of speakers experiencing quotidian contact with another language and contact.

1.5. Organization of the Study

The following chapter discusses the theoretical framework for this study in much greater detail. Because the goals of this study are 1) to identify frequent constructions with verbs of cognition and speech in oral discourse in a situation of language and sociocultural contact, and 2) to investigate the distribution of those constructions and their relationship to discourse subjectivity, Usage-Based Theory was the most appropriate model to conduct the analysis of the data. Nevertheless, rather than adhering to just one theoretical approach, the study envisions Functional-Cognitive Linguistics as a continuum and, consequently, it takes into account the different ways such approaches

⁶ An extensive list is available in the website of the Department of Spanish and Portuguese at the University of New Mexico (http://spanport.unm.edu/common/docs/nmcoss_pubs)

conceptualize the binomial form-meaning. With this in mind, besides the main aspects of Usage-Based Theory, Chapter 2 presents other relevant theoretical models, such as Construction Grammar, as well as a review of linguistic literature regarding structural patterns in spoken language, in general, and cognitive and speech verbs in particular. Chapter 3 describes the data and methodology used in the coding and analysis of the data. Also, it presents an outlook of the history and characteristics of New Mexican Spanish. Subsequent chapters address the analysis of the data; chapter 4 describes and discusses patterns emerging from data as a whole, while chapters 5 and 6 focus separately on relevant patterns for cognitive and speech verbs, respectively, paying special attention to those constructions that differ from the patterns of the semantic class.

Chapter 2. Theoretical Framework

2.1. Organization of the Chapter

This chapter provides an overview of the principle theories that form the foundation for the methods and analysis of this study. I begin with a description of usage-based-theories and then proceed to a discussion of construction grammar. Both sections review as well research on the roles of conversation and interactional practice, and linguistic subjectivity. In section 2.5, I show how these two theoretical models can be combined to look at the grammaticalization of structural patterns in language. I conclude the chapter with a discussion of how recent work on frequency and usage-based linguistics, construction grammar and the role of subjectivity are to be combined in the approach to the data analyzed here.

2.2. Usage-Based Theory

Usage-based theory, as indicated by its name, is built upon the understanding that usage has a critical effect on linguistic structure, and it postulates that the units and structure of language emerge out of specific communicative events. The theory arises from the confluence of several research perspectives that conceive language as an embodied and social human behavior, but it distinguishes from its sources for the inclusion of any relevant method of obtaining evidence, either small or large corpora, diachronic or synchronic data, comparison across languages, or child language development, among others (Scheibman 2001; Bybee 2007; Bybee 2010; Bybee and Beckner 2009). Following Elman and Bates' (1997) claim that language evolved from a variety of cognitive abilities used in non-linguistic domains, usage-based theory “seeks to derive the mechanisms of language from more general and basic capacities of the human

brain, including sequential and statistical learning, chunking, and categorization” (Bybee and Beckner 2009:829).

2.2.1. Repetition, Chunking and Knowledge of Usage

It has long been observed that repetition plays a fundamental role in the development of “procedural knowledge” (Anderson 1993); the implicit knowledge about how to do an activity or process, from typing to riding a bicycle to driving a car. Through practice, the sequential and simultaneous individual tasks that comprise an activity eventually become automatized to the extent that we no longer stop to think the order in which such tasks must be performed, let alone how to do each one of them individually.

A similar phenomenon occurs in cognitive process: “repeatedly engaging in a task leads to the formation of that process in the long-term memory” (Bybee and Beckner 2009). Scholars suggest that, by tracking recurrent behaviors, our cognitive systems divide the activity into chunks that make the process more efficient, thus improving our performance. Furthermore, the principle applies the same to isolated, simultaneous, or sequential events, which relates to people’s ability to infer patterns even if little input exists. In other words, by observing that two or more independent events frequently occur together, people are able to anticipate the occurrence of one of such events before it actually happens, hence taking cognitive automatization to a higher level in the sense that it involves not just tasks within a process, but a set of sequential and simultaneous processes. For usage-based theory, there is little reason, if any, to assert that experiential knowledge and knowledge of a language happen independently from each other; on the contrary, because language is not distinct from other cognitive domains, it follows the same principle of representation and automation: whether there is a reason to choose a

particular construction over other available ones, linguistic structure arises from recurring specific events experienced by speakers. Furthermore, because such experiences include not only purely linguistic information, knowledge of a language includes as well implicit, procedural knowledge such as frequency of occurrence and statistical patterns (Bybee 1998; Bybee 2000; Thompson and Hopper 2001). Thus, for Spanish speakers, the use of an impersonal *se* construction over a passive one to represent a specific event may not be related to a conscious intention to eliminate agency; it may simply be a matter of familiarity because the first construction is much more frequent than the second one.

2.2.2. Categorization

Another general capacity fundamental to usage-based theory is categorization: the recognition of two events as being similar to each other, even if some differences exist (Haiman 1997). Furthermore, because such recognition on the basis of some shared properties allows grouping continuously varied input into “equivalence classes”, categorization operates as well across domains (Bruner, Goodnow, and George 1956; Pierrehumbert 2001). For instance, research has shown that people quickly learn to classify stimuli from examples they are exposed to in experimental settings (Medin and Schaffer 1978; Posner and Keele 1968).

Grouping events together based on shared properties rather than in a checklist of sufficient conditions has two main implications for categorization; first, category membership is a gradient phenomenon because boundaries are not absolute (Labov 1973), and, second, categories have an internal structure with some members being “better members” than others based on more or less shared properties, as a variety of converging methodologies has demonstrated (e.g. typicality ratings, response times for

classification, etc.). For Usage-based theory, this gradient membership must include information related to representations of experienced tokens, rather than operating based on minimal abstractions. Strong evidence for the claims comes from several studies (Posner and Keele 1968; Malt and Smith 1984; Nosofsky 1988) that have showed that when asked to classify different entities, people base the categorization not only on one abstract feature, but they also include individual features extracted from their own experience and are aware when certain features co-occur within a category, no matter whether those features are coincidental or highly predictable for the category in question. Based on this sort of findings, the theory claims that, within a category, items are stored as exemplars rather than being organized around a prototype, without any of said exemplars being, necessarily, better than the rest; and that in language and other domains, “[s]pecific instances of learning are retained in memory alongside the generalizations that gradually emerge from them” (Bybee & Beckner 2009: 832).

2.2.3. Exemplars and Networks

Usage-based theory proposes a non-reductive model of categorization in which each individual instance of language experience affects cognitive representation, and is stored, with all sort of information associated to it, as an exemplar: a “large cloud of remembered tokens” of a particular category (Pierrehumbert 2001:140); thus, it is implicit that the more a token is used, the easier will be for speakers to remember it and, consequently, the larger its “cloud” will become. Recollection, however, is not a sufficient condition for the development of a cloud, nor these clouds are merely a bunch of remembered tokens. Rather, they constitute cognitive maps: upon receiving linguistic input, speakers compare it to previously stored exemplars and map it onto the one or ones

to which it best matches. The chosen exemplar, then, becomes stronger and more easily accessible as more input is mapped onto it. Consequently, the more an instance is perceived, the stronger it becomes its exemplar. Furthermore, because several similarities can be established at the same time based on different features, the same token is stored within several exemplars. On the other hand, if the input does not match any existent exemplar, it is stored by itself as a new one and it will serve as a referent for latter inputs (Bybee 2001; Pierrehumbert 2001), processing that allows usage-based theory to deal with innovative or uncommon linguistic items without having to consider them as “exceptions”. For production, a strong exemplar is more accessible, which makes it the easiest candidate to be picked when speakers need to express an event related to it, thus in turn resulting in higher frequency of use.

Another theoretical advantage resulting from this model of categorization is the ability to explain both diachronic and synchronic regularities and changes. On the one hand, a production-perception loop may increase an exemplar’s strength, thus helping it to resist change, as it is the case of highly frequent irregular English verbs such as *to be* and *to go*, or Spanish *ser* and *ir*. On the other hand, the more speakers use an innovation, independently of the motifs behind such increase, the easier that change will become part of the system, as it is currently happening with the semantic extension of Spanish verb *ignorar* ‘to be ignorant of’, which now includes too the meanings of ‘to overlook’ and ‘to disregard’.

Similarities among exemplars form networks, which are complex and of various levels by nature because a token shares properties with several linguistic units at the same level based on different features. For instance, as a single unit, the word ‘unbelievable’

establishes links with semantically similar items, (e.g., incredible, doubtful, etc.), but it also connects with other words through smaller, internal units such as the suffix ‘-ble’ (e.g. readable, wearable, feasible) or the prefix ‘un-’ (e.g. unrealistic, unexpected, unapt), independently of the semantic features of each of these words. Because of such a process, gradience and variability are built into the exemplar model: “cognitive representations will reflect any new variants or ongoing changes in the distribution and frequency of variants” (Bybee and Beckner 2009:832).

Gradience refers to the fact that, more often than not, it is difficult to distinguish the boundaries of a grammatical category. Rather, most of them form a continuum in which linguistic units move along from one category to another. Clear examples of gradience are the verb’s non-personal forms; depending on syntactic factors, they change their function from that of a verb to that of a noun, an adjective, or an adverb without changing their form. To demonstrate how a network representation allows for gradience, (Bybee 1998) gives the example of *Tuesday* and *Wednesday*, in which speakers recognize ‘-day as a morpheme, but not the initial part of the word. If we considered that a word must be exhaustively dividable into morphemes that can be part of other somehow related words, then we would have to explain what *Tues-* and *Wednes-* mean and why they are not linguistically productive. For the exemplar model the issue does not represent a problem because it does not require that all parts of a linguistic unit be connected to some other unit.

Because of such non-exhaustive approach, a network representation, then, results “quite appropriate as it allows access to the sequence as a whole, while maintaining the links that identify the component parts” (Bybee and Beckner 2009:836). Furthermore,

since the model takes into account speakers' memory representation, these complex networks are able to capture semantic and pragmatic phenomena associated with a particular token, thus allowing usage-based theory to explain how speakers can process and store word strings, including idioms, as whole units with properties of their own, while still processing each word separately in other contexts. For instance, given an idiom such as *pull strings*, speakers establish networks (phonetic, morphological, lexical, etc.) through the properties of both *pull* and *string*; however, their linguistic experience also provide information regarding when and how the sequence needs to be understood as a whole unit whose meaning, use and function differ from that of its component parts.

2.3. Constructions and Other Multiword Units

2.3.1. Constructions

Given the way exemplars are processed and organized, constructions constitute a very appropriate unit to represent grammatical patterns. Broadly speaking, constructions are pairings of form and meaning, “conventionalized sequences of morphemes or words than contain a position that can be filled by more than one item⁷” (Bybee and Beckner 2009:842). The first consequence of such definition is that constructions range from single words –since they consist of bound morphemes to general configuration such as intransitive, transitive and passive constructions. It is important to mention, though, that, according to Bybee (2010), the term is more often used to refer a morphosyntactically complex structure that is partially schematic. The second theoretical consequence is that,

⁷ In a later work, when explaining the role of chunking in morphosyntax, Bybee (2010:35) states that “[a]ll sorts of conventionalized multi-word expressions, from prefabricated expressions to idioms to constructions, can be considered chunks for the purpose of processing and analysis”, which seems to imply that constructions are a particular type of unit different than idioms and others. To avoid any methodological confusion, in this paper, construction is used as a generic term, as indicated in chapter 2.3

because morpho-syntactic changes affect meaning and viceversa, “[e]ven the most general syntactic constructions have corresponding general rules of semantic interpretation” (Cruse and Croft 2004:257), which allows this theoretical model to account for different syntactic structures required by different verbs, as in (1) - (2)⁸, or even the same one, as in (3), without having to propose subcategorization frames for each case.

- (1) a. Tina slept.
b. *Tina slept bananas.
- (2) a. David consumed the bananas.
b. *David consumed.
- (3) a. He drove over the speed limit.
b. He drove his kids to school.

Moreover, in all construction grammar versions⁹, the pairing’s “meaning” component represents, in one way or another, "all of the conventionalized aspects of a construction's function, which may include not only properties of the situation described by the utterance, but also properties of the discourse in which the utterance is found [...] and of the pragmatic situation of the interlocutors (Cruse and Croft 2004:258). Meaning, then, refers not only to lexical features, but also to “grammatical” functions (e.g. the contrast known vs. new referent indicated by the use of a definite or indefinite article), and even idiosyncratic and communicative functions such as those performed by discourse markers.

⁸ From Cruse and Croft (2004:230)

⁹ Construction Grammar (Kay and Fillmore 1999), Lackoff’s (1987; 1990) and Goldberg’s (1995) models, Langacker’s (1987; 1991) Cognitive Grammar, Croft’s (2001) Radical Construction Grammar.

2.3.2. The Lexicon-Syntax Continuum

A consequence of the concept of constructions being open to different structural units, as well as different kinds of meaning is that, rather than constituting an inventory of discrete categories, constructions form a continuum from single words to complex and syntactically arranged, at least to some degree, multiword sequences. Cruse and Croft (2004:255)¹⁰ propos a lexicon-syntax continuum (Table 1 below) ranging from specific lexical items containing just one word (e.g. *this* and *green*) to syntactic patterns reflecting verbal argument structure. Since the present research focuses on structural patterns including at least one verb, and even a single unit such as *digo* ‘(I) say’, *creo* ‘(I) believe’ or *dicen* ‘(they) say’ includes morphemes for person, tense, and other categories, the smallest constructions considered in the analysis is the complex but bound one.

Table 1. The syntax-lexicon continuum

Construction type	Traditional name	Example
Complex and (mostly) schematic	syntax	[SBJ <i>be</i> - TNS VERB <i>-en by</i> OBL]
Complex, substantive verb	subcategorization frame	[SBJ <i>consume</i> OBJ]
Complex and mostly substantive	idiom	[<i>kick</i> - tns <i>the bucket</i>]
Complex but bound	morphology	[NOUN- <i>s</i>], [VERB-TNS]
Atomic and schematic	syntactic category	[Dem],[Adj]
Atomic and substantive	word/lexicon	[<i>this</i>],[<i>green</i>]

From Cruse and Coft 2004: 255

As it can be seen in the table, each construction type is expressed by a binomial, which indicates that the continuum has two dimensions; one going from the specific (substantive) to the structural (schematic), and another one moving from the simple (atomic) to the complex.

The first dimension distinguishes between substantive and schematic constructions, with the former being those in which all elements are fixed, and the latter those in which one or more of the parts can be replaced by other items. Examples of

¹⁰ The continuum is based on Fillmore et al (Fillmore, Kay, and O’Connor 1988) classification of idioms.

substantive constructions found in the data include the sequences *es decir* ‘that is to say / you mean’, as in (4), and *quiere decir* ‘it means’. Schematic constructions include the transitive one (SBJ-VTR-DO) realized by utterances such as (5) and (6).

- (4) .. *antonces le dijo .. el hombre a –*
a –
a la vieja,
.. % es dijir que no la dejas .. casar conmigo,
le dijo.
No,
le dijo.
.. Primero quiero verla muerta que –
que –
que se case contigo. (311-3B2: 550)¹¹
‘The guy told the old lady: “So, you mean that you won’t let her to marry me”, he told her. “No”, she said. I rather see her dead than married with you.’
- (5) *Yo sé que tenían gente allá en .. Kansas. (47-1A2: 682)*
‘I know (that) they have relatives over there in ... Kansas.’
- (6) *Los míos siempre han hablado los dos idiomas. (117-1A3: 759)*
‘My kids have always spoken both languages.’

In the second dimension, atomic constructions are those that cannot be divided into meaningful parts. On the opposite pole, complex constructions are composed by two or more meaningful elements. Since a construction can be a single word, the basic elements to take into account in such bidimensional continuum are morphemes. Examples of atomic constructions are *this* or *week*, or Spanish adverbs (e.g. *ayer* ‘yesterday’) and nouns with common gender (e.g. *artista* ‘artist’).

Table 1 shows as well that substantive constructions are particular instances of schematic ones, as shown in (5) and (6) above. Such a relationship allow us to consider both regular syntactic expressions (e.g. *kick a football*) and highly idiosyncratic idioms

¹¹ Interviews from NMCOS are labeled according to the number assigned to the participant (e.g. 214) and the session of recording (e.g. 1B3). The number after the label indicates the line of the transcription containing the token under analysis. Also, they are transcribed according to Du Bois et al (1993), method that is described in Chapter 3.

(e.g. *kick the bucket*) as realizations of the same schematic construction (e.g. SBJ-VTR-OBJ).

Since the present research focuses on structural patterns including at least one verb, the smallest constructions to be considered will be complex but bound; even a single unit such as *digo* ‘(I) say’, *creo* ‘(I) believe’ or *dicen* ‘(they) say’ includes morphemes for person, tense, aspect and mood.

In an intermediate point of the continuum, Cruse and Croft (2004) situate collocations, combinations of words that, for some reason, speakers prefer over other expressions that seem to be semantically equivalent. For instance, although *toasted* and *roasted* basically describe the same process, each word is used in different combinations: *roasted meat* and *toasted bread*. In this sense, collocations are fixed constructions, at least within a linguistic community, because speakers use them as such; even if, in theory, they could have some variability (see also Matthews 1981; Nunberg, Sag, and Wasow 1994).

Collocations are sometimes called formulaic expressions and prefabs. That terms seems to be synonyms is evident in Bybee’s (2006:713) definition of prefab as “word sequences that are conventionalized, but predictable in other ways” and which “occur repeatedly in discourse and are known to represent the conventional way of expressing certain notions” (Erman and Warren 2000; Sinclair 1991; Wray 2005); for instance, *prominent role*, *mixed language*, and *think of*. Nevertheless, given the fact that prefab is used as well to referred to constructions which have one or more open slots, such as *X drives Y crazy*, or Spanish *quedarse* + adjective ‘become + adjective’, and whose meaning may be idiosyncratic, it may be appropriate to use both terms to name different

constructions. Moreover, since prefabs' idiosyncratic meanings often require a metaphorical interpretation, the distinction between them and idioms is not always clear; whereas the distinction between an idiom and a collocation as defined by Cruse and Croft seem to be more categorical.

Based on these definitions and claims, in the present research 'construction' will be used as a generic term for structural patterns ranging from those comprised just by a verb in a personal form with unexpressed subject (e.g. *creo* 'I think') to specific utterances consisting of a verb and one of more constituents (e.g. *me dicen que les hable en inglés* 'they ask me to speak to them in English') to syntactic schemas such the ditransitive construction Sbj-V-DO. When addressing specific constructions, their position in the lexicon-syntactic continuum as well as their function in the discourse will be determined based on three dimensions: productivity, transparency of meaning and analyzability (Bybee and Torres Cacoullos 2009). In chapter 3, I will describe and discuss what morphosyntactic features will be used to operationalize these three dimensions.

Productivity refers to how many different items that a construction accepts to fill one of its open slots. Thus, the schematic construction SBJ V_{TR} OBJ is more productive than SBJ *read* OBJ because the verb's slot can be filled by any verb taking a direct object, as in (7) while, in SBJ *read* OBJ, the range of items that can fill the subject and object slots is restricted by the meaning of the verb, as illustrated by (8). In turn, SBJ *read* OBJ 'is more productive than SBJ *read that* OBJ because, in the latter case, the object slot is restricted to clausal items.

- (7) a. I already have that book.
b. That book has received good reviews.
- (8) a. I already read that book.
b. *I already read that food.

- (9) a. I read that the book has received good reviews.
b. *I read that the book.

Transparency of meaning relates to compositionality versus idiosyncrasy of a construction. Thus transparency is higher in constructions whose meaning results from the sum of the individual parts (e.g. Boys never change) than in constructions in which such sum does not correspond to the meaning of the whole sequence (Boys will be boys).

Finally, analyzability indicates the extent to which individual parts of a construction “are associated with the etymologically same units in other constructions”. For instance, the Spanish construction *hacer enmiendas* ‘to make corrections’ is more analyzable than *hacer las paces* ‘to make peace (with someone)’ because the plural noun *paces* ‘peace-plural’ is rarely used outside of the construction. Such a restricted use may result in speakers, especially non-native ones, not recognizing the relation between *paz* and *paces*.

2.3.3. Idioms

Due to such an inclusive approach to meaning, construction grammar is able to address a particular kind of constructions which are problematic for componential models of grammar because their meaning, let alone their function, cannot be explained by highly general rules: idioms. By definition, idioms are multiword grammatical units with, at least, some degree of idiosyncrasy (Bybee 2007; Cruse and Croft 2004), as the ones shown in (10) below:

- (10) a. It takes one to know one.
b. bring down the house.
c. boys will be boys.
d. (X) blow X’s nose.

Rather than being an homogenous class, there are different kinds of idioms depending on their morphosyntactic properties, and their semantic and pragmatic meaning; nevertheless they all share a set of features first defined by Nunberg, Sag and Wasow (1994:492–493):

- (11) a. Conventionality: the meaning and use of the construction cannot be determined, at least not completely, from those of their parts.
- b. Inflexibility: the construction presents a restricted syntax; e.g. shoot the breeze vs *the breeze is hard to shoot.
- c. Figuration: the construction has a non-literal meaning; e.g. take the bull by the horns
- d. Proverbiality: the construction describes a social activity by comparing it with a concrete one; e.g. spill the beans.
- e. Informality: their language typically relates to informal speech registers
- f. Affect: they convey an judgmental or affective attitude towards the situation they are describing.

Among these properties, conventionality is particularly relevant for construction grammars and usage based theory because it implies that idioms must be stored as whole units in speakers' minds and, consequently, they are part of speakers' grammatical knowledge. As well, idioms can be classified in different ways based on the characteristics one focuses on. One of the earlier classifications for these grammatical units is Fillmore et al's (1988), which, in its final version, proposed the following set of distinctions:

- (12) a. encoding / decoding idioms
- b. idiomatically combined expressions / idiomatic phrases
- c. grammatical / extragrammatical idioms
- d. substantive / schematic idioms
- e. idiom with pragmatic point / idioms without pragmatic point

In the first distinction, encoding idioms refer to units whose meaning can be interpreted using general (i.e. compositional) rules, but it is still conventional in the sense that there is not a real reason justifying why a certain word was chosen for that idiom instead of another one within the same semantic group. For instance, although a person, even a non-native speaker of English, can figure out the meaning of *answer the door* and *bright red*, these constructions do not seem to be the natural-sounding way to express that

particular idea. On the opposite side of the distinction, decoding idioms cannot be decoded by the hearer based only on compositional rules; for instance, *kick the bucket* or its Spanish equivalent *estirar la pata*¹². An interesting characteristic of these idioms is that they conform to canonical morpho-syntactic patterns (e.g. SBJ V-TR OBJ), so if a hearer does not know that the construction is a convention, he/she could interpret it literally, with the consequent communicative confusion.

In the second distinction (originally in Nunberg, Sag and Wasow (1994)), an idiomatically combining expression is an idiom whose parts can still correlate to a literal meaning, while an idiomatic phrase is a unit in which such correlation does not exist. For instance, in *spill the beans*, *spill* can be analyzed as a metaphor for ‘divulge’ and *beans* as one for ‘information’ respectively, therefore it is an instance of the former type of idioms. On the contrary, it would be hard to establish the same kind of relation between *kick* and *die* in the idiom *kick the bucket*.

Grammatical idioms are units that, although parsable by the language’s general rules, are semantically irregular, that is to say, they are highly conventional. Both *spill the beans* and *kick the bucket* are examples of these idioms, even if they represent different types in another distinction. This latter phenomenon clearly illustrates how complex and multidimensional idioms are. Unlike grammatical idioms, extragrammatical ones do not conform to the language’s syntactic rules, behavior that in itself gives them an idiomatic nature. Examples of these idioms are *by and large*, *so far so good* and *all of a sudden*.

¹² The degree of idiosyncrasy of the phrase is evident if we compared with *estirar las piernas* ‘go for a walk lit. Stretch the legs’. The same verb and the same body part (although one using the colloquial term and the other the formal one) are used to convey completely different meanings.

The fourth feature distinguishes between idioms in which all elements are fixed (substantive), as *It takes one to know one*¹³, and idioms (formal) in which at least one slot can be filled by any syntactically and semantically appropriate element; for instance, in *(X) blows X's nose*, any nominal phrase referring to a nose's possessor can fill the X slot: *They all blew their noses*.

In the last distinction, idioms with pragmatic point are those which are used in particular communicative contexts beyond their purely semantic meaning; for instance, *Good morning* or *See you later*, are used to open and close, respectively, a conversation. Obviously, idioms without pragmatic point do not perform such type of functions, no matter how conventional they may be (e.g. *all of a sudden*).

From the previous five distinctions, Fillmore et al later propose a three-way categorization of idioms. The first category are idioms consisting of unfamiliar pieces unfamiliarly arranged, which results in lexically, syntactically and semantically irregular constructions. More importantly, the category brings to light the fact that certain words occur only in an idiom, as in *kith and kin* 'family and friends' (Cruse and Croft 2004: 235). This latter phenomenon is essential to the idioms' categorization since words that do not exist outside an idiom cannot be part of a syntactic category. Nevertheless, belonging to this category does not necessarily mean an idiom is an idiomatic phrase; it can be an idiomatically combining expression because "the parts of the construction can be made to correspond with the parts of its meaning" (Cruse and Croft 2004:236). For instance, the schematic idiom 'The X-er, the Y-er', has unfamiliar pieces, --the two occurrences of 'the'—arranged in a parallel syntactic structure following by a clause

¹³ The construction would no longer be an idiom, or have any sense at all, if, for instance, one changed the tense: **It took one to know one*.

indicating the degree expression (e.g. The bigger they come, the harder they fall), but the relation between the degrees of both parts of the construction can be observable. Thus, the fact that the syntactic arrangement of an idiosyncratic idiom fits that of a schematic idiom does not exclude its existence as a construction by itself.

The second category of idioms, familiar pieces unfamiliarly arranged, are lexically regular because they contain words that speakers use in utterances other than the idiom, but they are syntactically irregular because they do not conform to the patterns of the language and, consequently, their meaning cannot be obtained by compositional rules (e.g. 'in point of fact').

Finally, idioms consisting of familiar pieces familiarly arranged are those in which both the lexicon and syntax are regular, but the semantics is irregular; their elements can occur outside the idioms, they conform to the syntactic rules of the language, but they cannot be interpreted literally (e.g. 'pull X's leg').

On a similar line of thought, Du Bois (2003) argues for a conception of grammar that takes into account the structural irregularities and pragmatic nuances found in actual discourse. The author claims also that, in order to understand the system of grammar, it is essential to privilege the study of spoken language because it is the one that closely embodies what language is. Specifically, Du Bois regards spontaneous conversation as the ideal genre to address linguistic inquiries because its dialogic nature allows observing discourse as it is being produced, with its on-line planning, and interlocutors' production and interpretation. Moreover, because it is highly contextualized, conversation reflects the cognitive frames and processes that both speaker and hearer are employing to express messages and build the discourse. Thus, from a perspective based on typological research

on grammar and discourse-pragmatic functions, he makes three general assumptions regarding grammar and discourse: a) in communicative interaction, speakers exploit what is available in the grammar to achieve their communicative goal; b) recurrent patterns arising from discourse often escape what is predicted by grammatical rules; and c) grammatical structure tends to evolve following discourse patterns: “Grammars code best what speakers do most” (Du Bois 2003:49). An example of the first assumption is the group of Spanish impersonal *se* constructions often referred to as *se accidental* ‘accidental *se*’¹⁴. Because Spanish impersonal *se* constructions eliminate the agent’s role by not including a slot for its syntactic expressions, speakers prefer them over active constructions when expressing an event for which they do not feel responsible. The contrast is illustrated by the sentences in (13) below; while the speaker is both the subject and the agent of the event in (13a); his/her role is reduced in (13b) to that of a mere affected entity (*me* ‘to me’). The event, then, is expressed as an accident, not a conscious act.

- (13) a. *Perdí las llaves.*
 Lose-1sg-pret my keys
 ‘I lost my keys’
- b. *Se me perdieron las llaves.*
 SE-IO-1sg lose-3pl-pret my keys
 ‘I lost my keys (lit. My keys got lost)’

The second assumption is illustrated by verbs such as *creer* ‘to believe / to think’, *saber* ‘to know’, *decir* ‘to say’ and others. Although these verbs are classified as

¹⁴ Impersonal *se* constructions accept any verb with an implicit human agent. *Se accidental* constructions include only verbs implying non-intentional actions, e.g. to lose, to forget, to be late, to misplaced, to broke, to burn, etc.

transitive in normative grammars and dictionaries, in oral discourse they often occur without any syntactic complement, as in the following examples from NMCOS¹⁵.

- (14) E: *Y dónde nació su papá.*
 U: *Creo que nació en Buena Vista también.*
 E: (H) *Y su mamá?*
 U: *Ella yo creo que --*
*yo creo que también **no sé**=,*
no me acuerdo [donde=] XX – (214-1A1: 156)
 ‘E: And, where was your dad born? U: I think he was born in Buena Vista as well. E: And your mom? U: She, I think that... I think [she was born] there too, **I don’t know**, I don’t remember where...’

An example of Du Bois’s third assumption is the grammatical requirement of using the preposition “a” before animate and specific direct objects in Modern Spanish, as shown in (15) below. Because Spanish allows for the direct object to appear before the verb, even if it a marked word order, the rule resulted from the frequent inclusion of the preposition in utterances which speakers otherwise would considered ambiguous.

- (15) a. *Óscar vio una película.*
 Sbj-V-DO
 ‘Oscar saw a movie.’
 b. *Óscar vio a su amiga.*
 Sbj-V- “a” -DO
 ‘Oscar saw his friend.’

Du Bois’s assumptions and their implications, however, can be attested as well in less spontaneous oral genre such as the interviews in NMCOS. Given the purpose of these interviews (see chapter 3), speakers were asked directly for the names of a variety of animals, food and objects, which, it could be argued, do not only restricts interviewees’ spontaneity, but also set the structural framework they need to use in their answers. In reality, as demonstrated in chapters 4-6, even in this topic-specific

¹⁵ Examples from NMCOS will presented according to Dubois et al’s transcription method; in which interrogative marks, for instance, are not used as indicated by punctuation rules of the language. More details about the method and the symbols used to represent spoken features are included in appendix A.

interactions, interviewers and interviewees exploit a range of linguistic structures to elicit answers, keep the conversation going, check the interlocutor's attention and, more importantly for the purpose of this study, convey different attitudes and pragmatic nuances. For instance, when asked for the name of an object, speakers rarely use the verb *llamarse* 'to be, lit. to be called', which would provide an objective answer. Instead, they prefer the construction *Le dicen X* 'People call it X / It is called' which, depending on the context, may convey lack of certainty, lack of commitment towards the veracity of the answer or even a subtle disagreement regarding the name used by another linguistic community.

Another central claim of Du Bois's proposal is that analyses heavily based on the form-meaning pairings can describe linguistic phenomena only to a certain extent, and that "as applied to discourse may reach its limits once one undertakes the more abstract and highly grammaticized roles" because they "treat discourse tokens as mere means of arriving at, or confirming evidence of, a type description" (2003:51). Considering the observation regarding the methodological limits of form-meaning pairings, Du Bois's approach may seem irrelevant, if not contrary, to the goals of this study. Nevertheless, his proposal of a methodological approach that looks for the correlation between grammatical structure and discourse-pragmatic function, as well as its claim that "[t]he accumulated linguistic experience of a lifetime amounts to a body of utterances encompassing form, meaning, and contextualization" (2003:52) evidently agree with Cognitive Grammar's and usage-based theory's assumption that speakers experience and contextual information are necessarily included in the "meaning" component of constructions. Furthermore, his conclusion that grammar contributes to discourse an

architectural framework rather than meaning can be related to the fact that different constructions, from merely compositional ones to prefabs to idioms, can be realizations of, or emerge from, the same schematic construction.

2.4. Constructions, Grammar and Conversation

Ono and Thompson (1995) propose a model for English syntax that strongly suggests that rather than structurally matching the perfect sentences found in books of grammar, the linguistic units we usually refer to as phrases and clauses are constituted by those basic units observed in conversation. Moreover, they claim that these units are constantly evolving because of the cognitive and social constraints inherent to natural interaction.

They refer to two particular phenomena that provide clear and strong evidence for the existence of constructional schemas: adherence and co-construction. The first phenomenon refers to speakers' ability to reproduce and recognize schemas even when they are interrupted by other sounds, words or phrases. In the abovementioned study, Ono and Thompson noted that sometimes speakers come across intervening material when producing an instantiation of a schema (e.g. a clause); however, the interference barely has any repercussion in the realization of the schema. The fact that this occurs indicates speakers attend to schemas even across utterances. Example (16) from the data for the present research, illustrates the phenomenon. As it can be seen, even with the prosodic truncation at the end of line 2, the vocalization in line 3, and the discourse marker in line 4, the speaker produces the schema 'MAIN CLAUSE *that* SUBORDINATE CLAUSE' in line 2 and 5: *yo creo que la mujer se le fue con otro* 'I think his wife ran away with another man'.

- (16) 1 *No sé qué pasaba.*
 2 *yo creo que la mujer —*
 3 *(H) Hm.*
 4 *% (Hx) pues no sé=.*
 5 *Se le fue con otro, (214-1B3)*
 ‘I don’t know what was happening. I think his wife—Hm, well, I don’t know, she ran away with another man.’

The second phenomenon, co-construction, refers to speakers’ ability to collaboratively produce a schema alternating turn sequences. Ono and Thompson authors find, to do so, speakers use two major strategies: a) one speaker adds to a unit spoken by the interlocutor, and b) the second speaker completes the utterance that the first speaker had started. Such a collaborative production demonstrates that “abstract schemas are shared between the two interactants” (Ono and Thompson 1995:226). In other words, because speakers do not plan together what one of them will say, let alone how s/he will express it, the fact that participants are able to finish each other’s sentences demonstrates that they share information not just about the content of the discourse, but also, and more importantly, about the constructions the interlocutor may use to express his/her idea. Thus, even if many, if not most, actual utterances “may only roughly approximate the schemas which speakers seem to be using” in designing the output, speakers are not only capable of carrying on the conversation; they also are “remarkably adept at making adjustments between the perceived output and the schemas we hypothesize they are working with” (Ono and Thompson 1995:222). That is to say, despite the interruptions, overlapping, and reformulations observed in conversation, it is possible to identify the “cognitive routines” (i.e. constructional schemas) constituting grammar.

2.5. Usage-Based Construction Grammar in Grammaticalization

For usage-based theory, using constructions to represent grammatical patterns is particularly appropriate both because they are “direct pairing of form and meaning with no

intermediate level of representations” (Bybee 2010:26), and because the fact that constructions form a continuum, and move along it based on the interaction of several factors, clearly supports the claim that individual instances of experience impact cognitive representation. Furthermore, the existence of a continuum implies that contiguous constructions may share some of the features and functions, which allows usage-based theory to explain how new constructions arise from previously established ones, how new and old constructions compete with each other, and how they grammaticalize. These three latter phenomena will be discussed on detail on the following sub-sections.

2.5.1. New Constructions Emerge from Already Established Ones

As it has been stated several times in this chapter, linguistic expressions include information about the context in which they were produced or perceived. Thus, even if two or more constructions are formally related, one of them may emerge as a new construction if and when speakers associate it with a particular interpretation. The idiom ‘The bigger they are, the harder they fall’ is a good example of this phenomenon; although it is part of the schematic construction ‘The X-er, the Y-er’ and so it is related to other constructions such as ‘The sooner you start, the sooner you’ll finish’, it is interpreted as an idiom with non-literal meaning because speakers’ memory of representation has registered the implications and nuances connected to the communicative context¹⁶. In this study, the prefab *le dicen X a Y* ‘X is called Y’, which conveys an evidential nuance, arises from the schematic ditransitive construction ‘Sbj *decirle* DO IO’. Nevertheless, since the IO’s referent is an inanimate object, the verb is

¹⁶ See Fillmore, Kay, and O’Connor 1988 and Cruse and Croft 2004 for a further discussion.

interpreted not as transmission of information, but as the naming of a referent. Certainly, both meanings are included in the normative *Diccionario de la Lengua Española* (RAE), but the entry for naming is presented in the third place, suggesting is not the main use of the verb. Moreover, this sense occurs with only two subject-tense combinations, *dicen* and, less frequently, *decían*, which implies a strong correlation form-function.

2.5.2. Old and Emerging Constructions Compete

Naturally, a construction's new form or a new meaning does not emerge from one day to the next; it is a gradual and slow process which causes that, during a period, the new construction coexists and compete with the one from which it arose, or even the existence of two or more ways of expressing the same or similar meanings. A current example for this phenomenon comes from Spanish *imprimido* and *impreso*, *imprimir* 'to print' regular and irregular, respectively, past participles. Even though prescriptive rules of grammar clearly state that the regular form must be used in perfect tenses (e.g. *he imprimido la solicitud* 'I have printed the application'), and the irregular form as an adjective (e.g. *una solicitud impresa* 'a printed application'), speakers very frequently use the irregular form in both constructions. Thus, normative *HABER + imprimido* 'to have printed' coexists and competes with the new *HABER + impreso* construction. Eventually, the preference for the irregular form over the regular one could result in the latter being prescribed as archaic and grammatically incorrect, just as it happened to other regular past participles such as *rompido* 'broken'.

2.5.3. Grammaticalization of Constructions

Due to its frequency in particular contexts, a construction may become a grammatical unit. A classic example of this phenomenon is the development of 'going to'

in English, as well as the equivalent *ir a* in Spanish, from a construction expressing purpose into a marker of future: X is going to some place to do something > X is going to do something. Because of frequency of use, the construction experienced semantic inferencing and, eventually, the physical motion sense of *going* got reinterpreted as displacement in time, thus losing its motion sense and becoming an auxiliary verb (Bybee 2010). Moreover, the same phenomenon can result in the subjectivization of a construction, as it is happening with Spanish *dizque*. From being a syntactic schema *dice que* ‘he-she-it says that’, it experienced phonetic reduction, first, and then it was reinterpreted as an adverb meaning ‘apparently, presumably’, as in (17):

- (17) ...*que nunca mas festejada coronacion **dizque** se hizo...* (Davies 2002)
‘**Presumably** a more celebrated coronation never existed...’

Recently, though, the word is occurring in contexts in which the speaker is referring something he/she did, as in (18), where the speaker states that he was pretending to water the plants in the hallway in order to listen what his neighbors were talking about.

- (18) *Tras varios intentos que consistieron en salir al zaguán [...] **dizque** a regar las plantas en el momento preciso, coseché únicamente miradas de desconfianza [de los vecinos].*
‘After trying several times [to listen the neighbors’ private conversation] by going out the hallway, **supposedly** to water the plants, I only got suspicious looks [from the neighbors].’

Thus, because the speaker is indeed the source of the information, *dizque* does not longer indicate hearsay nor just ‘apparently’; now it also means ‘supposedly’, with an implication that a state or activity was a pretense or not even carried out. Furthermore, this novel sense has allowed, *dizque* to change its syntactic category thus reducing its scope; from modifying a whole sentence, now it can function as an adjective, often conveying the speaker’s disbelief or even sarcasm, (cf. Valdivia 2010b), as in (19) and (20).

- (19) *si este **dizque** movimiento de depuración sindical triunfa, ya podemos cortarnos la coleta...*
 ‘if this **so-called** movement for labor purge successes, we are screwed...’
- (20) *lo que estamos es frente a una simulación de un dedazo, que ahora será envuelto en un celofán, **dizque** democrático.*
 ‘...what we are seeing is an act of autocracy, one which will be wrapped with a **pseudo**-democratic cellophane.’

2.6. The Role of Repetition

Earlier in this chapter, frequency was mentioned as a central factor in the formation of development and strengthening of exemplars, in the creation of cognitive networks, and in language variation and change, among other phenomena. In particular, it was stated that increases in frequency resulting from production-perception loops may operate in opposite directions, sometimes easing and speeding language change, but other times helping an item to resist such phenomenon. This section will discuss the effects that both type and token frequency have on language, as well as the mechanisms allowing them to do so.

2.6.1. Type Frequency

The different and sometimes opposite effects of type and token frequency relates to how each one of them count linguistic units. Type frequency focuses on the number of different patterns, schemas, or lexical items occurring in a text, no matter the specific realization of the tokens. For instance, when observing the occurrence of verbs of cognition in a discourse, type frequency will count as one item all the tokens of the verb *saber* ‘to know’, independently of whether they occur in a non-personal or a personal form, and, in the latter case, independently of their subject and tense. Thus, since in the corpus for this research there are 21 different cognitive verbs but only 11 speech verbs, the former class has a larger type frequency than the later one.

Beyond raw numbers, type frequency largely determines a pattern's degree of productivity. A pattern applied to a high-type frequency category will be highly productive because its effect will easily propagate through an extensive network. This is why new coined verbs in Spanish adopt regular inflections: type frequency for regular verbs is much larger than for irregular verbs.

For a construction, the more different items that can fill one of its open slots, the more productive it will be. For instance, in a study on variation between Spanish aspectual constructions involving *estar* 'to be' and *andar* 'to go around', (Torres Cacoullos 2001) finds that although both *estar* 'to be' and *andar* 'go around' have grammaticalized as auxiliary verbs for Spanish progressive constructions (aux + V-ndo), each one of them tends to appear with different classes of main verbs. While the construction *estar* + V-ndo (e.g. *estar escuchando* 'to be listening') accepts a wide range of verbs, and, consequently, has a high type frequency, *andar* + V-ndo (e.g. *andar paseando* 'to be strolling around'), tends to occur with verbs conveying real or figurative motion, which results in a lower type frequency. Thus, a verb such as *esperar* 'to wait' appears more frequently with *estar* than with *andar*. In a related research Bybee and Torres (2009) find that the low degree of productivity of *andar* + V-ndo has resulted in a usage's decline; their historical data show that the verb represented 21% of all aux + V-ndo utterances in the 13th century (21%), but only 3% in the 19th century. Furthermore, since the most frequent realization of *andar* + V-ndo is *andar buscando*, the latter has been described as a set phrase (Spaulding 1926) or a case of lexical specialization (Squartini 1998).

It is important to say, though, that type frequency is not the only factor determining productivity. Torres Cacoullos (2001) finds that, in Mexican Spanish, the *andar* + V-*ndo* construction is often associated with non-educated registers; while in a corpus of educated speech the *estar/andar* ratio is about 21 to 1, in a corpus of popular speech¹⁷ the ratio is just a little bit higher of 3 to 1. Thus, because speakers are predisposed to avoid the construction in formal contexts, verbs occurring with it are generally restricted to those conveying ‘non-sophisticated’ actions such as *sembrar* ‘to sow, to farm’; as a consequence, the type frequency, and the productivity of *andar* + V-*ndo* is lower than those of *estar* + V-*ndo*. Productivity, then, is not just a matter of type frequency; it is a matter too of how open or restricted a category is:

"The more lexical items that are heard in a certain position in a construction, the less likely it is that the construction will be associated with a particular lexical item and the more likely it is that a general category will be formed over the items that occur in that position. The more items the category must cover, the more general will be its criterial features and the more likely it will be to extend to new items. Furthermore, high type frequency ensures that a construction will be used frequently, which will strengthen its representational schema, making it more accessible for further use, possibly with new items." (Bybee 2007:205)

2.6.2. Token Frequency

Rather than counting for the number of times a pattern or a lexical item appears in a text, token frequency refers to the number of times a particular realization of an item (from allophones to word strings) appears in a text. For instance, in this project’s corpus, *creo* ‘I believe / I think’) has a token frequency of 109 occurrences because that is the number of times that *creer* appear in that particular subject-tense combination.

¹⁷ Data for Torres (2001) come from *El habla de la Ciudad de México* (UNAM 1971) and *El habla popular de la Ciudad de México* (UNAM 1976).

Several studies have found that high-token frequency has a reducing effect on words and phrases because repetition causes neuromotor behavior to be more fluent: based on experience, speakers are able to perceive a unit even if not all of its sounds are fully uttered. For English, Bybee and Scheibman (1999) show that *don't* is reduced more often in *I don't know* than in other sequences because of, precisely, such a construction has a high token frequency. For Spanish, intervocalic [ð] deletion is more common in frequent words than in those occurring less often. Thus, is more likely to find the deletion in *cansado* > *cansao* 'tired' than in *aterido* > *aterio* 'frozen stiff'.

In other levels, e.g. morphology, high token frequency has a conserving effect: frequently used items are more resistant to change because their strength helps them to avoid the influence of productive patterns in the language. For instance, for Spanish Past tense, frequently used irregular verbs (e.g. *dije* 'I said') are less likely to become regular ones than verbs with a low token frequency (e.g. *produje* > *producí* 'I produced'). The phenomenon is illustrated too by linguistic units with two or more meaningful parts (i.e. *insane* or *pull strings*); there can be at least two ways of processing it: either as individual units combined in a larger one, or as a whole unit. Based on experience, speakers are able to apply the pertinent process in each context of use; nevertheless, the more frequent a sequence becomes, the more likely it will be processed as a whole unit even to the extent that speakers may no longer recognize one of the elements as a separable unit, as it has happened to the prefix in *insane*. As a result of this mechanism, sequences with a high frequency "are more entrenched in their morpho-syntactic structure, and therefore resist change on the basis of more productive patterns" (Bybee and Beckner 2009:840). In

grammaticalization, this effect causes originally complex sequences to lose semantic and syntactic transparency and to develop a grammatical meaning (e.g. *be going to*).

Another effect of high-token frequency, related to the conserving one, is autonomy. In said phenomenon, a complex form (i.e. a word or a string) loses its internal structure as it loses ties with etymologically related forms. The word *insane* illustrates too this phenomenon; as the frequency of *insane* increases, speakers' awareness regarding the prefix status of *in-* decreases and, eventually, the link between *sane* and *insane* is not completely observable.

2.7. Language and Subjectivity

As stated in the introduction, besides determining the role of language use on the emergence of different types of constructions, this research seeks to demonstrate that some of such constructions, namely those showing a higher degree of morphosyntactic fixation, often convey a variety of meanings related to the domain of subjectivity.

The concept and study of subjectivity is typically associated with Benveniste (1971)¹⁸, who defined it as the speaker's ability to posit herself as the subject of her sentence. An essential aspect of said definition is the distinction between the syntactic subject and the subject of the enunciation; so, the most basic mechanism to express subjectivity would be the use of personal pronouns, as well as spatial and temporal deictics, because these elements allow the speaker, or her environment, to become the utterance's referential anchor. Hence, from this point of view, a sentence would be subjective simply by including the pronoun *I*. Nevertheless, beyond the mere auto-reference of the speaker in the linguistic structure, Benveniste also launches the idea that

¹⁸ In reality, the study of subjectivity had already been approached by scholars such as Breal (1964) and Bühler (1990). See Traugott (2003:125) for further details.

some constructions express speakers' attitude and, more importantly, that language is predominantly marked by the expression of subjectivity: "In *I suppose* and *I presume*, there is an indication of attitude, not a description of an operation. By including *I suppose* and *I presume* in my discourse, I imply that I am taking a certain attitude with regard to the utterance that follows" (Benveniste 1971:228).

It is important to note that, rather than being uniformly defined, the concept of subjectivity varies according to the linguistic framework from which the phenomenon is approached. From a logical perspective, formalist approaches, for instance, see subjectivity as the expression of non-factual information, or information that cannot be verified using values of truth. Because of this, Formalism does not pay enough attention to the phenomenon. For functional and cognitive approaches, on the other hand, linguistic subjectivity needs to be separated from logical-philosophical subjectivity, and to be understood as the "way in which natural languages, in their structure and their normal manner of operation, provide for the locutionary agent's expressions of himself and his own attitudes and beliefs" (Lyons 1982:102), as well as the speaker's cognitive ability to take the interlocutor's perspective. Lyons (1994) later distinguishes between subjectivity and locutionary subjectivity, the former referring to the conceptual property of the speaker to be "a subject of consciousness or a thinking subject" (1994:11), and the latter referring to the concrete construction that speakers use to express their own attitudes and beliefs in a particular utterance. In addition, Lyons claims that subjectivity is affected by factors such as prosodic and paralinguistic modulation, thus introducing the idea of subjective gradualness.

Based on Lyons' locutionary subjectivity, functional linguists (Finegan 1995; Scheibman 2001; Traugott 1989; Traugott 1995) define subjectivity as the expression and representation of the speaker's attitudes and perspectives in the discourse. From this approach, the inclusion of *yo digo* 'I say' in (21) makes the utterance more subjective than the one in (22). Additionally, it is worth noting that *yo digo* in (21) is used not as speech act, but as a discourse marker; it is not simply uttering a fact, but mitigating an assertion.

(21) *Eso está mal, yo digo.*
"That's wrong, I think". (lit. 'I say')

(22) *Eso está mal.*
"That's wrong"

Unlike functional linguistics, cognitive linguistics conceives subjectivity not as speakers' involvement towards the utterance, but as an "asymmetry between a perceiving individual and the entity perceived" (Langacker 1990:7); that is to say, the degree of subjectivity lies in how speakers conceptualize their role in the event, and it is codified by the presence/absence, and type of grammatical material referring to the speaker. In words of Langacker (1990:10): "[i]f the speaker's involvement in a linguistic expression is merely as the conceptualizer, then the speaker/hearer role is almost wholly subjective because he/she is external to the scope of predication." From this point of view, the most subjective utterances are those conveying description of persons other than the speaker. In the opposite extreme, utterances including personal deictics are more objective because they profile the speaker as the referential point (ground, in Langacker's terminology). From this conception of subjectivity, (23) is subjective because the speaker is completely absent in the construction of the event; since 'across' expresses physical motion and direction, its meaning does not require the speaker's cognitive involvement.

On the other hand, (24) is an objective utterance because the meaning of ‘across’ lies in the spatial construals from the speaker’s perspective. She performs the cognitive movement from Vanessa’s to Veronica’s position. Both examples are Langacker’s:

(23) *Vanessa jumped across the table.*

(24) *Vanessa is sitting across the table from Veronica.*

Despite the differences, functional and cognitive approaches to subjectivity are not necessarily exclusive, but rather they can complement each other. To start, they agree that is practically impossible for a speaker not to appear in the utterances; because discourse consists of a chain of inferences (Verhagen 1995), utterances never occur in isolation; therefore, the speaker/hearer is always somehow involved in the construction, and interpretation of the events. Furthermore, since discourse rarely is purely objective or purely subjective, we can expect both conceptualization of subjectivity will coincide in some areas of the continuum: speakers may state their attitude by not only explicitly saying what they think or want, but also elaborating events in a way the information of the utterance appears to be from someone else. Hence, the speaker’s linguistic absence in the utterance may be seen both as her not being a participant of the event being expressed, and as a strategy to distance herself from it. For purposes of methodological cohesiveness, and because it seeks to explore the function of patterns frequently used by speakers, this study will follow just the functional approach when addressing subjectivity¹⁹.

¹⁹ Certainly, a study combining both approaches would provide invaluable results supporting the predominance of subjectivity in language. The endeavor, however, exceeds the goals and reach of the present one.

2.7.1. Subjectivity and Subjectification

Studies in several languages (Traugott 1989; Traugott 1995; Torres Cacoullos and Schwenter 2005; Company Company 2006a; Company Company 2006b) have demonstrated that constructions associated with the expression of subjectivity often result in changes in the grammar of the language; a process known as subjectification. Although subjectification shares some properties with grammaticalization (e.g. decategorization), they are different processes: while grammaticalized forms generalize and become obligatory (e.g. the grammatical meaning of ‘will’), subjective constructions do not generalize and do not necessarily acquire the status of obligatoriness.

Subjectification, just as the term subjectivity does, is understood differently depending on the linguistic approach. For Langacker (e.g. 1987, 1990, 1991) and his followers, subjectification is the shift from a syntactic subject to a speaking one. Accordingly, research looks for evidence in the emergence of constructions in which the speaker can perform an agentive role, or at least arise as the most prominent participant; as it occurs with ‘promise’ or ‘be going to’, as well as their Spanish equivalents *prometer* and *ir a*. From a more functional approach, Traugott and others (Traugott 1989; Traugott 1997; Traugott 1999; Traugott 2003; Company Company 2006a; Company Company 2006b) focus on how “meanings tend to become increasingly based in the SP/W’s [speaker / writer] subjective belief state or attitude toward what is being said and how it is being said”, with subjectification being “the mechanism whereby meanings come over time to encode and or externalise the SP/W’s perspectives and attitudes as constrained by the communicative world of the speech event rather than by the so-called ‘real-world’ characteristics of the event or situation referred to” (Traugott 2003:126–127). For said

perspective, then, subjectification involves those mechanisms by which “meanings come over time to encode or externalise the SP/W’s perspectives and attitudes as constrained by the communitive world of the speech event, rather than by the so-called ‘real-world’ characteristics of the event or situation being referred to” (Traugott 2003:126).

Besides demonstrating that subjective meanings are often encoded in grammar, diachronic studies show evidence that changes in meaning are a matter of degree, and follow a unidirectional path from non-subjective to subjective. For instance, Modern Spanish *a pesar de* ‘in spite of’ evolved from the frequent use of the noun *pesar* ‘sorrow, regret’ in constructions expressing emotional affliction by one of the participants of the event, as in (25) from Torres Cacoullos and Schwenter (2005). Although in Modern Spanish *pesar* still exists as a noun, its meaning is no longer transparent in the aforementioned subjective construction. Other examples of unidirectionality include performative uses of verbs such as ‘promise’ (Traugott 1997), the development of epistemic meanings of modals (Traugott 1989), and the subjectification of Spanish verbs such as *decir* ‘to say’ and *saber* ‘to know’ into discourse markers (Company Company 2006a).

- (25) *fue preso Daniel, a pesar del rey que lo quería anparar*
‘Daniel was imprisoned, to the regret of the king who wanted to protect him’

From a synchronic perspective, Verhagen (1995, 2005) showed that some Dutch verbs, e.g. *beloven* ‘to promise’, have different syntactic and semantic limitations when conveying subjectivity than when expressing objectivity. For Spanish, Cornillie (2004; 2005a) found a similar behaviour for the verbs *prometer* ‘to promise’ and *amenazar* ‘to threaten’. For *prometer*, he found that the subjective construal rejects animate subjects, which implies a low degree of commitment of such participant. At the same time, the preference for inanimate, non-agentive subjects is related to a major occurrence of

infinitival complements, as in (26). An animate and agentive subject, on the other hand, implies a major participation on the event; consequently, the construal is an objective one and it prefers finite complements as in (27). For *amenazar*, the patterns are similar, although the verb imposes fewer semantic rejections to the subject, which means the subject's referent can have a certain degree of commitment. Interestingly, for both verbs, the lack of animate and agentive subjects in the construals also gives the utterance a subjective reading from a functional approach. Thus, by rejecting human subjects, the subjective construal removes responsibility from the speaker, which can be considered an expression of attitude towards the information of the utterance:

(26) *El libro promete ser un éxito de ventas.* (Subjective construal)
 'The book promises to be a bestseller'

(27) *José prometió [llegar/que llegaría] a tiempo.* (Objective construal)
 "Jose promised to be on time"

Company (2006a; 2006b) argues that subjectification affects as well the syntactic properties of subjective expressions causing, specifically, cancellation of syntax and syntactic isolation. Even linguistic literature had already approached the effect of said phenomena in pragmatic particles, the author demonstrates that verbs undergoing subjectification also lose their capacity to relate to other units, subcategorize, to take complements and modifiers, among other properties. As a result, the subjective form is able to appear in parenthetical positions and to constitute a whole predication by itself. Through the analysis of five Spanish discourse markers (*¡dale!*, *¡dizque!*, *¡tate!*, *¡ándale!*, *¡sepa!*), the author concludes that the more syntactic capabilities a construction has, the less subjectivity it conveys and viceversa. She proposes as well that subjectification, although a syntactic change on its own, it is a subtype of grammaticalization since they share some properties, but differ on other ones. Syntactic cancellation and isolation can

be observed in constructions with Spanish verbs such as *decir* ‘to say’, which expresses a myriad of speech events²⁰, or *creer* ‘to believe’. When conveying subjective meanings, they can appear at the end of or inside a clause, as in (28) and (29) below:

(28) *Eso está mal, yo digo.*

“That’s wrong, I think”

(29) *El nuevo sistema de transporte, creo yo, no va a funcionar.*

“The new transportation system, I believe, it won’t work.”

This syntactic freedom, though, often comes with loss of syntactic properties. That is to say, subjective expressions can appear in different positions in the clause because they have lost their referential function. For instance, Garachana’s study on Spanish *encima* ‘on top’ (2008) illustrates loss of referential function accompanying subjectification. When performing a prepositional function, the phrase introduces locative complements, as in (30). However, when it functions as a discourse marker, it no longer accepts that type of complement, as in (31).

(30) *El libro está encima del escritorio.* (Referential, objective reading)

‘The book is on top the desk’

(31) *Encima de que llegó tarde, se enojó porque le pregunté dónde estaba.* (Subjective reading)

“On top being late; he got mad when I asked where he was”

From the previous discussion, it is clear that spoken language plays a central role when studying how subjective meanings are expressed in different kinds of interaction, and how they are encoded structurally in the language; thus more studies, diachronic and synchronic, are needed to contribute to the understanding of the interrelation between language use and structural patterns, as well as what aspects of discourse are universal and what are language-specific; particularly in situations where two or more languages

²⁰ Since its lexical meaning is not as specific as that of other verbs, *decir* is used as a synonym of *preguntar* ‘to ask’, *responder* ‘to reply, to answer’, *ordenar* ‘to command’, etc.

coexist daily. These issues constitute the main goals of this study: the expression of objectivity-subjectivity in the discourse of Spanish speakers living in a situation of contact language, and its effects on semantic, pragmatic and morphosyntactic changes. To provide a panorama of how this study will contribute to the understanding of these theoretical issues, the next section discusses relevant research on discourse, structural patterns and subjectivity in Spanish and other languages.

2.8. Discourse, Structural Patterns and Subjectivity

Studies in several languages have applied usage-based theory to examine how grammar emerges from actual discourse. From a usage-based approach, Scheibman (2002) explores the notion that language, especially in spontaneous conversation, is subjective to varying degrees, rather than mainly propositional (i.e. objective) as claimed by formalist linguistic approaches. Moreover, the author demonstrates that linguistic subjectivity can be observed in a variety of structural indications used by speakers when conveying information to their interlocutors, independently of the topic and genre of the discourse. For instance, the author finds that 1sg and 2sg predicates often occur in lexical collocations involving cognitive verbs and, more importantly, conveying subjective and pragmatic nuances: *I guess, I don't know, I mean, you know*. Subjectivity, however, does not emerge just from the use of such highly fixed epistemic collocations; Scheibman's analysis shows that in cooperative talk, speakers often employ personalizing properties to "fulfill face-maintaining functions in 2sg expressions" (2002:115), thus causing 2sg utterances to convey an interactive (i.e. empathetic) subjectivity. Furthermore, the fact that, in conversation, personal stance (i.e. 1sg utterances) does not exist independently of interaction (i.e. 2sg utterances) creates a complex relationship in which "details of

interaction are recruited for the negotiation of individual subjectivities, or the sharing of points of view” (2002:115).

Certainly, linguistic subjectivity in 1sg and 2sg utterances may not be surprising; after all, interlocutors are constantly adapting what they say, and how they say it, in order to engage the other participant in the conversation, to convince him/her about one’s arguments, or to follow sociocultural conventions, among other motifs. Nevertheless, Scheibman argues, and demonstrates, that subjectivity can be observed as well in third person predicates. Thus, results suggest that while the convergence of third person subjects with human referents, past tense, and material verbs (e.g. do, go, take, work, etc.) tends to express a high degree of propositionality, such is not the case for 3pl subjects with human, non-specific referents. Said generic subjects, highly frequent in Scheibman’s corpus, move away from the domain of mere referentiality “by anchoring the predicate to the generalizing subjects, thus producing an evaluation and not a description” (Scheibman 2002:116).

In terms of methodology, Scheibman’s results prove the importance, even necessity, of focusing on local patterns when examining how language works in actual discourse because “global distributional frequencies [...] are not typically attributable to any major structural trend” (2002:162). In terms of theoretical implications, the study makes evident that a comprehensive account of language cannot approach the semantic meaning of a lexical item as independent from its expressive and pragmatic use; otherwise, it would risk being a restrictive and misleading one. More importantly, the research raises the issue of whether it is appropriate to treat certain linguistic items (e.g. generic *they*, *I think*, etc.) as members of a formal category (e.g. person or tense), as their

functional properties do not really relate to those of other members of said category. For instance, the author claims that present and past tenses “are neither distributionally symmetrical nor contrastive members of the category tense” (2002:163) because most instances of the former occur in the expression of speakers’ point of view, and past tense mainly correlates to speakers’ reports about other actors and events. In other words, subjects and verbs are combined not arbitrarily, but heavily influenced by local contexts and previous experiences; so they should not be seen as separate constituents, as most grammatical treatments do (see also Bybee and Scheibman 1999).

Silveira’s (2007) study on subject-verb-tense patterns in Brazilian Portuguese finds that speakers tend to personalize their discourse by using specific constructions: when talking about themselves, speakers prefer 1sg predicates, usually with a cognitive verb in present tense; when talking about other referents, they tend to produce 3sg predicates with relational verbs in present tense. More importantly, the fact that, far from stating an impersonal, objective affair of states, the relational construction “projects an evaluation of a certain lexical NP, fact, or proposition” (Silveira 2007:255) provides evidence for the subjective nature of language.

For New Mexican Spanish, Valdivia’s (2011) results show a similar situation; for instance, while the verb *creer* ‘to believe / to think’ only appears in 1sg-present-affirmative constructions (i.e. *yo creo*) ‘I think’, usually at the end of a utterance, as in (32) below, *decir* ‘to say’ appears mostly in 3sg-preterit and 3pl-present combinations, as in (33) and (34):

- (32) *Yo no sé de dónde traerán capulín aquí.*
Yo no sé dónde se dará.
Muy lejos yo creo. (10-3A2: 175)
 ‘I don’t know where they bring capulin from. I don’t know where it may grow. Far away, **I think.**’

- (33) *Voy a matarlo,*
*dizque **dijo**.* (142-3B2: 197)
 ‘I’m going to kill him, *dizque* **he said**.’
- (34) *It’s a beautiful –*
es un libro muy bonito.
*eh **dicen**.* (88-1A3: 77)
 ‘It’s a beautiful—it’s a very beautiful book. **They say**.’

This kind of correlation between subject, verbal types and tenses proves that much of the language that speakers produce in interaction must not be understood in terms of preconceived rules. On the contrary, certain pre-constructed patterns constitute the core of speakers’ grammatical knowledge. Moreover, considering the differences, typological or of other sorts, between American English, Brazilian Portuguese and New Mexican Spanish, the alignment of Silveira’s and Valdivia’s results with those from Scheibman’s (2001; 2002) provides clear evidence for grammar being rooted in language use, instead of the opposite, which causes emergent constructions conveying speakers’ attitudes and point-of-view to be similar across languages.

The recurrent use of prefabs and collocations, and the development of epistemic meanings, are observable as well in studies focusing on particular verbal classes. Cognitive verbs, in particular, have been extensively studied not only because the speaker often appears as both the syntactic and the elocutionary subject of the utterance, e.g. *pienso* ‘I think’, but also because forming an opinion is associated with cognitive processes (Aijón Oliva and Serrano 2010; Bentivoglio and Weber 1999; Weber and Bentivoglio 1991; Tao 2001; Tao 2003; Travis 2006a; Valdivia 2013). A frequent finding in these studies is that cognitive verbs do not really constitute a homogenous class; they differ from each other in the type of subject-tense combinations they frequently appear, in the polarity of the sentences, etc. Furthermore, more than one construction, each one

conveying a different meaning, may arise from the same verb due to the confluence of syntactic and pragmatic factors. For instance, in a study of cognitive verbs in conversational Spanish, Travis (2006a) finds that not only a few verbs (*saber* ‘to know’, *creer* ‘to believe’, *pensar* ‘to think’, *imaginarse* ‘to imagine / to figure’, *acordarse* ‘to remember’) account for the vast majority of occurrences in the corpus, but, for each one of said verbs, most instances correspond to one or two constructions performing epistemic functions (e.g. *yo no sé* ‘I don’t know’, *¿sabes qué?* ‘you know what?’). As expected, these constructions lack many of the syntactic properties associated with the propositional function of the verb.

Valdivia (2013; 2015) finds similar patterns for New Mexican Spanish in the sense that frequent cognitive constructions conveying subjectivity involve just a few verbs (*creer* ‘to believe’, *saber* ‘to think’, *entender* ‘to understand’, *pensar* ‘to think’), in specific subject-tense combinations. The fact that constructions emerging from this dialect of Spanish show both similarities and differences from those in Travis’ study provides strong evidence for the role of interaction in the development of grammar; on the one hand, similarity between two different and separated dialects proves that speakers employ the mechanisms of the language to achieve their communicative goals, on the other hand, structural differences demonstrate the impact of factors specific to each communicative situation and linguistic community.

Aijón and Serrano’s (2010) study on the implications of using subject pronouns in Spanish 1sg constructions provides more evidence for the correlation between structure and meaning. Using oral corpora from two Spanish speaking communities (Salamanca and Canary Islands) and two different registers (media and conversation), the authors

seek to demonstrate that expressing or omitting *yo* 'I' in *creo* 'I believe' utterances is closely linked to discourse and pragmatic factors, such as the speaker's prominence. Their results show that pronominal subject (predominantly in preverbal) is more frequent (75%) than the unexpressed subject, which would imply that, for 1sg, unexpressed subject is marked rather than unmarked. Nevertheless, a detailed analysis of the data demonstrates that *yo creo* functions as an epistemic marker conveying speaker's assertion; it emphasizes the speaker's stance regarding the argument he or she is introducing, as well as his/her importance as an opinionated person, particularly when several people are expressing their opinion. Considering how important this function is, particularly in discussions, it is not surprising that *yo creo* is much more frequent than *creo*. For the latter construction, the low frequency does not allow for conclusive results regarding its function in the discursive. The authors notice, however, that *no creo* 'I don't believe' tends to appear without a pronoun even if it mitigates or relativizes the veracity and importance of the speaker's assertion.

The subjectification of constructions conveying speakers' point of view is attested too in verbs from semantic classes other than the cognitive one. Travis (2005) finds that 25% of the occurrences of *decir* in a corpus of conversational Spanish convey epistemic nuances, which indicates that, in interaction, *decir* is not a mere verb of saying, but it performs several other functions. More importantly for the matter of this study, there is a clear correlation between particular constructions (e.g. *dizque*, *mejor dicho* 'in other words, or rather' *yo digo* 'I say', *lo que yo digo* 'what I say', *¿qué te digo?* 'What can I tell you?'), and the expression of epistemic values. For instance, the author finds that speakers use *dizque* (from *dizen que* 'they say') as an interjection (35), and to introduce

nominal (36)-(37), and prepositional phrases, even if, according to Royal Spanish Academy (RAE), the word is an adverb meaning ‘apparently’ or ‘presumably’. Although Travis merely describes the construction’s meaning as purely subjective, one may propose that said subjective meaning conveys the speaker’s disbelief, irony or sarcasm towards the utterance, or a segment of it, as in (36) and (37) below. In fact, said nuance is observable too when the form modifies an adjective (38) or a clause; that is to say, when it follows the syntactic function of a canonical adverb. All of the examples below are from the referred study. The development of *dizque* into a discourse marker and its syntactic effects are examined in further detail in Travis (2006b), Company (2006a; 2006b), Olbertz (2007), and Valdivia (2010b).

- (35) *Me hizo quedar ahí, yo, dizque. Ay=.*
‘She made me wait there; I, dizque. Ay!’
- (36) *tengo uno, tengo dizque el capa, o algo así,*
‘I have one; I have a so-called cape, or something like that.’
- (37) *Y lo compramos dizque por ella.*
‘And we bought it supposedly because of her.’
- (38) *Unas peras dizque importadas, eso dicen.*
‘Some supposedly imported pears; that’s what they say.’

Following *dizque* in frequency of occurrence, *mejor dicho* ‘that is to say / or rather’ functions as an emphatic device correcting or explaining something just said. Structurally, the construction lacks morpho-syntactic variability and it is syntactically independent. In terms of prosody, it appears in its own intonation unit. A third construction *yo digo* ‘I say’ provides evidence as well for the gradual nature of subjectification; like *mejor dicho* and *dizque*, it does not show variability, but, unlike said markers, it has not reached complete syntactic nor prosodic independence.

Continuing with the emergence of non-propositional constructions with *decir* ‘to say’, Valdivia (2014) explores the use of the verb in interviews from NMCOS to identify what subject-tense combinations are developing prefab status, as well as to determine what morphosyntactic features play an essential role in said development. Results demonstrate that not only morphosyntactic features, but also prosodic ones provide a set of operational measures to determine whether a particular subject-tense combination constitutes a prefab, a collocation or, rather, a general schematic construction. For instance, the verb seems to maintain its propositional function in constructions reporting an event (e.g. *dije* ‘I said’ or *dijo* ‘he-she said’); nevertheless, the fact that said constructions often appear in the middle or after the reported event suggests that, rather than merely introducing information, they are being used as fillers, discourse planning devices, and other types of discourse markers. As it will be discussed in Chapter 6, the syntactic independence allowing these items to move to an internal, or even final position, while developing new semantic and pragmatic nuances is strongly intertwined with their occurrence in their own intonation unit.

The study also provides evidence regarding the role of established prefabs in the development of new ones. For instance, the emergence of *le(s) decían X* ‘they called it/them X’ conveying certain degree of evidentiality is largely due to its syntactic, semantic and pragmatic relation to the more frequent evidential construction *le(s) dicen X* ‘they call it/them X’. This latter construction, in turn, likely arose from the use of *dicen* ‘they say’ conveying evidentiality or epistemicity.

2.9. Summary

This chapter has presented the concepts and assumptions constituting the theoretical framework for this research: usage-based theory. It discusses as well the term construction referring to form-meaning pairings, its central role in various methodological approaches, the type of units it may include, and how they can be placed in a continuum from schematic to fixed units. More importantly, the chapter unifies different interpretations of the term ‘construction’ and provides the definition to be used in this research. Also, the chapter illustrates how subjectivity and subjectification are defined and studied by cognitive (e.g. Langacker’s) and discourse-based (e.g. Traugott’s) approaches to language. Lastly, the chapter discusses studies that have approached the two intertwined notions to be explored in this research: a) structural patterns arising from discourse form a continuum ranging from schematic to idiomatic constructions, and b) in actual discourse, particularly interactional and spontaneous genre, constructions convey subjectivity more often than objectivity.

Chapter 3. Data and Methodology

3.1. Data

Data for the research come from Bills & Vigil's (2008) *New Mexico and Colorado Spanish Survey* (NMCOS), a project that started in 1991²¹ with two main goals: 1) to create, at the University of New Mexico, a public archive containing the recordings, paper documents and other materials from the project, an archive that presents “a realistic sampling of New Mexican Spanish as spoken at the end of the twentieth century” (Bills and Vigil 2008:21); and 2) to produce a linguistic atlas²² documenting lexical variation in the Spanish of the region. To collect data pertinent to both goals, a three-part interview was designed to 1) find out about consultants’ background and language use, 2) obtain information about language proficiency, and 3) procure specific linguistic elicitation and free conversation.

Regarding the first goal, it was essential that the data represented regional and social diversity, that illustrated different performance styles, and that described speakers’ individual and cultural experiences. To collect such a corpus, Bills and Vigil divided the area of research into twelve sections from three regions: Central, Eastern and Western (see map 1 in Bills and Vigil (2008:23) for further details). A total of 357 consultants, ages 15 to 96, were selected for the NMCOS based on three criteria: consultants had to be adults, to have been born and raised in the region they represented, and to have learned Spanish, at least to some extent, in their homes during childhood. In addition to these basic criteria, researchers tried to create a group of informants that included equal

²¹ The project received a grant from the National Endowment for the Humanities for three and a half years.

²² Bills, Garland & Neddy Vigil. 2008. *The Spanish of New Mexico and Southern Colorado. A Linguistic Atlas*. Albuquerque, NM. University of New Mexico Press.

numbers of male and female, representing three age groups (18-39, 40-60, 60+).

Nevertheless, since “interviews were arranged informally through the project team’s networks of contacts [...] there were deviations from the desiderata in coverage of localities as well as in the age and sex distributions” (Bills and Vigil 2008:25).

For this dissertation, data were extracted from 18 interviews available for research in the Department of Spanish and Portuguese at the University of New Mexico. Except for one, in which there were two interviewees, interviews were one-on-one. The consultants are 11 women and 8 men between 45 and 96 years of age; with time of formal instruction ranging from minimal schooling to 17 years (see Table 2).

Table 2. Consultants' Demographic Information				
Interview no.	Gender	Age at time of interview	Formal instruction	Former or current occupation
219	F	96	3 years	weaver
010	F	88	two winters	housekeeper
020	F	82	book 5	cook, sheep shearing
004	F	80	3 years	rancher
142	F	80	16 years	teacher
076	F	74	book 8	bus driver
318	F	72	book 10	? ²³
147	F	66	17 years	teacher
117	F	62	8 years	cook
047	F	54	8 years	housewife
311.2	F	?	10 years	?
311	M	90	book 7	highway dept
270	M	87	12 years	?
088	M	45	12 years	truck driver
144	M	83	16 years	pastor
214	M	76	book 4	truck driver
156	M	73	10 years	?
190	M	71	12 years	mechanic/serv. manager
102	M	48	10+ yrs	wood artist

Data from interviewers have been included in the study not only because they are members of the community to be studied, but also, and specially, because their speech

²³ Occupation unknown

may show discourse functions only observable in their interaction with the interviewees' speech; for instance, the common use of questions seeking to elicit speech from the interviewee.

Data were transcribed according to Du Bois et al's (1993) system, in which each line represents an intonation unit (IU), defined as a speech segment uttered under one intonation contour. The system also identifies prosodic cues, broken speech, turns of speech, overlaps between features, pauses, and so on. Even though intonation units do not directly correspond to syntactic units such as phrases or clauses, they do appear to have a role in the way grammar is organized (Chafe 1994; Ono and Thompson 1995) and thus are incorporated into the analysis.

Using the software Monoconc (Barlow 2002), a total of 2134 instances of both cognitive or speech verbs were extracted from approximately 8 hours, 10 minutes of audio, corresponding to 87,800 words and 24,544 intonation units. Because the goal of the research implies examining the different ways speakers use both verbs of cognition and verbs of speech, all realizations of said verbs²⁴ were considered in the analysis independently of their form (i.e. personal or non-personal) and their syntactic function (e.g. verb, noun, adverb, etc.), as shown in (39) and (40) below, where an infinitive functions as a noun and a gerund functions as an adverb, respectively.

- (39) *depende con quién.*
*pero me gusta más **hablar** español.* (102-1A1: 133)
'It depends with whom. But I like to **speak** Spanish better.'

²⁴ Semantic classification of the verbs is based on (Halliday 1994). See chapter 4 for more details.

- (40) *pues,*
se fueron,
***pensando** que no iba a encontrar a ninguna persona que sacara --*
que cortara el zacate. (214-1B3: 318)
 ‘So, they went away **thinking** I wasn’t going to find anybody to pull – to
 mow the loan.’

With the exception of passive voice constructions, past participles functioning as adjectives (e.g. *la lengua hablada* ‘the spoken language’) were excluded too because of the difficulty in drawing the line between whether they are verbs used as an adjectives and whether they are adjectives indeed. Clauses excluded from the corpus were those in which the expression of the verb is truncated²⁵, as in (41), and those in which the verb or its arguments could not be accurately coded because the speech was unclear²⁶ (speech between <X X>). Originally, clauses with truncated complements, as (42), were also excluded; however, they were later included because of the effects said truncation might have in the way speakers construct and adapt their discourse. For instance, in (42) the speaker truncates the first utterance of *dijo* ‘she said’, and then restarts her idea specifying who the subject of *dijo* is (*una amiga mía* ‘a friend of mine’). The truncation and reformulation may be related, then, to the speaker’s need to include the phrase *se me hace* ‘it seems to me / I think’ between the subject and the verb, phrase with which the she introduces a nuance of doubt into the whole statement.

- (41) Example with truncated verb
Cuando se fueron,
*..(2.0) **dijo**,*
<X hallo el X>-- (144-3B2: 341)
 ‘When they left, she said, <X I find the X>²⁷--
- (42) *me **dijo** --*
.. una amiga mía se me hace que me dijo, (318-1A2: 592)
 ‘(She) told me, a friend of mine told me, I think,’

²⁵ Indicated with -- in Du Bois et al’s (1993) transcription method.

²⁶ Marked as <X X> in the transcription.

²⁷ For the translation of the examples, punctuation will be used in its traditional way.

3.2. Methodology

Once extracted and coded for the features explained in section 3.3 below, two levels of analysis were conducted. First, a global analysis approaches each verbal class as a whole and heterogeneous group, looking for differences, or similarities, in the way speakers use these types of verbs. Thus, it compares whether one verbal class is more recurrent than the other (i.e. token frequency), and how said frequency relates to the class' membership (i.e. type frequency). This general analysis seeks, as well, to identify structural tendencies of each class in terms of subject, tense, polarity and argument structure. Global patterns will be later contrasted to constructions emerging from local contexts to determine how subjectivity and subjectification relate to verbal classes as a whole, and to specific items. Results for this examination and their implications are presented and discussed in chapter 4 below.

The second level of examination focuses on subject-tense combinations for specific items in each verbal class. This local analysis seeks to demonstrate that recurrent use of a verb in specific morphosyntactic combinations may result in different degrees of structural fixation, and how such fixation may cause a particular combination to eventually develop semantic nuances and even functions associated with subjectivity. Based on said goals, this local analysis focuses on those specific collocations (i.e. subject-verb combinations) that show significant token frequency either in the whole data, within their subject-tense groups or within its type group. That is to say, a collocation may have not a very high token frequency regarding the whole data, but it may be among the most frequent items of the verb in a particular subject-tense group or even the most frequent realization of that particular verb. To compare global patterns and

local patterns, the local analysis approaches the same interrelations that the global analysis does; however, to achieve a better understanding of how specific constructions emerge from discourse, the local analysis also examines the interrelation between linguistic structures and features such as prosody (e.g. distribution of intonations units), discourse flow (e.g. turn changes), and others explained in section 3.3 below.

3.3. Coding

Considering this research focuses on the use of verbs of cognition and speech, and particularly on frequent subject-tense collocations within each verbal class, it might seem logical to use the clause (i.e. the syntactic structure formed by a verb and its arguments) as the basic unit of analysis. Nevertheless, said unit resulted not to be the best for the coding and, specially, for the analysis for several reasons. First, because clauses often are arguments of another clause, using the same verb to refer to the main predicate and to its argument might be confusing when discussing patterns, particularly when both clauses include items of the same verb, as in (43).

- (43) *tú fuistes la que **dijites**,*
*le **dijo** de la puerta,*
*le **dijo**. (10-8A: 445)*
‘You were the one who **told** that, she **told** her from the door, she **told** her.’

Second, due to the complexity of oral speech, sometimes it was difficult to decide, from a structural point of view, where the scope of a verb ended, or whether a verb actually formed a syntactic unit with the information it introduces or refers to. For instance, in (44) the item *me dice* ‘he tells me’ in line 4 may be interpreted as referring to the quote in line 3 or just to *grandma* in line 5. Also, in line 6, *me decía* ‘he called me / he told me’ may be introducing the quote in line 7, or clarifying that the speaker’s grandson used to call her *grandma*, not *mamá* ‘mom’, as originally stated in line 3.

- (44) 1 *L: Yo quiero a este abogado que está.*
 2 *.. Y me dice,*
 3 *<VOX bue=no mamá VOX>,*
 4 ***me dice,***
 5 *grandma,*
 6 ***me decía,***
 7 *Voy a ver qué hace mi mama por mí. (219-1B1: 514)*
 ‘I want the lawyer we have now. And he tells me: “Well mom”, **he tells me**, grandma, **he told me**, “I’ll see what my mom can do for me” or ‘... **he tells me**: “grandma”, **he called me**: “I’ll see what my mom can do for me.”’

The complexity of the discourse is also illustrated in (45) below, in which *yo creo* ‘I think’ in line 5 appears as a parenthetical unit commenting on one segment (lines 4 and 6) of the whole utterance.

- (45) 1 *tuvieron bilingual.*
 2 *and ah --*
 3 *y luego eh= --*
 4 *hace como dos años,*
 5 *yo creo,*
 6 *o tres,*
 7 *que lo quitaron.*
 8 *ya ese program lo quitaron aquí que sepa -- que sé yo. (117-1A2: 392)*
 ‘... they had a bilingual program. And... and then... like two years ago, **I think**, or three, it way suspended. The program was suspended. As far as [I] know... I know.’

Another candidate for unit of analysis was the intonation unit, which besides being the basic entity of transcription, is one of the features explored in the study as playing an important role in the development of subjectivity and subjectification. Nevertheless, it was discarded as the basic unit of analysis because not all IU’s include a verb belonging to the classes to study, or do not include a verb at all. Both situations are illustrated by line 6 and line 7 of example (45) above.

Due to these issues, and because the goal of this research is to demonstrate how structural patterns emerge from the use of the verbs in actual discourse (i.e. beyond the sentence), I decided to follow Scheibman’s (2001; 2002) methodology and to use UTTERANCE, as the basic unit in my analysis. Scheibman defines the term as a “chunk of

language produced by real speakers”, and she uses it only for units with verbs in a finite form (2002:25). Unlike her, however, I also use the term for units including non-finite forms of the verbs under analysis²⁸. In addition to fitting better the goals of the research, using utterance as the unit of analysis allows a clearer and more precise comparison between its results and those from previous research on structural patterns (Scheibman 2001; Scheibman 2002; Travis 2006a; Silveira 2007; Valdivia 2014).

Seeking to establish an objective measurement for the emergence of structural patterns, and its position in a construction’s continuum (which includes the development of properties associated to discourse and pragmatic markers), several morphological, syntactic, and prosodic features were operationalized and used to code each utterance in the corpus.

3.3.1. Subject’s Form (Person and Number)

The category includes values for first, second and third person both singular and plural; for second person singular, it distinguishes between the form for *tú* (familiar ‘you’) and for *usted* (formal ‘you’). From now on, these subjects will be referred as 1sg, 2sg-*tú*, 2sg-*Ud*, 3sg, 1pl, 2pl and 3pl. The coding is based primarily on the deictic reference of the utterance’s subject, if any, and then in the verb’s declination. Thus, cases in which there is a conflict between the form of an expressed subject and the grammatical person it refers to, the latter was favored. I applied this criterion mainly because, in New

²⁸ In her work, Scheibman uses utterance just for units with finite verbs mainly because she examines predicates with all semantic classes; thus, including non-finite forms represented theoretical issues that were beyond the goal of her research. In my case, because one of my goal is, precisely, to examine how speakers use cognitive and speech verbs, and how such usage impacts grammar, the term utterance include both personal and non-personal realizations of said verbs. Because of so, clauses such as *ella sabe hablar español* ‘she knows how to speak Spanish’ were counted and coded as two utterances, one focusing on *saber* ‘to know’ and the other one focusing on *hablar* ‘to speak’.

Mexican Spanish, the form *ha* has been documented not only for third person singular (46), but also for first person singular²⁹ as in (47) below (Bills and Vigil 2008:146).

- (46) *B: Buena suerte [2este señor2].*
O: [2Le ha dicho2]?
Z: yeah. (311-4A1: 177)
 ‘B: He’s lucky, this man. O: **Has he told** you. Z: Yeah’
- (47) *P: Y ellos te dicen,*
.. Pa' ellos no es nada.
ellos no más están platicando contigo.
A: .. [@ @ @]
P: [Eso me ha fijado] mucho. (88-1A2: 356)
 ‘P: And they say to you. For them is nothing. They are just chatting with you.
 A: @ @ @. P: **I have noticed** that a lot.’

Subjects expressed by nominalized elements (e.g. adjectives, clauses, etc.) are coded as third person singular or plural, as in (48) in (49) respectively, according to the agreement they establish with the verb.

- (48) *D: Este señor .. qué es? ((MOSTRANDO ILUSTRACIONES))*
usted <X qué le parece X>.
*S: **Parece que es un doctor** no? (102-2B1:407)*
 ‘D: This man, what does he do? What do you think? S: **It seems he is a doctor**, no?’
- (49) *A: Así que --*
Sí.
y hablaba eh=
-- con los trabajadores en español?
o en inglés?
R: En inglés.
Todos los más hablaban en inglés. (20-1A1: 675)
 A: ‘So – Yes. And, with the workers, did you use to speak in Spanish or in English?’ R: ‘In English. All the others spoke in English.’

When the verb is part of a modal or a periphrastic construction, the subject is coded based on the auxiliary verb; so, the subject of (50) is coded as 2sg-*Ud*, and the subject of (51) as 1pl. The same criterion was applied when one of the verbs under study appears in a non-personal form, either as a complement or a modifier, providing that it is

²⁹ The canonical form in present perfect for 1sg is *he*, ‘I have’ as in *Me he fijado* ‘I have noticed’.

co-referential to the main predicate, as in (52) to (54). The motive for such a decision is the interest in observing not just the behavior of the verb as an independent unit, but as a part of a construction as a whole. Whether the verb is part of the predicate's nucleus, or whether it performs a non-verbal function (e.g. noun, adverb, etc.) is indicated in another category (see section 3.3.7)

(50) Subject = 2sg-Ud

A: *Y me puede= platicar usted de= dónde nació=,
... en qué año nació,
dónde creció?* (318-1A2:14)

'A: And, **can you tell** me where were you born? In what year, where you grew up.'

(51) Subject = 1pl

W: *... Y luego=,
que?
de qué estábamos hablando?* (190-3B1:45)

'W: And then... What? What **were we talking** about?'

(52) Subject=1sg

A: *Y cuándo aprendió a [hablar inglés]?*

O: *[(H)] Ah=,
pos,
.. **aprendí a hablar inglés** --
%% antes de entrar a la escuela.
Porque teníamos muchos ah --
muchos ahm --
ahm --*

.. Neighbors. (117-1A2:112)

'A: And when did you learn to speak English? O: Well, **I learned to speak** English before starting school. Because we had a lot of neighbors.'

(53) Subject = 1sg

Y: *Y todos los días,
.. cuando caminaba,
nos dijía historias.*

A: *[Mhm].*

Y: *[en] español.
muy bonito.
(H) Y cómo **quisiera yo acordarme** esas historias para escribirlas.
más que las escribiera en inglés.*

A: *[Mhm].*

Y: *[Pero sí],
unas historias muy bonitas nos dijía.* (147-1A2:655)

'Y: And everyday, while she'd go for a walk, she told us stories. A:Mhm. Y: In Spanish. It was very pretty. **If only I could remember** those stories to

write them down! A: Mhm. Y: But yeah. Some pretty tales, she used to tell us.'

(54) Subject=3sg

X: *ésa fue la que se creció un --
que= se crió con ella.*

*.. y llegaba **hablándole** en inglés,
y la agarraba de las orejas ella,
<VOX no no no oye.*

aquí me vienes a hablar n=o güirigüiri.

aquí me vas a hablar en mexicano. (47-1A2:233)

'X: She was the one who grew up a... the one who was raised by her. And she [the interviewee's daughter] would come home **speaking** in English and she [the interviewee's mom] would grab her from the ears while telling her: "here you won't babble; here you have to speak in Mexican.'

Finally the category includes the value 'non-applicable' (NA) for cases in which the cognitive or speech verb does not have a subject, as in the construction *cómo decir* 'how to say it' illustrated in (55); because it appears in an impersonal construction, as in (56) and (57); because the non-finite verb is not the construction's main verb³⁰ neither co-referential to the main predicate, as is in (58)-(60) below; or because it functions as another verb's subject, as in (61).

(55) T: *hace varios años,
(H) que no tenemos ni una lluvia fuerte,
ni una --
ni una nieve .. alta tampoco.*

F: *Hm.*

T: *polvitos no más.*

F: *[Yeah].*

T: *[you see].
y así ya hace varios años.
No sé porque.*

F: *yeah.*

*El ahm --
.. (TSK) y **cómo decir** el eh --
y más antes sí?*

T: *Pues eh,
me acuerdo,
XX han venido s- años secos también,
no más que se acuerda uno de algún año muy nevoso ... (270-1B2:317)*

'T: We haven't had a storm, not a single one, neither a blizzard, for years. F: Hm. T: Just flurry. F: Yeah. T: You see. It has been like that for years. I

³⁰ That is to say, it is not part of a modal or periphrastic construction.

don't know why. F. Yeah. The... **what I want to say?** The... And before you did? T: Well, I remember... we have had droughts too, but one usually remembers just years with a lot of snow'

- (56) R: *[Todos] le hablan en inglés.*
todos le [2hablan en inglés2].
 A: *[2Pues, ahí está2].*
 R: *Mhm.*
 A: ***Hay que hablarle** .. más en [español].* (20-5A3:265)
 'R: Everybody speaks to her in English. Everybody speaks to her in English.
 A: Well, that's why [she cannot speak Spanish]. R: Mhm. A: She **must be spoken** to in Spanish.'
- (57) R: *[(H) No],*
*yo no más pienso que ellos .. quieren que se .. **hable** no más en inglés.* (20-5A3:71)
 'R: No. I think that they want only English **to be spoken**.'
- (58) *no recuerdo si le mandían una natita,*
notita,
o mandaría a algún amigo o= pariente,
*(H) **a decirle** al –*
al ah –
al padre del –
del novio que –
que –
.. sí .. o no. (144:3B2:245)
 'I don't remember whether one would send the fiancé's father a little note, or whether one would send a friend or relative **to tell** the fiancé's father yes or not.'
- (59) X: *y mi mama no le gustaba que hablaran inglés.*
*A mí no me hagas **hablar** a güirigüiri aquí,*
ni güirigüí.
 A: *@@@@*
 X: *a mí me van a hablar mexicano.* (47-4A2: 38)
 'X: My mom didn't like to be spoken in English. "Don't make **speak** 'güirigüí' here". A: (Laughs). X: When you talk to me, you have to speak Mexican.'
- (60) *ya oí yo, **oí dijir** que tenían bilingual.* (117-1A2:357).
 'I heard so. I heard someone **saying** they had a bilingual program.'
- (61) *y él era una persona muy inteligente.*
.. Muy trabajadora,
Y=, le gustaba mucho= --
***platicar** con la gente=,* (20-1A1:417)
 '...and he was a very intelligent person. Very hardworking. And she really liked **to chat** with people'

3.3.2. Subject's Realization

The category distinguishes between unexpressed and expressed subject, as well as the morphological realization of the latter. Thus, the included values are unexpressed subject (\emptyset), full nominal phrase (FNP), personal pronoun (i.e. *yo* 'I', *tú* 'you', *él* 'he', *ella* 'she', etc.), indefinite pronouns (e.g. *algunos* 'some people', *otros* 'other people', *nadie* 'nobody', etc.), other pronouns (e.g. demonstrative, interrogative, relative pronouns, etc.), clause, and other subject. Each kind of subject is illustrated in (62) - (68) below.

- (62) Subject = unexpressed (\emptyset)
*Por qué (\emptyset) **no nos dices** también cómo las .. cocinas, estas comidas. (4-2B2: 181)*
'Why **don't you tell us** too how you cook them? These foods'
- (63) Subject= FNP
*U: porque **los mexicanos** no hablamos el= el real Spanish,*
E: Aha.
U: hablamos todo mixed,
no? (214-1A1:557)
'U: because **Mexicans, we** don't speak the real Spanish. E: Aha. U: We speak mixing Spanish and English, no?'
- (64) Subject = Personal pronoun
*(H) Y por eso digo **yo** que,*
if you do that, (190-3B2: 177)
'That's why **I** say that, if you do that,...'
- (65) Subject = indefinite pronoun
*y **otros** dicen tina. (10-3A2: 34)*
'and **other people** call it washtub'
- (66) Subject = relative pronoun
había mucha gente que --
***que** decía muy buenas historias. (142-3B2: 124)*
'there were a lot of people who... **who** told very good stories'
- (67) Subject = clause
[pues],
*Sabe usted que a mí se me hace **que cinco dólares,***
por cortar unos zacate es muy bueno=. (214-1B3: 435)
'Well, it seems to me that five dollars for pulling some grass is very good money, you know?'

- (68) Subject = other (adverb)

*Ésa la sacaron,
la mataron,
la tiraron de por allá en Pecos,
.. y no se sabe **nada**.*

...No se puede probar nada todavía. (219-1B1: 609)

‘Her, she took her out. They killed her, they threw her over there by Pecos.
And nobody knows anything. Nobody can probe anything yet.’

3.3.3. Subject’s Position

In addition to their morphological realization, expressed subjects were coded for their position in relation to the verb, i.e., as preverbal, (63) above, postverbal, as in (69) below, and middle position. The latter code was included to account for verbal constructions in which the subject appears between the main predicate and a verb in a non-finite form, as in (70). Rather than coding these cases just based on the finite verb, I decided to include the code because the ‘intra-verbal’ position could provide information regarding the degree of unit-hood of both verbs.

- (69) *Le digo yo a él que no tiene que ir uno a gamble,* (311-4A1: 290)

‘I tell him that it’s not necessary that one gambles.’

- (70) *Y: Y todos los días,
.. cuando caminaba,
nos dijía historias.*

A: [Mhm].

*Y: [en] español.
muy bonito.*

- (H) Y cómo quisiera yo acordarme esas historias para escribirlas.* (147-1A2: 155)

‘Y: And everyday, while she’d go for a walk, she told us stories. A:Mhm. Y: In Spanish. It was very pretty. If only I could remember those stories to write them down! A: Mhm. Y: But yeah. Some pretty tales, she used to tell us.’

By measuring different features of subject realization, the three previous categories test whether specific subject-verb combinations are losing morpho-syntactic variability, an indicator that a structural pattern is becoming a construction of some sort. Furthermore, determining the degree of variability is important for this research because structural fixation often has semantic-pragmatic implications. For instance, Torres

Cacoullos and Travis (2011:249) suggest that post-verbal *yo* in New Mexican Spanish occurs frequently in fixed discourse formulae marking contrast (e.g. *digo yo* ‘Say-I’). Also, in a study on the use *creo* ‘I believe / I think’ in Peninsular and Canarian Spanish, Aijón and Serrano (2010) observe that although both *(yo) no creo* ‘I don’t think so’ and *(yo) creo* ‘I think’ function as discourse markers, the former tend to appear with unexpressed subjects (71), while the latter prefer pronominal subjects (72). Both examples are from Aijón and Serrano’s paper.

- (71) *No lo creo, no lo creo, el gobierno ha mantenido acuerdos por lo tanto **no creo** que sea imprescindible la colaboración con CC.*
‘I don’t think so, I don’t think so. The Government has reached agreements, so **I don’t think** collaboration with CC is indispensable.’
- (72) ***Yo creo** que los constructores han ido hasta donde han podido ir. Eso es realmente lo que ha pasado, **yo creo**.*
‘**I think** the builders have done everything they could. That is really what has happened, **I think**.’

3.3.4. Subject’s Referentiality

The need to consider the referentiality of the subject as a factor of possible influence on the emergence of structural patterns arose early in the pilot study, mainly because of the diversity of forms in which third person subjects can be realized. Thus, while many personal pronouns (e.g. *yo* ‘I’, *tú* ‘you’, *él* ‘he’, *ella* ‘she’ etc.), and nominal phrases such as *mis hijos* ‘my kids’ or *los míos* ‘my ones’ refer to specific individuals, pronouns such as *uno* ‘one’, *algunos* ‘some people’, and phrases such as *la gente* ‘people’, refer to non-specific individuals, either because the speaker do not know who the referents are or because s/he considers their identity is not essential to the event being narrated. Furthermore, unexpressed subjects are used in Spanish not only when their referent is known, but also as a mechanism to reduce the agency of the clause by presenting the referent of the subject as non-specific. On this matter, in the pilot study,

speakers often used Ø-*dicen* ‘they say’ referring to a non-specific group of people; by doing so, the utterance expresses a sense of evidentiality:

- (73) *[No sé cómo se] llamarán.*
Te- --
.. *Tepocates les **dicen** en español.* (214-1A2:37)
‘I don’t know how they are called. Tepocates, **they call** them in Spanish.’

In the pilot study for this dissertation, the distinction between specific and non-specific referents was included in the values for subject’s form³¹; nevertheless, when more verbs were brought into the analysis, it became necessary to create a separate category for referentiality in order to avoid having a large amount of labels for subject’s form, and, specially, mixing two features into one value. Referentiality of subject, then, includes the values: speech act participant (SAP), specific SAP plus other referents (SAP+), non-speech act referents (non-SAP), inanimate entities, generic, and events.

SAP indicates the subject is an actual participant in the conversation; that is to say, the interviewee and the interviewer, which in the data always are first and second person as shown in (74) and (75) below. SAP+ indicates plural subjects referring to the interviewee or interviewer, plus other people who are not participating in the conversation, as in (76).

- (74) Referentiality = SAP
A: *Cómo se hubiera sentido usted si le hubieran enseñado puro inglés?*
R: .. **Yo creo** que no me hubiera sentido muy bien porque,
... *A mí me parece que dejar uno su idiomia del todo,*
Sí es bueno aprender uno a hablar inglés. (20-5A3:216)
‘A: How would you have felt if only English had been taught to you? R: **I think** I wouldn’t have felt very well because... I think that completely ignoring one’s own language... Yes, it’s good to learn to speak English.’

³¹ For instance, there were different values for third person subjects depending on whether they refer to a specific (3s, 3p) or to a generic entity (3s-gen, 3p-gen).

- (75) A: [*Usted*] **cree** que esa ley es .. justa?
 K: .. [No].
 A: [Nomás] porque es ley?
 K: No=,
 le --
 no.
 no creo porque,
 pues,
 ésa es nuestra idomia.
 por qué vamos a dejarla. (318-1A3:563)
 ‘A: **Do you think** that is a fair law? K: No. A: Just because it is the law? K:
 No. I don’t think so because, well, that is our language. Why should we stop
 using it?’
- (76) Referentiality = SAP+
 A: *Qué hablaban en la casa?*
 X: *Mexicano.*
 A: *Toda --*
 .. *Toda la vida?*
con sus hermanos?
 X: *Con mi hermano y mamá.*
Oh yeah,
*a ella no le gustaba que **habláranos** así en inglés.* (47-1A2:205)
 ‘A: What language do you speak at home? X: Mexican. A: All... All of your
 life? With your sibilings? X: With my brother and my mom. Oh yeah. She
 didn’t like that **we spoke** like that, in English’

Non-SAP’s referents are specific, animate, entities that do not participate in the conversation, as in (77). Although they usually are third person subjects, they also can be first and second person subjects whose referent do not match the participants in the actual conversation; for instance, interlocutors in a speech event referred to by the interviewee, as in (78), where the speaker uses a verb in 2sg referring to himself because, in the event he is narrating, he was the hearer.

- (77) Referentiality = non-SAP
 C: *Y no **hablaba tu papá** de los antepasados?* (144-4B2:357)
 ‘And your dad, did he speak to you about your ancestors?’

- (78) *When I was growing up,*
que me --
(H) que me dijieran pues,
*no **hables** en mexicano,*
no pues.
ahora sí me han dicho. (318-1A3:275)
 ‘When I was growing up I wasn’t told not to speak in Mexican. Nowadays, I have been told that.’

Inanimate referents are those indicating specific beings other than people and animals, as in (79). Generic subjects designate a class of referents, as in (80), or a subject that does not refer to anybody in particular, as in (81).

- (79) Referentiality = inanimate
*yo tengo un libro **que dice** todo. (76-1B2: 264)*
 ‘I have a book **that says** everything (about Spanish grammar).’
- (80) Referentiality = generic
U: [No sé cómo se] llamarán.
Te- --
*.. Tepocates les **dicen** en [es]pañol, (214-1A2:37)*
 ‘U: I don’t know what their name is. Te—**they call** them tepocates in Spanish.’
- (81) Referentiality = generic
viven como antes.
...Tienen sus dioses.
pa' la lluvia y pa' esto y pa' lo otro.
.. pa' el maíz y,
*.. **Quién sabe** qué tantos dioses tienen. (142-3A2:149)*
 ‘They live like they did before. They have their gods. A god of rain, a god of this, a god of that. A god of corn. **Who knows** how many gods the have!’

Subjects coded as ‘event’ or ‘language’ do not designate entities; events indicate the linguistic expression of a situation (event or state). In the data, these subjects tend to appear as clauses, finite or non-finite, in impersonal constructions, as in (82) and (83). The label ‘language’ indicates the subject is a linguistic unit referring to an element or segment of speech, as in (84).

- (82) Referentiality = event

O: *Ahora mismo él* --
he is an epileptic,
y vive solo.
ha [vivido solo].

A: [Mhm].

O: *pero si miro su apartamento,*
*parece que **una mujer lo está cuidando el apartamento.***
.. Porque él es tan limpio.
para él, pa' su casa y pa' todo. (117-1A3: 506)

‘O: Right now he.. he is an epileptic, and lives alone. He has lived alone. A: Mhm. O: But, if you see his apartment, it seems **that a woman is taking care of it.** Because he is so clean. In his person, his house, in everything.’

- (83) [pues],

*Sabe usted que a mí se me hace **que cinco dólares,***
por cortar unos zacate es muy bueno=. (214-1B3: 435)

‘Well, it seems to me **that five dollars** for pulling some grass is very good money, you know?’

- (84) *Y le dijeron una palabra que no entendía y fue a*

-- *a*

-- *a ver a Rubén .. Murillo,*

*que le explicara qué era lo que quería decir **la palabra esa.** (144-4A2:167)*

‘And they told him a word he didn’t understand. So he went to see ... to see Ruben Murillo; he wanted him to explain what **that word** meant’.

Finally, the referentiality of verbs labeled as NA for subject was coded based on the referent of their implicit agent. This criterion was mainly applied to address two major cases: impersonal constructions with *se*, and verbs complementing a preposition or a verb. Although syntax prevents verbs to take a subject in the former situation, or allow for the use of non-finite forms in the latter one, there is an implicit agent behind the event. Thus, in (85) there is a generic referent because the event expressed by *hablar* ‘to speak’ refers to any person in that situation, even if a subject cannot be included because of the use of *se*, and in (86) there is a specific referent because the event’s agent is a group of identifiable and specific people (*los nietos* ‘the grandchildren’).

- (85) Referent = generic

*yo no más pienso que ellos .. quieren que **se .. hable** no más en inglés. (20-5A3: 71)*

‘R: No. I think that they want only English to be spoken.’

(86) Referent = non-SAP

E: *el inglés lo habla un poco, y lo entiende.*

L: *bu- --*

bueno.

ahí medio les cucharéo a los nietos.

E: *Ahá. sí. (H) [y] –*

L: *[Los] hago **entender**.* (219-1A1: 280)

‘E: English, you speak it a little bit, and you understand it. L: Well, I kind of speak it with my grandchildren. At least I make **them understand** me.’

3.3.5. Polarity

Originally, this category had four values: affirmative, negative, interrogative, and interrogative-negative. At the end, to get a better picture of how speaker use different type of utterances, only two values were included, affirmative and negative, and a separate category focusing on their purpose was created (see 3.3.6). Coding for polarity aims to attest relationships, if any, between utterances’ assertion or negation, and particular semantic nuances or pragmatic functions. Regarding this topic, Travis (2006a) finds that Colombian speakers of Spanish often use *no sé* ‘I don’t know’ as a kind of adverb highlighting the presented information, rather than as a main predicate indicating lack of knowledge (e.g. *yo no sé si te parezca malo, que estoy hablando* ‘I don’t know if you think it’s wrong, that I am talking to you’). Also, the author finds instances in which the collocation works as a mitigator since it minimizes a utterance conveying disagreement (*Pero es que tan--.. ay, yo no sé, tan bobos*). In the database for this study, the collocation seems to have the same function, as shown in (87), where the speakers use *yo no sé* ‘I don’t know’ to mitigate the assertion he has just expressed about the lack of jobs.

(87) A: *Qué pasaba en Nuevo México?*

K: *% Seguro que no había .. trabajos.*

yo --

Yo no sé. (318-1A2: 76)

‘A: What was going on in New Mexico? K: For sure there weren’t jobs. I-- **I don’t know**’

3.3.6. Type of Sentence

It codes how speakers enounce the utterances. The category includes declarative, interrogative and imperative utterances. Also, the category includes values for impersonal constructions with *se*³². When necessary, values are combined to indicate the utterance has characteristics from different types of sentences. This situation is illustrated by (88), which is both an interrogative and an impersonal construction. Like the previous categories, the present one looks for evidence showing that specific subject-verb combinations are increasing their degree of fixation, in this case, for intonation.

(88) Type of sentence: interrogative-se

A: *sus –
cómo se dice?
sus books.*

K: *Yeah.*

A: ... *Llevar sus cuentas.* (318-1A2: 475)

‘A: His—how do **you say it**? K: Yeah. A: Keeping the books.

3.3.7. Verb’s Function and Form

This category focuses on the syntactic function that the cognitive or speech verb has in its utterance, as well as in its realization. For function, the values are main verb, auxiliary, subject, complement, predicate and adjunct. Main verb indicates an item functioning as the predicate’s nucleus, either by itself, as in (89), or as a part of a construction. For the latter case, an indication was added regarding the type of construction (e.g. periphrasis) and the form of the verb (e.g. infinitive) when necessary. Thus, utterances such as (90) were coded as ‘main verb / periphrasis: infinitive’, while

³² The value includes both impersonal and passive *se* constructions. Following Maldonado (1999; 2007) and others, I consider passive *se* constructions to be a type of impersonal because of the existence of an implied human agent. Although a separate category could be created to distinguish between active, middle and passive voice constructions, I decided not to do so because of only a few of the utterances in the data are in the two latter voices.

utterances such as (91) were coded as ‘main verb / periphrasis: gerund’³³. It is important to mention that, following Gómez Torrego (1999), periphrases of infinitive include constructions with verbs classified as modals in English (e.g. *poder* ‘can’, *deber* ‘must’, etc.).

Auxiliary refers to a cognitive or speech verb that occurs in a personal form, and accompanies a main verb. In the corpus, the only verb functioning as an auxiliary is *saber* ‘to know’, as in (92).

- (89) Verb’s function and form = main
 A: [Usted] **cree** que esa ley es .. justa?
 K: .. [No]. (318-1A3:563)
 ‘A: Do **you think** that law is fair? K: No.’
- (90) Verb’s function and form = Main verb/periphrasis: infinitive
Oh, sí son.
.. eran.
*lo que iba a **dicir** porque,*
ahora la plebe vive el día pues,
a como que le dan ganas. (190-3B1:195)
 ‘Oh, they are indeed. They were. What **I was going to say**, because, now, kids live day by day, they do what they want’.
- (91) Verb’s function and form = Main verb/periphrasis: gerund
 A: *Que algunos hoy **estaban diciendo** que –*
este,
Ross Perot –
Perot,
No quiere ni al Clinton ni al Bush. (20-5A3:1)
 ‘A: Today, some people **were saying** that... this guy, Ross Peror. He doesn’t like Clinton nor Bush.’
- (92) Verb’s function and form = Auxiliary
 U: [Ella] tenía= .. de --
yo creo ya tenía como diez y nueve.
 E: Ah. joven.
 U: [Diez y] nueve años.
 E: [joven].
 U: *De diez y ocho a diez y nueve estaba por ahí, ((AHÍ))*
*no **sabría decir**. Porque ya ahí, había acabado la escuela.* (214-1A1: 1179)

³³ There were no cases of periphrases of participle.

‘She was... I think she was like nineteen years old. E: Ah, young. U: Nineteen years old. E: Young. U: She was around that place since she was eighteen until she was nineteen, I couldn’t say.’

The label ‘subject’ refers to nominalized verbs functioning as the subject of a main predicate, as in (93). Complement indicates a non-finite verb complementing another one, as in (95)-(96). For the latter, it was also indicated that the main verb requires or includes a preposition. The label ‘predicates’ refers to utterances in which the cognitive or speech verb complements a copulative one, as illustrated by (97). Although, in the latter case *decir* does not complement *es*, in a strict sense, since it is not really an attribute, the verb was labeled as predicate for matter of patterns’ operationalization.

- (93) Verb’s function = Subject

*Me siento igual de hablar inglés que español porque,
.. pues,
depende con quién.
pero me gusta más **hablar** español. (102-1A1:133)*

‘I feel the same speaking English or Spanish because... well, it depends on whom I’m talking to. But I like better to **speak** Spanish.’

- (94) Verb’s function = adverb

*X: ésa fue la que se creció un --
que= se crió con ella.
.. y llegaba **hablándole** en inglés,
y la agarraba de las orejas ella,
<VOX no no no oye.
aquí me vienes a hablar n=o güirigüiri.
aquí me vas a hablar en mexicano. (47-1A2:233)*

‘X: She was the one who grew up a... the one who was raised by her. And she [the interviewee’s daughter] would come home **speaking** in English and she [the interviewee’s mom] would grab her from the ears while telling her: “here you won’t babble; here you have to speak in Mexican.”

- (95) Verb’s function = complement

*Tenía mucho dinero,
y viene y se casó con una m- --
mu- --
mujer que la **hizo creer** que la quería,
<@ y luego a los tres meses ya no la quiso @>. (214-1B2: 870)*

‘He had a lot of money. So he went and married a wom—a woman to whom he made **believe** he loved her. And then, three months later, he didn’t love her anymore.’

- (96) Verbs's function = prepositional complement

A: *Ah bueno.*

Vamos a hablar un poquito de= la lengua.

.. Por ejemplo=,

*qué lengua aprendió a **hablar** primero.* (117-1A2:104)

'A: Ah. Ok. Let's talk a little bit about language. For instance, what language did you learn to **speak** first?'

- (97) Verb's function = complement of a copula

... entonces le dijo .. el hombre a --

a --

a la vieja,

*.. % es **dijir** que no la dejas .. casar conmigo,*

le dijo. (311-3B2:550)

'So he told her... the man told the old woman, **you mean** (lit. 'it is to say') you won't let her marry me, he told her.'

3.3.8. Tense, Aspect and Mood (TAM)

The hypothesis for this category is that variation, or lack of, in tense, aspect and mood may introduce pragmatic meanings into the discourse. On this topic, Scheibman (2002) found that speakers often use 'thought' not to refer to a situation in the past, but to frame a current situation when they want to convey surprise or disagreement (e.g. 'I thought we got our electricity fixed'). For Spanish, Travis (2006a) finds a similar phenomenon; speakers often use *yo pensé* 'I thought' not only to communicate a past idea, but to indicate they were mistaken (*Tú estabas -- yo pensé que estabas más despierto* 'You were -- I thought you were awake').

The labels used in each case correspond to Spanish tenses' names, each of which includes the aspect; for instance, 'preterit' indicates past tense and perfective aspect, while 'imperfect' indicates past tense and imperfective aspect. When not explicitly indicated, the verb's mood is indicative, as in (98). Direct commands (100) and exhortations (101) were both classified as 'imperative'. Given the syntactic variety of utterances in which the verbs under study appear, coding TAM required taking into consideration both the form and syntactic function of said verbs. Thus, in the case of

items in a non-personal form, they were coded according to the auxiliary verb's TAM if they were indeed part of the predicate's nucleus, as in (102); otherwise, they were coded as 'infinitive' or 'gerund', as in (103) and (104). The exception to these criteria is the construction *ir a* + INF 'to be going + INF'; when expressing future, as in (105), the construction's TAM was coded as periphrastic future rather than as a present because *ir* is a grammaticalized verb. When the construction does not express future tense, as in (106), it was coded as all other periphrases, even if it has an exhortative meaning.

- (98) TAM = present (mood = indicative)

*Andaba por leña mi tío y,
 .. en la laderita lo pescó un oso,
 y lo --
 .. Lo mordió todo.
 Le mordió las piernas,
 y le mordió la ca=ra,
 .. los brazo=s,
 .. Y,
 no más y --*

*No más Dios **sabe** cómo vivió.* (142-3B2: 396)

'My uncle was looking for firewood, and, on the hillside, a bear caught him and... and it bit him everywhere. It bit him on his legs, on his face, on his arms. And... only, only God **knows** how did he survive'

- (99) TAM = present-subjunctive

*le digo.
 aprende,
 le digo.
 maybe --
 puede que algún día de estos necesiten en un trabajo a alguien que **sepa** las dos ..
 idiomas.* (117-1A3: 810)

'I say to him. Learn. I say to him. Maybe... Maybe one day some job will need someone who **speaks** both languages.'

- (100) TAM = imperative

*K: Nunca,
 no.
 yo no me acuerdo que nunca nos dijieron eh --
 .. you know,
 cuando --
 When I was growing up,
 que me --
 (H) que me dijieran pues,
 no **hables** en mexicano,
 no pues.*

ahora sí me han dicho. (318-1A3: 275)

‘Never. No. I don’t remember anybody ever telling us, you know, when... when I was growing up, that I... that someone would tell me, well, “**don’t speak** in Mexican, no. Nowadays, yes, I have been told so.’

(101) TAM = imperative

A: *Cómo la conservaban **digamos** para el invierno?* (10-3A2: 250)

‘A: How did you preserve it, **let’s say**, for the winter?’

(102) TAM = present

A: *Y me **puede= platicar** usted de= dónde nació=,
... en qué año nació,
dónde creció?*

.. y cuánto tiempo lleva aquí en .. Dolores. (318-1A2:14)

‘A: And **can you tell me** where were you born, in what year, where did you grow up, and how long have you been here in Dolores?’

(103) TAM = infinitive

C: *es bueno tene- --*

*Mantener algunas de estas tradiciones,
.. o es mejor .. **olvidar** las tradiciones de antes,
y adquirir [tradiciones nuevas]? (4-5B2:84)*

‘Is it good to have... to keep some of these traditions, or is it better to **forget** traditions from before and acquire new traditions?’

(104) TAM = gerund

X: *ésa fue la que se creció un --
que= se crió con ella.*

*.. y llegaba **hablándole** en inglés,
y la agarraba de las orejas ella,
<VOX no no no oye.*

aquí me vienes a hablar n=o güirigüiri.

aquí me vas a hablar en mexicano. (47-1A2:233)

‘X: She was the one who grew up a... the one who was raised by her. And she [the interviewee’s daughter] would come home **speaking** in English and she [the interviewee’s mom] would grab her from the ears while telling her: “here you won’t babble; here you have to speak in Mexican.’

(105) TAM = periphrastic future

aquella habla no más en inglés.

Yo estoy hablándote aquí en mexicano a ti.

It's too bad .. si ella no me entiende a mí.

*porque si ella habla en inglés yo la **voy a entender** a ella.*

porque yo sé los dos. (117-1A3:868)

‘that one only speaks English. I am talking to you in Spanish. It’s too bad if she does not understand me because, if she speaks in English, **I am going to understand her**, because I know them both.’

(106) TAM = present

A: *Ah bueno.*

Vamos a hablar un poquito de= la lengua.

.. Por ejemplo=,

qué lengua aprendió a hablar primero.

y cuándo aprendió [a] –

O: *[El] mexicano.* (117-1A2: 104)

‘A: **Let’s talk** a little bit about language. For instance, what language did you learn to speak first? And when did you learn... O: Mexican.

3.3.9. Complement Type

This category codes the type of the element, if any, complementing the verb. It is important to mention that the term 'complement' is not used in its strict syntactic meaning (i.e. direct and indirect object) but in a broad sense: it indicates that a linguistic element or a segment of speech is being introduced, or referred to, by the verb, although it may not form a syntactic unit with it. Constructions with specific discourse functions often deviate from the canonical syntactic structure associated with the lexical meaning of the verb, for instance, showing less variability in terms of objects they occur with (Verhagen 1995; Company Company 2004; Cornillie 2005a; Cornillie 2005b); consequently, frequent collocations with non-typical complements, or no complements at all, may be performing metalinguistic, or even pragmatic, functions. Values included in the category are direct object, speech, prepositional complement, adjunct and none.

The value ‘direct object’ is used in its traditional way; it indicates that the linguistic unit accompanying the verb is its argument, as in (107). ‘Speech’ indicates an element that cannot be considered a direct object because of the lack of a subordinator, as in (108) or, even, because the verb and its complement do not really constitute a clause. That is to say, although the verb does refer to the speech segment, it functions as a parenthetical more than as main verb, as shown in (109).

(107) Complement = Direct Object (DO)

(H) .. *no podías saber que era gringa ella.*

podías decir que era mexicana,

because tenía hasta el accent como nosotros. (117-1A2:238)

‘You wouldn’t know she was gringa. **You could assure** she was Mexican because she even had the same accent we have.’

(108) Complement = speech

Me levanté teniéndome,

y me dice ésta,

<VOX pues, qué le pasó mamá VOX>?

Y le digo,

yo no sé qué me pasaría. (219-1B1:46)

‘I got up holding myself [on the bed]; so **she asked**: “what happened to you, mom?” And I say to her: “Well, I don’t know what happened to me.”’

(109) Complement = speech

D: *pero que antes de eso aquí había libros en [español].*

S: *[Oh yeah].*

D: *[2aquí había2] calendarios.*

S: *[2había libra2] --*

[calendarios había].

D: *[había todo].*

[2Mhm2].

S: *[2las2] boletas de votar todavía.*

Todavía aquí en el= estado,

yo creo,

en Nuevo México y el --

por lo menos en el condado de [Taos]. (102-3A2: 356)

‘D: But before that, there were books in Spanish here. S: Yeah. D: There were calendars here. S: There were bo... there were calendars. D: There was everything. Mhm. S: Even elections ballots. Even here in the state, I think, in New Mexico and the... at least in Taos County.’

Prepositional complements are adjacent constituents required by the verb (e.g. *acordarse de* ‘to remember’, *pensar en* ‘think about’, etc.), as in (110). Although these complements are necessary to form a grammatical unit when the preposition is expressed, they are not arguments of the verb because their occurrence is not required when the information they convey is understood from the context; case in which the preposition is absent. The phenomenon is illustrated by (111) below.

(110) Prepositional complement

C: *No **te acuerdas de otro chiste de esos?** (142-3B2:224)*

‘**Do you remember** any other joke like that?’

- (111) Complement = None
*y algunos tenían vestidos*³⁴.
No muchos,
pero algunos tenían vestidos.
Y compraban levás.
*... **Yo ni me acuerdo.***
yo creo quizás la compraban --
.. separadamente porque ah --
porque casi era una cosa rara que una persona llevara un vestido. (144-3A2:115)
 ‘And some [men] wore *suites*. Not many, but some wore *suites*. And they would buy dress coats. **I don’t even remember.** I think maybe they bought them separately because... because it was almost a weird thing that someone would wear a suit.’

Since formally these latter utterances do not have a complement, they were coded as NA (not applicable). The same criterion was followed for verbs in intransitive patterns (112), and verbs ‘sharing’ the complement with other verb, as in (113). Finally, despite not being arguments of the verb, adjuncts were included in the category because their inclusion in the utterance introduces a particular nuance and, given so, they do form some kind of unit with the verb, at least from a semantic and discursive point of view. This type of constituent is shown in (114) - (116).

- (112) O: *no podían hablar como nosotros porque,*
el -- el language de ellos era .. inglés.
desde que ellos nacieron,
ya lo hablaban.
*como **hablan** los gringos.*
 A: <X No sé X>.
 O: *Tenían su accent and everything. (117-1A2:194)*
 ‘O: They couldn’t speak like us because their language was English. Since they were born, they spoke it as *gringos* **do**. A. I don’t know. O: They had their accent and everything.’
- (113) L: *Y luego en la tarde cuando llegó mi marido de la escuela,*
*.. **le digo**, <VOX Yo seré tonta VOX>,*
*le **dije**. Porque no estoy educada.*
pero tú tan educado,
.. y pusites los telares,

³⁴ In normative Spanish, *vestido* ‘dress’ usually refers to women’s clothing, although DRAE’s first entry defines the word as “Prenda o conjunto de prendas exteriores con que se cubre el cuerpo” (*Piece or group of clothing used to cover one’s body*). For New Mexican Spanish (1983), Cobos defines the word as ‘man’s suit’. Related to the said definition, DRAE includes the compound *vestido de ceremonia* for men’s formal attire.

le [dije],

E: [@@@]

L: *a=á en la casa.* (219-1A2:278)

‘L: And later, in the afternoon, when my husband got home from school, **I say** to him, “I might be dumb”, I told him, “because I am not educated; but you, so educated and yet you installed the looms at home”, I said to him.’

(114) *no estamos diciendo nada de ti,*

le dije.

*Yo le estoy **diciendo** a la Elisabeth del puzzle,*

le dije,

en mexicano. (318-1A3:476)

‘“We’re not saying anything about you”, I said to her. “I am talking to Elisabeth about the puzzle”, I told her in Mexican.’

(115) *él,*

sabe qué estás ha-

*-- **diciéndole** en mexican,*

pero en inglés le responde pa’trás. (47-1A3:74)

‘He knows you are speak- talking to him in Mexican (sic), but he answers back in English’

(116) Y: [*Asina*] *me **dijo**,*

just take a snack. (147-2A2: 316)

‘That’s what he **told** me, just take a snak’ (*asina* ‘in this way’)

The last type of complement is shown in (117). As it can be seen, it functions as a kind of attribute because, rather than expressing an event, the construction seems to mention a quality; in this case, an alternative name for the entity being talked about. Although the element seems to directly complement the verb, I decided not to classify it as a direct object given its correferentiality with an indirect object pronoun (e.g. *le* ‘to it/him/her’).

(117) *Muchos le **dicen** ovispa*³⁵,

*ovispas y otros les **dicen** abejas,* (214-1A2: 49)

‘A lot of people **call** it wasp, wasps. And some others **call** them bees...’

³⁵ *Ovispa* is the form for *avispa* in NM Spanish (see Cobos 1983). In this variety of Spanish, the term refers to a honeybee rather than to a wasp.

3.3.10. Complement Realization

This category codes the realization of the complement; its purpose is to identify whether any particular morpho-syntactic unit tends to appear with specific subject-verb combinations. The values include full nominal phrases (FNP), nominalized elements, pronouns (e.g. *lo/la* ‘it/him/her’, *eso* ‘that/so’), adverbs (e.g. *cómo* ‘how’, *así* ‘this way/like that’), phrases, clauses, sentences, and discourse.

FNP refers only to those phrases whose nucleus is an actual noun, as in (118) and (119). Nominalized elements (e.g. infinitive verbs, possessive pronouns, etc.) were coded in a separate group, even though they functionally are FNP’s. The decision was taken seeking to identify patterns emerging from the relation between subject-verbs combinations and their complement. For instance, the co-occurrence of *QUERER*³⁶ ‘to want’ and the infinitive *decir* ‘to say’, is often interpreted as a cognitive intention or explanation (e.g. *quiero decir*, ‘I mean’, *quiere decir* ‘it means’), rather than as a desire, as it would be with other infinitives (e.g. *Quiero descansar* ‘I want to rest’) or even if the event with *decir* were to include a recipient (e.g. *Quiero decirle que me visite* ‘I want to tell her to visit me’). The contrast can be observed in (120) and (121) below. This kind of phenomenon could be overlooked if these infinitives would have been classified as a FNP.

- (118) *ellos sí sabían (THROAT) cuando estaban chiquitos español y, (H) y no sabían el inglés.* (318-1A2:588)
‘They did know Spanish when they were kids, and and **they didn’t know English.**’

³⁶ The use of small caps indicates the item can be used in different forms; for instance, tenses, person, etc.

- (119) *le digo.*
aprende,
le digo.
maybe --
*puede que algún día de estos necesiten en un trabajo a alguien que **sepa** las dos ..*
idiomas. (117-1A3: 810)
 ‘I say to him. Learn. I say to him. Maybe... Maybe one day some job will need someone who **knows both languages.**’
- (120) *ya como cuando crecieron ya ahora todas hablan es-*
--Todas pueden hablar.
Quiero decir.
Que lo hablen es otra cosa. (147-1A2:392)
 ‘Now, when they grew up, now all of them speak Sp... They all can speak [Spanish]. **I mean**, whether they actually do it it’s a different thing’
- (121) *Le=s dijo, vengan pa’ acá,*
*les **quiero hablar.*** (88-1A3:133)
 ‘(He) said to them: “come here; **I want to talk** to you.”’

For pronouns other than personal ones (e.g. *lo* ‘it/him’, *la* ‘it/her’), it was indicated whether they are demonstrative, relative or any other type, as illustrated in (122) and (123). The label ‘phrase’ refers to syntactic constituents, and it was complemented by an indication specifying their type (prepositional, adverbial, etc.). With a few exceptions, phrases in the data are prepositional ones, as in (124).

- (122) Complement realization = pro-demonstrative
había doce .. hijos de familia,
pero dos se murieron,
*... Eso **dijo** Eduardo.*
Cuántos nietos tiene usted?
Sabe cuántos? (311-4A1: 7)
 ‘There were twelve kids in the family, but two of them died. That’s what Eduardo **said**. How any grandchildren do you have? Do you how many?’
- (123) Complement realization = pro-relative.
Entonces me dio más coraje y le digo,
no necesita dicirme,
le dije.
yo sé hablar en inglés.
*.. y entiendo todo lo que me **estás diciendo.*** (318-1A3: 442)
 ‘So that made me even angrier, so I said to her, “he doesn’t need to tell me”, I told her, “I can speak English and I understand what you’re saying to me.”’

(124) Complement realization = phrase-phrepositional

C: *Y no **hablaba** tu papá de los antepasados?*

J: *No,
hombre,
XX.*

C: *No.*

J: *muy --
... (TSK) .. cómo quiero decir.
... No ten- --
no --
no teníamos --
... ahm, ... concepto.
del --
de --
de --
de la historia. (144-4B2:357)*

‘C: Didn’t your dad **talk** about your ancestors? J: No, man. C: No. J: Very... what can I say? We didn’t hav-- we didn’t have a concept ab-- about history.’

Complements coded as clauses are those having a verb in a finite form, if and when the verb was subordinated to another one, as in (125). If there is no subordination, the complement was coded as sentence (simple, compound or complex), as in (126) and (127).

(125) *no podías saber que era gringa ella.*

***podías decir** que era mexicana,*

because tenía hasta el accent como nosotros. (117-1A2:238)

‘You wouldn’t know she was gringa. **You could assure** she was Mexican because she even had the same accent we have.’

(126) *y .. tenía un mente brillante.*

% buenos grades sacaba.

*y **dijo**,*

ya no quiero.

le faltaban .. seis pa’ gradar. (88-1A3:683)

‘And he had a bright mind; good grades he used to get. But he **said**, “I don’t want [to study anymore].” He only needed six³⁷ more to graduate.’

³⁷ It is not clear if the speaker is talking about six months or about sixth grade.

- (127) Y= *le dije a mi dad,*
nos vamos a mudar solos,
porque yo estoy teniendo mi familia,
y no cabemos aquí. (117-1A3:525)
 ‘And **I said** to my dad, “We’re are getting our own place because I have a family and we there’s not enough space for all of us here.’

Finally, complements coded as speech in the type category (section 3.3.9) were the hardest to code for realization because they do not always form a syntactic constituent and sometimes it was not even clear whether the verb and the complement actually form a syntactic unit, just a discursive one or none. Nevertheless, for matters of methodological uniformity, these complements were coded based on the element that the verb seems to refer to. Thus, in (128), *muy lejos* ‘very far away’ it is the part of the discourse to which *yo creo* ‘I think’ directly refers; therefore, it is was taken as the complement and coded as adverbial phrase for its expression. If the complement did not correspond to any morphosyntactic unit, it was coded as discourse, as *pa’ qué* ‘what for’ in (129).

- (128) A: .. *Sí,*
hace poquito comimos aquí alguien,
una .. señora de Chamita.
.. nos sirvió capulín.
 N: *Capulín.*
Yo no sé de dónde traerán capulín aquí.
Yo no sé dónde se dará.
Muy lejos yo creo. (10-3A2: 175)
 ‘A: Right, we ate that here recently. A lady from Chamita fed us chokecherry. N: Chokecherry. I don’t know where they get chokecherry. I don’t know where it might grow. Very far away, **I think**.

- (129) U: *Y está--*
gente de hoy en día,
jóvenes,
se casan y --
y --
y duran un tiempo y luego pronto se apartan no?
.. Y ahora, se están apartando hasta má=s pronto.
Ahí mira papel lleno de puro divorcio.
 R: *Es todo que pasa.*
 U: *Yo no sé pa’ [qué]?* (214-1B2: 846)

‘U: And it is... people today, Young people, they get married and... and are together for some time, but shortly after they separate. And now, they are separating even sooner. Look at the newspapers, full with divorces. R: It because of all the things that happen. U: **I don’t know** what for.’

3.3.11. Connective

The category codes the type and form of the unit, if any, linking the verb with its complement. This is particularly relevant for clausal complements because, in a strict sense, the occurrence or absence of a subordinating conjunction (e.g. *que* ‘that’, *si* ‘if, whether’) relates to whether the utterance’s verb is the syntactic head of the constituent it refers to, as in (130). On the contrary, when juxtaposed to its ‘complement’, the verb may lose its governing function and, as a consequence, appear after the utterance or even in the middle of it, as in (131) and (132) respectively.

(130) *Ella no **sabe** que es mexicana.* (318-1A3:675)
 ‘She doesn’t **know** that she is Mexican.’

(131) *I: Yo todavía no nacía **yo creo**.* (142-3B2:49)
 ‘And she hadn’t been born yet **I think.**’

(132) *S: Yeah. Fue en esos tiempos.
 en los cincuenta.
 ... cuando .. se puso trabajoso **yo creo** pa' todo el chicano no? (102-3A2:343)
 ‘Yeah, It was in those years, in the fifties, were things got complicated for, **I think**, all chicanos, no?’*

The preposition used by prepositional complements is coded in this category too. The purpose of doing so, even if the type of connector is implied in the type of complement, is to identify any variation in the use of prepositions by the same verbs and, if it is the case, to explore semantic nuances resulting from such variation. For instance, the verb *hablar* sometimes appears with a phrase headed by *en* ‘in’ and other times with a phrase headed by *de* ‘about’, which affects the meaning of the construction, as showed in the examples below.

- (133) ...*hay veces que estoy platicando con= somebody que no sabe en – en % inglés.*
cuando menos acuerdo les he dicho en inglés,
*y luego tengo que **dijir** en mexicano.* (318-1A3:535)
 ‘Sometimes I am chatting with someone who doesn’t know English. After a while, I realize I have been talking to them in English and then I have to **say** it again in Mexican.’
- (134) S: *Había los almanaques que daban en la tienda.*
 ((skipped lines))
Y eran en español no?
 D: [Sí].
 S: [Te] **dijían** de la luna,
te dijían cuándo sembraras, cuándo [cosecharas]. (102-3A2:272)
 ‘S: There were almanacs that the stores gave as gifts. And they were in Spanish, you know? D: Yes. S: **They told you** about the moon phases, they told you when to plant the seeds, when to harvest.

3.3.12. Word order

The category codes for the position of the verb in relation to the constituent it refers to. Coding for this feature seeks to observe whether changes in words order follows what is established by the grammatical system of Spanish, as well as to determine whether different word orders relate to morpho-syntactic variation. For instance, while it is acceptable to place a lexical direct object before its verb, such a displacement requires, from a prescriptive perspective, pronominal reduplication, as in the invented example shown in (135). Therefore, having a pre-verbal, non-pronominal complement with no reduplication provides evidence that it less dependent from the verb (136).

- (135) *El español lo **hablo** en la casa; el inglés, en el trabajo.*
 ‘Spanish, I speak it at home; English, at work.’
- (136) [Pero él le enseñó el inglés],
*me **dijo**.* (20-1A1: 411)
 ‘But he taught you English, **you told** me.’

3.3.13. Indirect Object

This category observes whether the verb has an indirect object or not. Values are similar to those of subject's form in the sense that they focus on first, second, and third person singular or plural; but an indication was added when the referent was not specific. Naturally, the value NA (not applicable) was used for those utterances lacking such constituent (137). Based on its semantic roles, the presence of an indirect object usually indicates that the event expressed by the verb is addressed to a recipient, either specific (138) or generic (139), or that it is experienced by an animate entity; therefore, its occurrence is closely related to the speaker's communicative purpose. For instance, *te* 'to you', the generic referent in (139), suggests the speaker is talking about a general idea that is not addressed to anyone in particular, rather than talking about a specific speech event. Furthermore, in addition to the implications derived from the semantic role of its referent, observing patterns for indirect object is important because, in some dialect of Spanish, discourse markers such as *ándale*, *ándenle*, *dale* (Company Company 2006a), *híjole*, and others include pronominal third person indirect objects although it is not always clear what its referent is, let alone its semantic role.

- (137) Y **dicía** la gente que era –
.. *Que era una bruja.* (142-1B2: 86)
'And people **said** she was... she was a witch.'

- (138) *Pues,*
si éranos tantos de familia,
(H) y teníanos tanto que hacer,
.. *los **dicía** papá que, ((NOS))*
.. tal y tal,
tienen que ordeñar,
y tal y tal tienen que asistir los marranos,
y tal y tal tienen que ir a partir leña,
.. y sacar la ceniza,
.. y,
y eso. (190-3B1)

'Well, since we were a big family, and we had a lot to do, Dad **used to tell** us such and such, have to milk the cows; such and such have to take care of the pigs; such and such have to split the wood and remove the ashes, and, and that [kind of things].'

- (139) *S: [Te] **dijían** de la luna,
te **dijían** cuándo sembraras,
cuándo [cosecharas].* (102-3A2: 272)
'S: [Farmers' almanacs] **told you** about the moon phases, they **told you** when to plant the seeds, when to harvest.'

3.3.14. Prosodic Distribution of the Complement Regarding the Verb

This category observes whether the verb and the element it presents appear in the same or in different intonation units. Constituents are marked as sharing the intonation unit with the verb not only in cases such as in (140) and (141), where the whole element do so, but also in utterances in which the verb and at least a lexical element of the complement are in the same IU, as in (142).

- (140) *[Rastros] **les dicen**,
o les **dicen huellas**.* (214-1B2: 143)
'**They call them** traces or **they call them** footprints.'

- (141) *no más llegué y me **dijo** ella qué es lo que hiciera.* (10-8A: 349)
'As soon as I got there, **she told** me what to do.'

- (142) *Pero pa' allá **entiendo** que iban,
que iban a la –
a la caza de los cíbolos.* (270-1B2:522)
'But **it's my understanding** that they used to go... they used to go to... they used to go buffalo hunting'.

This coding and the one in section 3.3.15 below relate to Chafe's *one new idea constraint* (1994: 108), according to which an intonation unit cannot have more than one piece of new information because of the cognitive effort it would mean for speakers. Since such type of information is expressed by lexical units³⁸, a non-pronominal complement can co-occur with the verb in the same intonation unit only if at least one of

³⁸ By lexical units, I refer to those containing either a verb or a noun.

them expresses given or accessible information, or if the verb is a low content one, as *decir* (cf. Chafe 1994). Said verbs represent a low activation cost because they do not express an independent idea of their own, but provide the framework for the event expressed by the predicate. Following this constraint, then, in (140) above, *rastros* and *huellas* are new information in their respective intonations units, while the referent of *les* 'them' is a generic entity and, therefore, it has a low cognitive load, so it qualifies as accessible information. For (141), only *qué es lo que hiciera* 'what I had to do' is new information, so it can share the IU with both *llegué* 'I arrived', idea previously mentioned in the conversation, and *dijo* 'she said' without any problems.

Why, then, do a verb such as *decir* sometimes appear in a separate intonation unit from its complement, as in (143) and (144) below? A possible explanation is that the verb actually expresses new information; we know someone does something, but we do not know what. Another possible explanation is that the assumed high cognitive load of *decir* lies on its use as a device to report a piece of speech as if it were actually happening at the moment of the enunciation.

- (143) *Y le digo,*
yo no sé qué me pasaría.
mira,
mi pie estaba azul [azul], (219-1B1: 46)
 'and then **I tell** to her: "I don't know what happened to me, look." Mi foot was all blue.'

- (144) *Pero éste dice,*
si no aprendo inglés no voy a subir. (88-1A3: 29)
 'But **this one says**: if I don't learn English, I won't move up.'

Also, the frequent occurrence of a collocation in its own intonation unit may facilitate its syntactic displacement and, eventually, syntactic independence; phenomena associated with the development of discourse and pragmatic functions. For instance, in

(145), with the inclusion of *yo creo* ‘I think’ at the end of the utterance, the speaker mitigates the assertion he just presented.

- (145) K: *y luego,*
seguro que no había trabajo,
yo creo.
Yo no sé. (318-1A2:105)
‘And later, I’m sure there weren’t any jobs, I think. I don’t know.’

3.3.15. Prosodic Distribution of the Complement

This category analyzes whether the whole complement occurs in one or more intonation units. The hypothesis for this category is that the more information a complement adds to the discourse, the more intonations units it is likely to need to be expressed, which could affect the syntactic scope of main verb. In (146), it is easy to observe that *que ya no puedo estudiar* ‘that I cannot study anymore’ is the complement of *yo creo* ‘I think’ because it is uttered in one IU, which it shares with the main verb. On the contrary, the ‘complement’ in (147) takes four IU’s to be expressed, which makes it difficult to determine whether *yo creo* refers to the whole piece of information, just to *se puso trabajoso* ‘it got difficult’ or just to *pa’ todo el chicano* ‘for all chicanos’.

- (146) Y: **yo creo** *que ya no puedo estudiar.*
A: [@@@@@]
Y: [ya no está] *mi mente pa’ estudiar.* (147-1A2:471)
‘Y: I don’t think I can study anymore. A: [@@@@@]. Y: My brain is no longer good for studying’.
- (147) *Fue en esos tiempos.*
en los cincuenta.
... cuando
.. se puso trabajoso yo creo pa’ todo el chicano no? (102-3A2:343)
‘It was then, in the fifties, when it got really hard **I think** for all Chicanos, no?’

3.3.16. Transition Continuity of the IU Containing the Verb

This category observes whether the IU containing the verb has continuing (148), ascendant (149), final (150), or truncated transitional continuity (151). Examining this feature seeks to determine whether, when performing non-lexical functions, collocations tend to show a particular transitional contour. Truncated utterances were originally excluded from the analysis; however, they were reintroduced because truncation is often followed by reformulation; so, whether speakers resume, reformulate their idea, and how they do it, may provide important information about the behavior of cognitive and speech verbs in the discourse. For instance, in the eighth line of (151), speaker E's utterance is truncated because speaker U intervenes to offer his opinion about the information that speaker E states he has forgotten; the intervention, then, does not affect but collaborates in the construction of the discourse. On top of that, it is interesting that rather than just giving the information, speaker U adds *yo creo* 'I think' and *no?* 'isn't it?'. By doing so, the speaker participates in the conversation in a non-imposing nor pretentious way.

- (148) L: *Pero ya como de --*
 .. de ocho años tenía yo,
 ***yo creo**, ya comenzaron los libros a usar el --*
 en inglés. (219-1A1:392)
 'L: But around... when I was eight years old, **I think**, textbooks started to be
 in English.'
- (149) A: *No quería que hablaran español?*
 P: *Yeah. Yeah. su nombre era García.*
 A: *Y eso es contra la ley.*
 ***sabía usted?** (88-1A3:150)*
 'A: He didn't want all of you to speak Spanish? P: Yeah, yeah. His name was
 García. A: That is against the law, **did you know?**

- (150) M: *[Oh, sí].*
Yo le mandaba .. birthday cards,
y él también a mí.
 B: *Ah= [okay].*
 M: *[tú sabes].* (76-1A1:41)
 ‘M: Oh, yeah. I would send him birthday cards, and he to me too. B: Ah, ok.
 M: **You know.**’
- (151) U: *[Pero=] no está lejos de ahí de --*
Flagstaff?
 E: *No no está muy lejos.*
 U: *Ah=.*
 E: *<P No está muy lejos P>.*
 U: *two hundred y algo millas?*
[o menos]?
 E: *[Mhm].*
No me acuerdo cuántas millas [2pero2] –
 U: *[2two hundred2] and ninety yo creo,*
qué no? (214-1A1:94)
 ‘U: But was it far from Flagstaff? E: No, it’s not very far. U: Ah. E: It’s not too far. U: Two hundred and something miles? Or less? E: Mhm. I don’t remember how many miles are but... U: Two hundred and ninety, **I think**, right?’

3.3.17. Change of Turn

This category observes whether there is a change of speech turn after the IU including the verb under analysis. Coding for this feature seeks to identify collocations performing interactional functions such as back channeling, invitation to participate in the discourse, interlocutor’s reply, etc. Change of turn was coded as existing either when the speaker clearly intended to finish his/her turn (152)-(153), or when the utterances got truncated because of the intervention of the interlocutor (154). In the examples below, the intonation unit where change of turn occurs is indicated with an arrow (←).

- (152) C: *Y quiénes eran los mejores para decir historias?*
*.. que te **acuerdes** tú.* ←
 I: *Oh, había mucha gente que –*
que decía muy buenas historias.
pero ahora no sé quiénes serían. (142-3B2:132)
 ‘C: Who were the best storytellers? That **you remember**? I: Oh, there were many people who.. who told very good stories. But I don’t recall now who they were.’

- (153) E: *Cuántos años tenía cuando se casó?* [Wendy].
 U: [Ella] tenía= .. de --
yo creo ya tenía como diez y nueve. ←
 E: Ah. joven. (214-1A1:1173)
 ‘How old was she when she got married? Wendy. U: She was.. **I think** she was nineteen years old. E: Ah, she was young.’
- (154) U: *Esta es una huella de a --*
<X tal X> animal dicen,
y otros dicen, este es un rastro de animal.
 E: Hm.
 U: Se --
se dice de dos modos.
 E: De dos modos.
 U: Pero no --
ahora yo no sé cuál será la -- ←
 E: Pero=,
el --
el --
no más la --
la= p- --
palabra que usted prefiere. (214-1B2:183)
 ‘U: This is an animal’s footprint, they say. And other people say “this is an animal track”. E: Hm. U: There are two ways of saying it. E: Two ways. U: But I don’t.. Right now **I don’t know** which one might be the... E: But, the...the... just the.. the name you prefer’
- (155) J: *Y le dijeron una palabra que no entendía y fue a --*
a --
a ver a Rubén .. Murillo,
que le explicara qué era lo que quería decir la palabra esa.
 C: [yeah].
 J: [**me**] acuerdo yo. ←
 C: *Y no te acuerdas de la palabra?* (144-4A2:181)
 ‘And they told to him a word that he didn’t understand, and he went with... with Ruben Murillo and ask him to explain the meaning of that word. C: Yeah. J: **I recall**. C: Do you remember the word?’

3.3.18. Function

This category coded for the function that each subject-verb combination has in the discourse. The major values included were lexical, interactional and epistemic. Lexical refers to utterances whose function directly derives from the verb’s meaning, thus communicating an objective interpretation, as in (156). The interactional function indicates the construction’s purpose is to manage the flow of the conversation, as *que te*

acuerdes tú ‘that you remember’ in (157), where the interviewer is trying to get the interviewee involved in the topic they are talking about. Epistemic indicates the utterance conveys some sort of speaker’s attitude (opinion, doubt, etc.). Because the utterance’s function relates to specific collocations and subject-verb combinations (e.g. quotes, indirect speech, argumentative, etc.), I will further discuss them in the corresponding sections of the local analysis chapter rather than trying to summarize here what specific nuances epistemic constructions may convey.

(156) Function = lexical

*No les **platicaba** historias o asina tu papá?*

.. Cuando eran niños. (144-4A2:46)

Did your dad tell you stories or things like that? When you were kids.’

(157) Function = interactional

C: Y quiénes eran los mejores para decir historias?

*.. **que te acuerdes tú.***

I: Oh, había mucha gente que –

que decía muy buenas historias.

pero ahora no sé quiénes serían. (142-3B2:132)

‘C: Who were the best storytellers? **That you remember?** I: Oh, there were many people who.. who told very good stories. But I don’t recall now who they were.’

(158) Function = epistemic

*P: **Qué puede hacer uno?***

*A: **Quién sabe.***

.. A la mejor sí se puede hacer algo. (88-1A3:517)

‘P: What can one do? A: **Who knows.** Maybe something can be done indeed.’

Chapter 4. General Patterns: Verbs of Cognition and Verbs of speech

This chapter presents and discusses morphosyntactic and semantic structures (e.g. subject-tense, subject-type of verb, verb-argument structure) for all tokens of verbs of cognition and verbs of speech in the data. By comparing patterns across said verbal classes, and exploring correlations between those patterns and the discursive genre from which data were extracted, this level of analysis seeks to determine whether speakers use each verbal class, as a whole, differently from each other; for instance, whether they prefer one type of verbs when talking about events from the past than when talking about current events; whether they use 1sg subjects more frequently with one class than with the other, or whether they convey subjectivity with one class more often than with the other one.

4.1. Type and Token Frequency

Table 3 shows type and token frequency for both classes of verbs. The first thing worth noticing is that speech verbs occur more frequently (1320/2134: 61.9%) than cognitive verbs do (814/2134: 38.1%), despite the former including fewer members (n= 11) than the latter (n= 21). These results differ from previous studies on structural patterns on spoken language (Scheibman 2001; Scheibman 2002; Silveira 2007)³⁹, thus proving that grammar is a living entity that arises from specific interactions.

Table 3. Verbal Class and Frequency				
	Token frequency	Type frequency	Average token frequency	Most frequent verbs
Cognitive verbs	n= 814	n = 21	n= 38.7	<i>saber</i> ‘to know’: n= 324 <i>creer</i> ‘to believe / to think’: n= 150 <i>acordarse</i> ‘to remember’: n= 105
Speech verbs	n= 1320	n= 11	n= 120	<i>decir</i> ‘to say / to tell’: n= 853 <i>hablar</i> ‘to speak’: n= 375

³⁹ In Scheibman’s studies, cognitive verbs were approximately 4% more frequent than speech ones; in Silveira’s, token frequency for both verbal classes is very similar.

Important too is the distribution of verbs among each class. For speech verbs, two items represent more than 90% of the class: *decir* ‘to say / to tell’ (853/1320: 64.6%) and *hablar* ‘to speak’ (375/1320: 28.4%). Moreover, *decir* occurs twice as frequently as *hablar*. The question arises, then, about the extent that such high token frequency is caused merely by the topics of the interviews, and whether it reflects a variety of structural patterns. For the cognitive class, three verbs account for more than 70% of the utterances: *saber* ‘to think’ (324/814: 39.9%), *creer* ‘to believe, to think’ (150/814: 18.42%), and *acordarse* ‘to remember’ (105/814: 12.9%). Like in the speech class, the most frequent verb (*saber*) occurs twice as many times as the second most frequent one (*creer*). Previous studies have shown the high token frequency of certain cognitive verbs as well; with frequency often relating to the occurrence of formulaic expressions (Scheibman 2001; Scheibman 2002; Silveira 2007; Travis 2006a; Valdivia 2013; Weber and Bentivoglio 1991; Bentivoglio and Weber 1999). A similar phenomenon has been observed for *decir*; its frequency in spoken language relates to the variety of speech acts it can express (e.g. to say, to tell, to declare, to ask, to command, etc.), and, consequently, the assortment of contexts in which it occurs (Briz Gómez 1993; Fernández Bernárdez 2002). Certainly, the predominance of *decir* in the corpus relates to the characteristics of the interviews, given that one part of them consists of speakers’ personal narratives in which they report their own and other people’s speech. As for *hablar*, it does not typically emerge as a frequent verb in studies about structural patterns; furthermore, although there are several formulaic expressions that include it, they usually are idioms⁴⁰

⁴⁰ For instance, *HABLAR bien* ‘to speak properly’, *HABLAR en chino* ‘to speak gibberish’, *HABLAR en castellano/cristiano* ‘to speak clearly’, *HABLAR directamente* ‘to speak frankly’, *HABLAR a tontas y a locas* ‘to speak willy-nilly’, *ni hablar* ‘forget it / no way’.

rather than constructions whose meaning and pragmatic implications depend on the specific context in which speakers use them. Thus, the question arises of whether the frequent use of *hablar* merely relates to the topics of the interviews, and whether the repetition of certain sequences may form prefabs and other constructions. I will explore these questions in subsequent sections.

4.2. Subject

Table 4 shows the distribution of the data according to subject and verbal class. As illustrated, the most frequent subject in the data is 1sg (823/2134: 38.6%), followed by 3sg (594/2134: 27.8%) and 3pl (350/2134: 16.4%). In addition to token frequency, other noteworthy difference between these subjects regards their occurrence within each verbal class; while 1sg subjects show just a slight tendency towards cognitive verbs (430/823: 52.2%), 3sg subjects tend to appear mostly with speech verbs (428/605: 70.7%), tendency that is even stronger for 3pl subjects (285/351: 81.2%).

Table 4. Utterances by verbal class and person⁴¹

	1sg	2sg (tú)	2sg (Ud.)	3sg	1pl	2pl	3pl	NA ⁴²	Total
Cognition	430 (52.8%)	70 (8.6%)	55 (6.8%)	175 (21.5%)	6 (0.7%)	2 (0.2%)	66 ⁴³ (8.1%)	10 (1.2%)	814 (100%)
Speech	393 (29.8%)	33 (2.5%)	78 (5.9%)	419 (31.7%)	68 (5.2%)	16 (1.2%)	284 (21.5%)	29 (2.2%)	1320 (100%)
Total	823	103	133	594	74	18	350	39	2134

⁴¹ In this table, person is categorized without distinguishing generic and specific referents. Such a distinction will be addressed when discussing local patterns and specific constructions.

⁴² It includes impersonal utterances such as *ellos quieren que se hable nomás en inglés* ‘they want people to speak only in English’, which do not have subject despite the verb is in 3sg.

⁴³ It includes a case in which the subject is *la gente* ‘people’; from a prescriptive point of view, the verb for this type of nouns should be in 3sg.

Important differences are observable also when focusing on how each subject combines with each semantic class. 1sg-cognitive is the most frequent combination within its class (430/814: 52.8%), as well as in the whole data (430/2134: 20.1%). When occurring with speech verbs, 1sg provides the second most frequent combination in the class (393/1320: 29.8%) and the third most frequent in the whole data (393/2134: 18.4%). Such results, then, suggest that 1sg subjects play a larger role when speakers use cognitive verbs than when they use speech verbs. In other words, when talking about beliefs and opinions, speakers seem to talk about themselves; when reporting speech events, they talk about themselves less frequently.

3sg-speech is the most frequent combination within its verbal class (419/1320: 31.7%), and the second most frequent in the whole data (419/2134: 19.6%). Within the cognitive class, 3sg is the second most frequent one (175/814: 21.5%), although the combination occupies the fifth position in the whole data (175/2134: 8.2%). Numbers suggest that when discussing speech events, speakers quote and report another person's utterances more frequently than they do their own.

Finally, 3pl-speech is the third most frequent combination of its verbal class (284/1320: 21.5%), and the fourth most frequent in the whole data (284/2134: 13.3%). The subject is not particularly relevant for the cognitive class; it occupies the fourth place in frequency, accounting for 8.1% (66/814) of the class, and just 3.1% (66/2134) of the whole data.

4.3. Tense

Table 5 summarizes the distribution of the data according to tense, and it allows to observe main differences regarding what tenses are more frequent and relevant for

each verbal class. To start, although present tense, the most frequent in the data (1238/2134: 58%), is similarly distributed among cognitive (669/1238: 54%) and speech verbs (569/1238: 46%); and both present-cognitive and present-speech are the most frequent tense-class combinations in the whole data, accounting for 31.3% (669/2134) and 26.7% (569/2134) of the utterances, respectively. Nevertheless, the tense represents a much larger portion of all cognitive verbs (669/814: 82.2%) than of all speech verbs (569/1238: 46%).

Table 5. Utterances by verbal class and tense⁴⁴

	Present	Preterit	Imperfect	Other	Total
Cognition	669 (82.2%)	29 (3.6%)	77 (9.5%)	39 (4.8%)	814 (100%)
Speech	569 (43.1%)	393 (29.8%)	258 (19.5%)	100 (7.8%)	1320 (100%)
Total	1238	422	335	139	2134

Unlike present tense, preterit tense, the second most frequent in the data (422/2134: 19.8%), occurs mainly in utterances involving a speech verb (393/422: 93.1%); said combination occupies the second place of frequency in its class (393/1320: 29.8%), and the third place in the whole data (393/2134: 18.4%). To a lesser degree, imperfect tense, the third most frequent in the data (335/2134: 15.7%) also tends to appear with speech verbs (258/335: 77%); the combination is the third most frequent among all speech utterances (258/1320: 19.5%), and the fourth most frequent in the whole data (258/2134: 12.1%). Even adding preterit and imperfect utterances, these results suggest that, in these interviews, speakers do not talk as frequently about past events as they do about present ones; moreover, when they do talk about past events, they

⁴⁴ The category of tense includes utterances with simple verbal structure, periphrastic constructions and modal constructions. For the last two, tense and person are based on the auxiliary and modal verbs.

tend to report information, usually other people's, rather than express thoughts and beliefs

4.4. Class, Subject and Tense

Table 6 summarizes the distribution of the data according to class, subject and tense. Following separate tendencies for subject and tense, the most frequent subject-tense combination in the data is 1sg-present (589/2134: 27.6%); it tends to occur with cognitive verbs (382/589: 64.9%), although its frequency with speech verbs is also noteworthy (207/589: 35.1%). This subject-tense combination is followed by 3sg-preterit (239/2134: 11.2%), which mainly happens with speech verbs (225/239: 94.1%).

Table 6. Distribution according to verbal class, tense and subject

	Present		Preterit		Imperfect		Other		Total
	Cogn.	Speech	Cogn.	Speech	Cogn.	Speech	Cogn.	Speech	
1sg	382	207	11	116	26	43	11	27	823
2sg-tú	64	17	0	6	2	3	4	7	103
2sg-Ud.	54	42	0	15	1	12	0	9	133
3sg	119	109	14	225	25	64	17	21	594
1pl	5	26	1	6	0	27	0	9	74
2pl	0	1	0	0	2	12	0	3	18
3pl	39	151	1	23	20	94	6	16	350
NA ⁴⁵	6	16	2	2	1	3	1	8	39
Sub-total	669	569	29	393	77	258	39	100	2134
Total	1238		422		335		139		

Slightly behind in frequency is the combination 3sg-present (228/2134: 10.7%), which is similarly distributed between cognitive (119/228: 52.2%) and speech utterances (109/228: 47.8%). Interestingly, despite being the third most frequent tense in the corpus,

⁴⁵ It includes utterances in which a verb appears in a non-personal form without being part of a periphrasis nor a modal construction, e.g., .. *o es mejor .. olvidar las tradiciones de antes, y adquirir [tradiciones nuevas]?* 'Or is it better to forget old traditions and adopt new ones? (4-5B2: 84). For tense, these tokens were classified based on the tense of the clause's main verb.

imperfect tense does not provide any of the top three subject-tense combinations; 3pl-imperfect, the most common combination within that tense group (114/2134: 5.3%), appears in the sixth position, occurring mostly with speech verbs (90/114: 78.9%).

4.4.1. Cognitive Verbs

Accordingly to the patterns above presented, the most frequent subject-tense combinations for cognitive verbs involve 1sg subjects. The most frequent one is 1sg-present, which accounts for 46.9% (382/814) of the class, 46.4% (382/823) of all 1sg utterances, and 30.9% (382/1238) of all present tense utterances in the data. This particular combination is the most frequent in the data as well (382/2134: 17.9%). Given the predominance of the combination both in the whole class and within its subject and tense categories, it is likely that, in the interviews, speakers spend a significant part of the discourse expressing what they know and think. By doing so, they do not only present themselves as important participants in the conversation, but also provide favorable conditions for the emergence and development of constructions conveying subjectivity.

The second most frequent subject-tense combination for cognitive verbs is 3sg-present, which accounts for 14.7% (119/814) of the verbal class; it is the second most frequent for 3sg utterances (119/594: 20%), and the fourth in the whole data. Thus, although somehow important within its class and subject categories, the combination is not as relevant in the whole data; suggesting that speakers do not talk about other what another person thinks or believe very often, either because they do not know so, or because it is not relevant for them and the conversation. Besides 1sg-present and 3sg-present, no other combination has a relevant token frequency in the cognitive group, at least not in relation to the whole class, let alone the whole data.

4.4.2. Speech Verbs

Within the speech class, 3sg-preterit is the most frequent combination (225/1320: 17%), and the second most frequent in the whole data (225/2134: 10.5%). As well, it is the most recurrent for its tense (225/422: 53.3%) and person groups (225/594: 37.9%). Closely behind, 1sg-present is the second most frequent combination for speech utterances (207/1320: 15.7), and the third most frequent one (207/2134: 9.7%); it accounts for 16.7% (207/1238) of present tense and for 25.2% (207/823: 25.2%) of 1sg categories.

Following this combination, 3pl-present is the third most frequent of the class (151/1320: 11.4%); although it emerges as the most frequent for its person group (151/350: 43%), and the third most frequent among present tense utterances (151/1238: 12.2%). The fourth and fifth combinations in the speech class show a similar token frequency. 1sg-preterit represents 8.8% (116/1320) of the class; it is the third most frequent among 1sg subjects (116/823: 14.1%), and the second most frequent of all preterit utterances (116/422: 27.5%). The combination is followed by 3sg-present, which accounts for 8.25% (109/1320) of the class, and it is the third most frequent for its subject group (109/594: 18.4%).

These patterns would suggest that, when talking about speech events, speakers tend to report another individual's speech when narrating past events, but to report their own, or other people's when the conversation deals with present events. Moreover, considering that preterit and present tenses contrast not just in terms of time, but also of aspect, their association with different subjects implies that speakers are using what is available in the language to alter the distance of a speech event depending on who the

subject is. Thus, by using a perfective tense to express what another person said, but an imperfective tense to express what they say, speakers implicitly convey that the latter speech events are more relevant since they are still in existence in the context of the conversation. I will come back to this hypothesis when examining constructions with specific verbs of speech (Chapter 6).

4.5. Subject Referentiality

Table 7 shows the distribution of the data according to subject's referentiality. Combining both verbal classes, subjects refer to speech acts participants (i.e. the interviewee and interviewer) in almost half of the cases (1038/2134: 48.6%), followed by specific, non-speech participants (640/2134: 30%) and generic referents (296/2134: 13.9%). Nevertheless, when observing verbal classes separately, important differences emerge suggesting speakers use each of them to do different things.

Table 7. Distribution according to verb class and subject referentiality

	SAP	SAP-plus	non-SAP	Generic	Inanimate	Event	Language	Total
Cognitive	537 (66%)	10 (1.2%)	124 (15.2%)	81 (9.9%)	4 (0.5%)	57 (7%)	1 (0.1%)	814 (100%)
Speech	501 (37.9%)	72 (5.4%)	516 (39.1%)	215 (26.4%)	8 (0.6%)	1 (<0.1%)	7 (0.5%)	1320 (100%)
Total	1038	82	640	296	12	58	8	2134

At first sight, SAP referents seem to be well distributed between both cognitive (537/1038: 51.7%) and speech verbs (501/1038: 48.3%); however, it is important to notice that they represent a larger proportion of all cognitive utterances (537/814: 66%) than of all speech utterances (501/1320: 40%). When using cognitive verbs, then, speakers either express their thinking or ask their interlocutor about his or hers; when using speech verbs, the interviewee and the interviewer do not appear as the source of information very frequently. Such patterns agree with what was found for subject

distribution in the sense that speakers appear as the subject of cognitive verbs more often than as the subject of speech verbs.

Non-SAP referents show a very different behavior; they tend to appear mainly with speech events (516/640: 80.6%) and, in fact, are the most frequent ones for said class (516/1320: 39.1%), although without representing a percentage as large as that of SAP referents for cognitive verbs. For this latter class, non-SAP referents are the second most frequent ones (124/814: 15.2%). A similar pattern is observed for generic referents: they appear mostly with speech verbs (215/296: 72.6%), being the third most frequent group for said class (215/1320: 16.3%). For cognitive verbs, generic referents barely account for 10% (81/814) of the utterances.

Finally, although subjects referring to events represent a small portion of the whole data (58/2134: 2.7%), it is worth noticing that they appear almost exclusively with cognitive verbs (57/58: 98.3%), even if they account for a mere 7% (87/814) of the class. Such results show that, when talking about ideas and opinions, speakers tend, on the one hand, to posit themselves as the subject of the verb (e.g. *yo creo* ‘I think’), but, on the other hand, they sometimes express those ideas as if they were indirect participants in the event (e.g. *se me hace que...* ‘it seems to me that’); that is to say, recipients rather than agents or, even, experiencers.

4.6. Type of Complement

Table 8 presents the distribution of the utterances according to type of complement. As mentioned in the section of coding, the term ‘complement’ refers not only to objective complements, but to any syntactic constituent frequently appearing with the verb, and adding to it a particular semantic nuance. Besides the fact that most

utterances take some kind of complement (1713/2134: 80.3%), it is noteworthy that verbs of speech take complements more frequently (1132/1320: 85.8%) than verbs of cognition do (581/814: 71.4%). Having said that, these results do not mean that speakers frequently use verbs of speech in canonical constructions, but they do so less often with verbs of cognition; simply, speech verbs seem to introduce information and to form some sort of unit (either syntactic or discursive) more frequently than verbs of cognition do. Relevant as well is the fact that cognitive verbs, unlike speech ones, barely take adjunct and attributes as complements; but the opposite phenomenon happens regarding prepositional phrases.

Table 8. Verb class and type of complement

	Cognition	Speech	Total
Adjunct	6 (0.7%)	152 (11.5%)	158 (7.4%)
Attribute	1 (0.1%)	61 (4.6%)	62 (2.9%)
Direct Object	401 (49.3%)	461 (34.9%)	862 (40.4%)
Prepositional	98 (12%)	16 (1.2%)	114 (5.3%)
Speech	75 (9.2%)	442 (33.5%)	517 (24.2%)
NA	233 (28.6%)	188 (14.2%)	421 (19.7%)
Total	814 (100%)	1320 (100%)	2134 (100%)

Another relevant result for the subjectification of cognitive and speech constructions is that NA appears as the third most frequent type of complement in the data (421/2134: 19.74%). Said pattern indicates a significant portion of the verbs do not introduce any information, either because they appear in intransitive utterances, or because they are used as discourse markers. Furthermore, utterances with no complement form the second most frequent group for cognitive utterances (233/814: 28.6%) and the third most frequent for speech utterances (188/1320: 14.2%).

Direct object is the most frequent complement in the whole data (862/2134: 40.4%); unlike other types, direct object does not show a strong tendency towards cognitive (401/862: 46.5%) nor speech utterances (461/862: 53.5%). The fact that a

substantial percentage of utterances takes a direct object is not surprising considering that, in the data, the most frequent items for each class are transitive verbs according to *Diccionario de la Lengua Española* (DEL) (Real Academia Española 2014)⁴⁶. What it is relevant, though, is the fact that, despite being the most frequent type of complement for both verbal classes, direct objects complement a larger percentage of cognitive utterances (403/814: 49.5%) than of speech ones (459/1320: 34.8%). The result seems to contradict what was said at the beginning of this section; however, it is important to notice that those results involve all types of complements, while here patterns refer just to direct object. Results imply as well that cognitive verbs would frequently appear in SVO, which might result in certain resistance to enter into a subjectification path. This hypothesis will be tested when examining local patterns for each semantic class.

Speech is the second most frequent type of complement in the data (517/2134: 24.2%); as expected, it appears mainly with verbs of speech (442/517: 85.5%), accounting for 33.5% (442/1320) of the class. Beyond the fact that this type of complement is closely related to the semantics of speech verbs, the results suggest that, in their discourse, speakers often present or make reference to verbal events. Relevant as well is the fact that speech complements are not among the most frequent ones for cognitive utterances; they merely represent a tenth of the whole class (75/814: 9.2%). Although relatively small, such a percentage implies speakers use cognitive verbs not just to express what they know or think, but sometimes as a way to present or comment information too.

⁴⁶ Consulted on 4/14/14

Worth mentioning as well are patterns concerning adjunct complements. Despite their low token frequency in relation to the whole data (158/2134: 7.4%), it is noteworthy that they appear almost exclusively with verbs of speech (152/158: 96.2%), accounting for almost a tenth of the class (152/1320: 11.5%): such behavior may be due to the occurrence of item-specific constructions. A similar situation is observed for attributes: they account for a very small percentage of the whole data (62/2134: 2.9%), but they occur almost exclusively in speech utterances (61/62: 98.4%). Finally, prepositional complements also account for a small percentage of the whole data (114/2134: 5.3%), and they tend to appear with cognitive verbs (98/114: 86%), representing 12% of the class (98/814).

4.7. Word Order: Verb-Complement

Table 9 shows word order for verb and complement. Overall, when present (1713/2134: 80.3%), complements tend to appear after the verb (1205/1713: 74.1%), both in cognitive (456/581: 78.5%) and speech utterances (749/1132: 66.2%), although the tendency is larger for the cognitive class. Noteworthy too is the fact that utterances with complement-verb (CV) order occur much more frequently in speech utterances (300/413: 72.6%) than in cognitive utterances (113/413: 27.4%). This would mean that speech utterances frequently occur either with pronominal complements (i.e. they present known information) or after the new information they are presenting. The latter case is particularly interesting because it means the verb is “presenting” information that had already been said; thus, the construction’s presence may be performing other functions.

Another interesting result concerns those utterances in which the verb appears in the middle of its complement; although it is not as frequent when observing all

complements in the data (95/1713: 5.5%), the fact that the phenomenon occurs mainly with verbs of speech (83/95: 87.4%), supports the aforementioned hypothesis about verbs of speech having functions other than presenting information.

Table 9. Verb-complement order			
	Cognitive	Speech	Total
VC	456 (78.5%)	749 (66.2%)	1205 (70.3%)
CV	113 (19.4%)	300 (26.5%)	413 (24.1%)
CVC	12 (2%)	83 (7.3%)	95 (5.5%)
Total	581 (100%)	1132 (100%)	1713 (100%)

4.8. Polarity

Finally, Table 10 shows the distribution of the data based on utterances' polarity; that is to say, whether they are affirmative or negative. As it happens with other features, important differences between the verbal classes are observable even at the general level of analysis. Affirmative utterances are the most frequent type in the data (1744/2134: 81.7%) and for both classes. Nevertheless, they account for a larger percentage of the speech class (1214/1320: 92%) than of the cognitive class (530/814: 65.1%). On the contrary, negative cognitive utterances are four times more frequent than negative speech utterances. These differences suggest each verbal class –or frequent items within them— plays a specific function regarding how speakers state the information they are presenting, which may result on some constructions experiencing loss of syntactic variability.

Table 10. Distribution according to type of utterance			
	Cognitive	Speech	Total
Affirmative Statement	530 (65.1%)	1214 (92%)	1744 (81.7%)
Negative Statement	284 (34.9%)	106 (8%)	390 (18.3%)
Total	814 (100%)	1320 (100%)	2134 (100%)

4.9. Summary

Even though a global analysis does not take into consideration how individual items behave in specific syntactic contexts, patterns do show relevant differences between verbs of cognition and speech in terms of subject and tense (separate and cross-referenced), subject referentiality, type of complement, and polarity. Even though it could be argued that what speakers do with each semantic class depends on the lexical meaning of the verbs in each group (i.e. cognitive verbs convey thinking and opinionating; speech verbs convey oral events), the fact that the same subjects and tenses account for very different percentages in each verbal class, and, moreover, that subject-tense combinations that are frequent in one class are not in the other clearly indicates that such evident argument is, in the best case, inaccurate. How speakers use members of each verbal class, and the constructions emerging from such usage depend largely on communicative intentions and needs pertinent to specific interactions. Thus, these general patterns do not only give us a broad picture of what speakers are doing in the discourse when using cognitive and speech verbs; they also make evident the role of local context for the emergence of specific morpho-syntactic patterns.

Regarding subject, 1sg arises as a very strong exemplar for cognitive verbs because it represents more than half of said semantic class; for speech utterances, there are three strong exemplars (3sg, 1sg, 3pl), although none of them account for more than one third of the category. Thus, despite being based on token frequency in the whole data, these results illustrate meaningful differences about the behavior of these verbal categories suggesting where collocations may emerge from. For instance, considering 1sg subjects account for a large portion of all cognitive verbs and that said person has only

two options for their realization (i.e., Ø or *yo* ‘I’), the combination 1sg-cognitive will likely produce fixed or semi-fixed collocations, especially if certain items have a much higher token frequency than other members of the class⁴⁷.

Regarding subject-tense-class patterns, three combinations emerge as the weightiest ones in the data: 1sg-present-cognitive, 3sg-preterit-speech, and 1pl-present-speech, suggesting that a large part of the interviews concerns present events expressed from the speaker’s perspective. Even though no other subject-tense pattern seems to be particularly relevant when observing the corpus and the verbal classes as a whole, there are indeed some combinations accounting for an important portion of their subject and tense combinations; for instance, 3pl-imperfect-speech is the most the most frequent combination for its tense category (94/335: 28%) and the second most frequent among all 3pl subjects (94/350: 26.9%). Thus, despite its low token frequency in relation to the speech class (94/1320: 7.1%) and the whole data (94/2134: 4.4%), it is not unlikely that a construction may emerge from this particular combination.

Other noteworthy patterns involve both forms of 2sg subjects. As seen in Table 6 above, most 2sg-*tú* subjects appear in present tense utterances (81/103: 78.6%), with a strong tendency towards the cognitive class (64/81: 79%). On the other hand, although 2sg-*usted* also appears mainly in present tense utterances (96/133: 72.2%), its tendency towards the cognitive class (54/96: 56.3%) is not as strong as that of 2sg-*tú*. Thus, the fact that the same group of speakers uses both forms of 2sg differently depending on the utterance’s verbal class suggests the existence of formulaic expressions involving one or

⁴⁷ On the contrary, 3s and 3p subjects would seem less likely to produce fixed construction because they show a larger morphological variability and their referent can be specific, generic, and even abstract entities.

both forms of 2sg⁴⁸. Furthermore, considering that, as a class, verbs of speech are more frequent than verbs of cognition, the fact that their most frequent subject-tense combination is not as predominant in the class, nor in the whole data, suggests more diversity of constructions than that expected for the cognitive class.

For subject referentiality, global patterns suggest that, when using cognitive verbs, utterances concern just the participants in the conversation, either the speaker or the listener. On the contrary, when communicating speech events, utterances mainly involve what other people say.

Differences were observed too for the type of complements that each verbal class take, and the position in which it appears. Thus, when observing the corpus as a whole, speech verbs take a complement more frequently than cognitive verbs do; however, when focusing on specific types of complements, results show that cognitive verbs are complemented by a direct object more often than speech verbs are, implying the former appear in canonical transitive constructions in a larger rate than the latter ones do. The implication is supported too by the fact that the majority of utterances with a complement present a SVO order.

⁴⁸ For instance, Travis (2006a) finds that speakers of Colombian Spanish use *¿sabes qué?* ‘You know what?’ to introduce suggestions that, otherwise, would be interpreted as orders; for New Mexican Spanish, Valdivia (2013; 2015) finds the use of *¿(me) entiendes?* ‘do you know what I mean?’ as a discourse marker to verify if the interlocutor is following the conversation.

Chapter 5. Cognitive Verbs

The present chapter presents and discusses results for cognitive verbs in the data. First, the analysis focuses on general patterns: distribution according to subject, tense, type and realization of the complement, and type of utterance. Such examination seeks to identify patterns that complement those described in Chapter 4 above. Second, based on such results, the analysis focuses on combinations showing certain degree of morphological fixation; for instance, verbs that usually occur with one subject or in a particular tense. Said combinations will be analyzed as well for features such as prosodic distribution, co-occurrence of grammatical or lexical entities, and their function in the discourse.

5.1. Cognitive Verbs: General Patterns

As mentioned in section 4.1 above, cognitive verbs have a token frequency of 814 occurrences distributed among 21 types, which means an average of 38.7 occurrences for item. Nevertheless, only four items have a token frequency larger than, or around, 10% of the class: *saber* ‘to know’ (324/814: 39.8%), *creer* ‘to believe/to think’ (150/814: 18.4%), *acordarse* ‘to remember’ (105/814: 12.9%) and *entender* ‘to understand’ (81/814: 9.9%): together, these four verbs represent 81.1% (660/814) of the class. Two facts worth noticing arise from these results; first, *saber* occurs twice as often as *creer*, which may reflect the type and fixation of patterns emerging from each verb. Second, *pensar* ‘to think’ does not only appear in the fifth position among the most frequent items in the class, but it also accounts for less than five percent of it (34/814: 4.2%). Such a low frequency is important because it suggests English discourse markers such as *I think* and *I don’t think so* have not been calqued into Spanish, which refutes the assumption that

some people may have regarding the varieties of Spanish in the United States being contaminated by constant contact with English.

5.1.1. Cognitive Verbs: Predicate Type

As a class, cognitive verbs tend to appear in utterances with simple predicates (747/814: 91.8%), as shown in Table 11. The remaining utterances are distributed mostly in two types of periphrases with infinitive: one in which a cognitive item is the main verb (28/814: 3.4%), and the other in which it is the auxiliary verb (24/814: 3%). Despite its low token frequency, the latter construction clearly illustrates how lexical items grammaticalize due to language use: a) it is exclusive of the cognitive class; b) it only occurs with *saber* ‘to know’, and c) the cognitive verb conveys ability rather than knowledge. I will return to the construction further in this chapter.

Table 11. Cognitive Verbs by Type of Predicate	
Predicate type	Token frequency
V	747 (91.8%)
Aux + INF _(cognitive)	28 (3.4%)
Aux _(cognitive) + INF	24 (3%)
Other ⁴⁹	15 (1.8%)
Total	814 (100%)

5.1.2. Subject and Tense⁵⁰

Table 12 shows the distribution for cognitive verbs according to subject and tense. For this latter feature, only the three most frequent tenses from the global analysis are listed individually. Following general patterns, most utterances occur with 1sg subjects (430/814: 52.8%); the second most frequent person is 3sg (175/814: 21.5%) and the third most frequent one is 2sg-*tú* (70/814: 8.6%), although with less than ten percent of the

⁴⁹ Includes constructions such as *estar* + adjective (e.g. *estar seguro/a*), periphrases with gerund (e.g. *estar pensando*), etc.

⁵⁰ As earlier mentioned, tense is coded based on the flexed verb or the utterance; thus, data in Table 12 include utterances in which cognitive verbs appear in a non-finite form.

utterances in the class. Regarding tense, only present emerges as relevant for cognitive verbs, since it accounts for four fifths of the class (669/814: 82.2%). Already relevant as the most frequent tense in the whole data, the predominance of present tense among cognitive verbs is even higher because it indicates speakers express their opinion and knowledge framing them in the moment of the speech, even if, as it will be discussed later, they may be talking about events from the past.

Unlike what general patterns for the whole data show, preterit and imperfect tenses are not very recurrent in cognitive utterances. Imperfect appears around ten percent of the time (77/814: 9.5%) and preterit merely 3.6% (29/814); in fact, the occurrence of the latter is lower than that of all the other tenses combined (39/814: 4.8%)⁵¹.

Table 12. Cognitive Verbs by Subject and Tense

	Present	Preterit	Imperfect	Other	Total
1sg	382	11	26	11	430 (52.8%)
2sg-tú	64	0	2	4	70 (8.6%)
2sg-Ud	54	0	1	0	55 (6.8%)
3sg	119	14	25	17	175 (21.5%)
1pl	5	1	0	0	6 (0.7%)
2pl	0	0	2	0	2 (0.2%)
3pl	39	1	20	6	66 (8.1%)
NA	6	2	1	1	10 (1.2%)
Total	669	29	77	39	814

Regarding subject-tense combinations, only two of them emerge as important ones in terms of token frequency; both of them involving present tense, as expected considering the predominance of the tense in this verbal class. 1sg-present (382/814: 46.9%) is not only the most frequent combination in the class, but also for both its subject (382/430: 88.8%) and its tense group (382/669: 57%). The second most frequent combination in the class, although with a lower percentage of occurrences, is 3sg-present

⁵¹ That is to say, excluding preterit and present tenses.

(119/814: 14.6%); it is the second most frequent for present tense (119/669: 17.8%) as well, and it accounts for most 3sg subjects (119/175: 69.2%) in the class. Even though no other combination has a large token frequency in relation to the whole verbal class, there are a couple of them that emerge as the predominant ones within their person group: 2sg-*tú*-present and 2sg-*Ud*-present. The first one represents the majority of *tú* subjects in the class (64/70: 91.4%) and the second one accounts for practically all of *usted* subjects with cognitive verbs (54/55: 98.2%).

Unlike present tense utterances, imperfect tense ones, the second most frequent for the class, do not lean towards a particular subject; rather, they are distributed between three different persons: 1sg (26/77: 33.8%), 3sg (26/77: 33.8%) and 3pl (20/77: 26%). Based on such distribution and token frequency, it is unlikely to find different constructions, if any, showing loss of syntactic properties; at least not to the same degree than those arising from 1sg-present combinations. The same can be expected from preterit utterances; even though two combinations account for most of the tense, 3sg-preterit (15/29: 51.7%) and 1sg-preterit (11/29: 37.9%), they represent a very small percentage in the whole cognitive class: 1.8% (15/814) and 1.6% (11/814) respectively.

5.1.3. Subject and Predicate Type

Table 13 shows the correlation between predicate type and subject. As it can be seen, the predominance of predicates with a single verbal nucleus persists across practically all subjects. In fact, for some persons, the pattern represents almost all of the utterances; for instance 2sg-*tú* (68/70: 97%) and 2sg-*Ud* (54/55: 98.1%).

Following general patterns for subject, when functioning as single nucleus, cognitive verbs appear primarily in 1sg (406/747: 54.8%), and then in 3sg (175/747:

23.4%). Accordingly, 1sg-V is the most frequent pattern for predicates in the cognitive class (406/814: 49.5%) and for 1sg-cognitive utterances (406/430: 94.4%). 3sg-V is the second most frequent pattern for predicates in the class (147/814: 18%) and the most frequent within its person group (147/175: 84%).

Table 13. Cognitive Utterances: distribution by subject and type of nucleus

	1sg	2sg- <i>tú</i>	2sg- <i>Ud.</i>	3sg	1pl	2pl	3pl	NA	Total
V	406	68	54	147	3	2	57	10	747
Aux + INF _(cognitive)	12	1	0	10	3	0	2	0	28
Aux _(cognitive) + INF	4	1	0	12	0	0	7	0	24
Other	8	0	1	6	0	0	0	0	15
Total	430	70	55	175	6	2	66	10	814

Although barely significant in relation to the whole class, ‘aux + INF_(cognitive)’ predicates follow general patterns for subject; they occur mainly with 1sg (12/28: 42.9%) and 3sg (10/28: 35.7%). Rather sporadic as well, ‘aux_(cognitive) + INF’ predicates, appear usually with 3sg (12/24: 50%) and then with 3pl subjects (7/24: 29%). As mentioned earlier, this pattern always involves the verb *saber* ‘to know’; so, it is noteworthy that this is the only type of predicate which does not follow general patterns for subject: speakers use a periphrasis when expressing what other people know how to do, but not as frequently when talking about what they know how to do.

5.2. Cognitive Verbs: Local Patterns

Following general patterns for the whole data and the class, the two most frequent cognitive verbs tend to appear in 1sg, present tense, and simple predicate utterances. *Saber* ‘to know’, the most frequent lexical item (324/814: 39.8%), accounts for 40 % (172/430), 29.9% (243/814) and 35.7% (291/814) of those categories, respectively. With a lower token frequency, *creer* ‘to think / to believe’ accounts for 18.4% (150/814) of the

class, but it presents stronger patterns than those of *saber*. It mainly appears in 1sg (112/150: 74.7%) utterances, and predominantly in present tense (142/150: 94.7%) with simple nucleus utterances (148/150: 98.7%).

5.2.1. 1sg-Cognitive Constructions

Cross-referencing the aforementioned features, the item *sé* ‘I know’ emerge as the most frequent for *saber* with 44.1% of the occurrences (143/324), usually in negative utterances: *(yo) no sé* ‘I don’t know’. The latter combination is the most frequent in the class (123/814: 15.1%), its subject (123/172: 71.5%) and its predicate type (123/291: 42.3%). In fact, it accounts for the majority of negative utterances with *saber* (123/187: 65.8%), negative 1sg-cognitive utterances (123/197: 62.4%), and a substantial portion of all negative utterances in the cognitive class (123/281: 43.8%). Closely behind in token frequency, *creo* ‘I think’ show stronger tendencies. Its 111 tokens account for practically all 1sg-cognitive occurrences (111/112: 99%), most present tense (111/142: 78.2%) and most simple predicate utterances in the class (111/148: 75%). Based on these patterns, both constructions represent good candidates to undergo structural fixation.

5.2.1.1. *(yo) no sé* ‘I don’t know’

This collocation appears without an expressed subject in the majority of the utterances (72/123: 58.5%), as illustrated by (159). When expressed (51/123: 41.5%), the pronoun *yo* (‘I’) always appear in preverbal position (160). Thus, even though *yo no sé* is less frequent than *no sé*, the fixed position of the pronoun suggests loss of syntactic mobility, which may promote the development of the unit into a collocation and, even, a discourse marker.

- (159) E: *Y dónde nació su papá.*
 U: *Creo que nació en Buena Vista también.*
 E: *(H) Y su mamá?*
 U: *Ella yo creo que --*
*yo creo que también **no sé**=,*
no me acuerdo [donde=] XX – (214-1A1: 156)
 ‘E: And, where was your dad born? U: I think he was born in Buena Vista as well. E: And your mom? U: She, I think that... I think [she was born] there too, **I don’t know**, I don’t remember where...’
- (160) A: *Los papás de él,*
de aquí también?
 X: *Mhm.*
 A: *De --*
de= --
de aquí= de Albuquerque?
o de otra parte también.
 X: *Pues,*
yo no sé=.
yo sé que tenían gente allá en .. Kansas.
.. Yo no sé--
Pero yo,
todo el tiempo .. supe que mi grandma y mi grandpa vivían aquí. (47-1A2: 684)
 ‘His parents, are from here too? X: Mhm. A: From..from here, Albuquerque? O from other place too? X: Well, **I don’t know**. I know they have relatives over there in Kansas. **I don’t know**... But I, I always knew that my grandma and my grandpa lived here.

In both cases (i.e. with or without *yo*), the construction tends to take direct objects (83/123: 67.5%), usually clauses (61/83: 73.5%) in post-verbal position with a subordinating conjunction (*si* ‘whether’) or an interrogative (60/83: 72.3%), as illustrated by (161) below.

- (161) L: *[Y no] podía.*
por mi pie de lo que me dolía.
Pues,
ésta en la --
en la mañana me dice,
pues,
qué le pasó.
Yo no sé qué me pasó,
le dije,
mira mi pie. (219-1B1: 61)
 ‘And I couldn’t [walk] because my foot hurt. Well, in the... this morning she asked me: “what happened to you”. **I don’t know** what happened to me, I told her, “look at my foot”.

Based on these canonical patterns, one may feel inclined to think that *(yo) no sé* ‘I don’t think’ conveys propositional (i.e. objectives) values derived from its lexical meaning: lack of knowledge regarding the information the speaker is asked about. Nevertheless, when observing the information introduced by the collocation, one finds features that suggest the construction actually introduces subjective nuances. For instance, while in (163) the speaker uses the collocation to actually express that he does not know where his father’s relatives were from, in (164) the event introduced by *yo no sé* appears in future perfect tense, which introduces a nuance of inference instead of certainty, and seems to convey a tone of politeness. If the verb were instead in present perfect (i.e. *ha oído* ‘you have heard’), the utterance would sound more direct and thus less polite. In (165), the interviewer starts a statement (*usted les – los* -- ‘you [control] them’) but then he truncates it. Later, when reformulating his statement, he uses *no sé* ‘I don’t know’ to present his idea as a doubt rather than an assertion. By doing so, he avoids calling his interviewee ‘bossy’, which might have been interpreted as disrespectful and, consequently, it would have represented an obstacle for the conversation. The strategy is successful and the interviewee does not only agree about her giving advice to her family, but she also takes the interviewer’s original idea and states that she indeed often controls her family. Thus, even though both the argument structure and the syntactic order of the utterances in which *(yo) no sé* occurs suggest that such collocation is used to convey propositional meaning, in most cases the collocation co-occur with features that point to the opposite direction: the construction does actually, at least to a certain degree, convey speakers’ attitudes regarding the information being presented.

- (162) P: *Se] puede ir por Bernalillo áhi por la --*
el eh= --
por el --
Placitas?
A: *Oh, cierto.*
P: *[Mhm.*
Mhm].
A: *[Una vez nos fuimos] por allá también.*
P: *.. Y luego a Phoenix.*
.. No s- --
% .. No sé qué pasó áhi.
.. Estaba muy chiquillo.
y luego me .. trajeron pa' Albuquerque XX. (88-1A2: 76)
‘P: One can go through Bernalillo, there... through Placitas. A: Oh, that’s right. P: Mhm. Mhm. A: Once, we went through there too. P: And then to Phoenix. I d- **I don’t know** what happened there. I was just a child. And then they brought me to Albuquerque XX.’
- (163) T: *but I think that he made some trips to Kansas City.*
F: *[Oh sí]?*
T: *[oh not] Kansas City.*
eh to Kansas.
F: *Kansas.*
T: *eh,*
Western Kansas.
eh no sé si a Dodge City Kansas,
o cuál,
you see.
F: *Oh. (270-1B2: 579)*
‘...but I think that he made some trips to Kansas City. F: [Oh sí]? T: [oh not] Kansas City. eh to Kansas. F: Kansas. T: eh, Western Kansas. Eh **I don’t know whether** to Dodge City Kansas or somewhere else, you see. F: Oh.’
- (164) *Y la comunidad, siempre tenían alguien que era mejor para hablar que otros?*
J: *... Posible que sí=.*
Yo creo que los maestros eran los que --
Por ejemplo en el valle del XX Ortega,
yo no sé si.. tú habrás oído de él.
(H) *Ése era uno que se podía expresar,*
..you know,
bien. (144-4A2: 133)
‘And in the community, was always someone who spoke Spanish better than most people? J: Maybe. I think the teachers were the ones who... For instance, the Valley where Ortega is from... **I don’t know** if you have heard about him; he could speak well, you know?’

- (165) L: *A mí, .. todavía mi familia no me gobierna.
me gobierno sola.*
E: *Qué bueno.*
L: *@@[@@]*
E: *[Qué] bueno.
@@ Qué bueno.
Más bien,
(H) usted les --
los --
No sé si usted los gobierna a ellos pero,
.. de cierto les da consejos.*
L: *.. Y los gobierno un tanto. (219-1B1: 192)*
‘To me... my family does not control me yet. I do what I want. E: That’s good. L: [Laughs]. E: That’s good. That’s good. Rather, you ... **I don’t know** if you control them, but, surely, you give them advices. L: And I control them to a certain degree.’

Furthermore, the fact that in a third of the occurrences (40/123: 32.5%) the collocation does not introduce a direct object provides more evidence for the hypothesis that speakers are using the verb to convey meanings beyond propositional statements, practice from which novel patterns may be emerging. As it can be seen in (160) above, none of the two instances of *(yo) no sé* is the syntactic head of the utterance expressing what the speaker does not know, as evident in the lack of a subordinating conjunction or an object pronoun (*lo* ‘it’). The fact the verb does not follow its canonical transitive pattern seems related to the speaker’s intention to lessen the importance of the unknown event and rather focus on what he takes as a true thing, therefore more important for him, that his grandparents lived in Albuquerque.

Evidence for the development of *(yo) no sé* into a discourse marker comes too from the position of the collocation in its utterance. In many cases, the chunk containing *(yo) no sé* appears inside a larger utterance with which it does not really form a syntactic unit; for instance, the speaker starts a statement, he/she then interrupts it and introduces the utterance with the collocation to express lack of knowledge about some element of the information being presented or about to be mentioned; by doing so, rather than

merely expressing lack of knowledge, the speaker is indicating he/she is not completely sure about the information and, therefore, cannot be held responsible if something is not accurate or, even, true. In (166), for instance, while talking about the violence in Albuquerque, the speaker mentions that his mother had read something in the paper, and he adds the phrase *o no sé qué estaba leyendo* ‘lit. Or I don’t know what was she reading’ before reporting what his mother said about the news she had read. With the phrase, the speaker implies that where his mother read the news is not important; that information is relevant just because it functions as a setting to bring the interlocutor’s attention to the mother’s opinion.

- (166) W: *the other day,*
ah uh --
on the way to visit my mom,
 B: *[Mhm].*
 W: *[you know]?*
y estaba ella --
% she had read in the paper,
*o **no sé qué estaba leyendo**,*
de las cosas que estaban pasando ahora.
(H) y dijo,
dijo mi mamá,
<VOX hijo,
sure I'm glad que yo no estoy --
(H) ..putting any kids to school now VOX>.
dijo,
porque mira aquí lo que está pasando en Albuquerque. (190-3B2: 852)
 ‘the other day, ah uh -- on the way to visit my mom, B: Mhm. W: You know? And she was... she had read in the paper, or **I don’t know what she was reading**, about what is happening; and she said, my mom said, “son, sure I’m glad that I’m not putting any kids to school now”, she said, “look what is happening in Albuquerque.”

A similar situation can be observed in (167), where the speaker talks about different ways to get to a destination; then, trying to explain why his family moved from Phoenix to Albuquerque, he says *no sé qué pasó ahí. Estaba muy chiquillo* (‘I don’t know what happened there. I was just a kid’). Again, although by itself the collocation seems to

merely convey lack of knowledge, in the context of the whole utterance, it conveys the speaker's attitude; in this case, some sort of apology for not being able to provide an accurate recount of his story.

(167) P: [*Se puede ir por Bernalillo áhi por la*
 -- *el eh=*
 -- *por el*
 -- *Placitas?*

A: *Oh,*
cierto.

P: [*Mhm.*
Mhm].

A: [*Una vez nos fuimos*] *por allá también.*

P: .. *Y luego a Phoenix.*

.. *No s- --*

% ..**No sé** *qué pasó áhi.*

.. *Estaba muy chiquillo.*

y luego me .. trajeron pa' Albuquerque XX. (88-1A2: 76)

'P: It is possible to go through Bernalillo; by Placitas? A: Oh, that's true. P:

Mhm. Mhm. A: We went that way once too. P: And then to Phoenix. I don't..

I don't know what happened there; I was just a kid. And then they brought me to Albuquerque.

The discursive function of (yo) *no sé* is suggested as well by its prosodic distribution regarding other elements of the utterance. First, the collocation often appears in its own intonation unit (48/123: 39%), either because there is no complement for its verb, as in (168), or because it closes an idea that the speaker does not want or can to continue, as in (169) and (170). The latter example makes a strong case for the syntactic freedom of (yo) *no sé* because the speaker's apparent lack of certainty has already been established twice, first by *seguro que no había trabajo* 'for sure there were no jobs' and then by *yo creo* 'I think', which, intonationally, closes the speaker's statement. Thus, the inclusion of *yo no sé*, reinforces such negation, without affecting the syntax of the utterance.

- (168) X: *Mi daddy era de aquí.*
 .. *Y mi mamá era –*
 A: *Los papás de él,*
de aquí también?
 X: *Mhm.*
 A: *De --*
de= --
de aquí= de Albuquerque?
o de otra parte también.
 X: *Pues,*
yo no sé=.
yo sé que tenían gente allá en .. Kansas.
.. Yo no sé --
Pero yo,
todo el tiempo .. supe que mi grandma y mi grandpa vivían aquí. (47-1A2: 686)
 ‘X: My daddy was from here, and my mom was – A: His parents were they from here too? X: Mhm. A: From, from, from here, Albuquerque? Or somewhere else? X: Well, **I don’t know**. I know they have relatives over there.. in Kansas. I don’t know... but I, all the time, knew my grandma and my grandpa lived here.’
- (169) K: *Casi hablo mucho en inglés,*
porque toda mi familia habla en inglés,
you know.
 A: *Mhm.*
 K: *Ahora mismo notas que cada rato te digo something en inglés porque,*
ya le viene a uno de= --
yo no sé.
a –
...hay veces que estoy platicando con= somebody que no sabe en --
en % inglés.
cuando menos acuerdo les he dicho en inglés,
y luego tengo que decir en mexicano. (318-1A3: 543)
 ‘K: Pretty much, I speak in English because all my family speaks in English, you know. A: Mhm. K: Right now you can see I often say something in English because it comes to my mind. **I don’t know**. Sometimes I am chatting with somebody who doesn’t know English and, without realizing, I say something in English and then I have to say it again in Mexican [i.e. Spanish].’
- (170) K: *.. Él vino de la guerra.*
él fue la primera guerra.
 A: *Mhm.*
 K: *y luego,*
seguro que no había trabajo,
yo creo.
Yo no sé. (318-1A2: 105)
 ‘K: He came from the war; he went [to] World War I. A: Mhm. K: Then, surely there were not jobs, I think. **I don’t know**.’

Second, when in its own intonation unit, *(yo) no sé* tends to be uttered with final transitional continuity, as in (168) - (170) above. Moreover, with such transitional feature, the collocation often appears at the end of the utterance, as in (171) below.

- (171) A: *y qué tal habla=n --*
 qué % tal le parece usted que habla español.
 X: .. *Qué tal?*
 A: *Mhm.*
 .. Qué tan bien lo habla.
 X: .. *No muy bie=n.*
 Se me hace a mí.
 yo no sé.
 @@
 A: *En comparación con su=s abuelitos=.*
 X: *Oh.*
 Oh yeah.
 ellos hablan mejor que yo. (47-1A3: 17)
 A: ‘And how well do they... how well do you think is your Spanish. X: How well? A: Mhm. How well do you speak it? X: Not very well. It seems to me. **I don’t know.** A: Compared to your grandparents. X: Oh, yeah. They speak it better than I do.’

When the collocation shares the intonation unit with its complement or another lexical unit, it usually presents or refers to known information, so its purpose is not to express lack of knowledge or awareness but to remark the importance of what is already known. In (172), for instance, both the speaker and her daughter know that there is something wrong with the speaker’s leg because they can see it is all bruised. Thus, when the speaker states *yo no sé que me pasaría*, with the subordinate verb in the conditional form (i.e. *pasaría* ‘lit. ‘it would happen’), she is asking a rhetorical question rather than expressing real doubt, as it would be the case if she had used the preterit form (*pasó* ‘it happened’) one. That is to say, even if the speaker does not know why her leg is all bruised, her intention is not to communicate such lack of knowledge, but to convey surprise regarding how bad her leg looks, thus remarking that something is wrong. Once again then, there are elements allowing *yo no sé* to express discursive nuances beyond the

lexical meaning of *saber*; even though it forms a syntactic unit with the clause it introduces and, therefore, it is the sentence's main verb.

- (172) L: *Otro día en la mañana,*
 E: *Mhm.*
 L: *Me levanté teniéndome,*
y me dice ésta,
<VOX pues,
qué le pasó mamá VOX>?
Y le digo,
yo no sé qué me pasaría.
mira,
mi pie estaba azul [azul],
 E: *[Hm].*
 L: *todo mi pie.*
 E: *Mhm.*
 L: *Hu=,*
se asustaron tanto. (219-1B1: 61)
 'L: The other day in the morning. E: Mhm. I got up from bed holding me [from the furniture] and she told me: "What happened to you, mom?" I told her: "**I don't know** what might have happened to me. Look!" My foot was all bruised. E: Mhm. L: My whole foot. E: Mhm. L: They got all scared.'

Like in the previous example, in (173) *yo no sé* syntactically functions as the sentence's main verb; this time, though, the subordinate verb is in the indicative mood rather than in subjunctive. By choosing the realis domain over the irrealis, the speaker is not merely communicating his lack of knowledge, but also stating, and stressing, a fact: bears often come down to town.

- (173) C: *So,*
era muy peligroso más antes?
 I: *Oh sí.*
.. Pues,
.. había muchos osos silvestres.
.. Bajaban al --
.. Todavía bajan.
 C: *(COUGH)*
 I: *al valle.*
.. Yo no sé porqué los osos bajan tanto al valle.
.. Donde hay casas. (142-3B2: 424)
 'C: So, was it very dangerous before? I: Oh, yes. Well, there were a lot of wild bears. They used to come down to, they still do, to the valley. **I don't know** why bears come down to the valley that often, where there are houses.'

Finally, there are a few cases in which independent collocations seem to be emerging from (yo) *no sé*: *no sé cómo* (3 occurrences) and *o no sé qué* (3 occurrences). Although it is not very frequent in this particular data and it is produced by the same speaker, it is worth noticing that *no sé cómo* appears with unexpressed subject and in its own intonation unit. Regarding its function, speakers seem to use the construction as a discursive mechanism to “buy time” while looking for the best way to express what they want to say.

- (174) O: *I never --
that's what I tell --
that's what I tell my children now.
A mis hijas.
Y yo les digo, l- --
l- --
l- --
l- --
les digo,
you know.
.. Les digo,
no sé cómo,
Que me dicen que me miro como me miro?
.. You know?*
- A: *Pues sí.
@ <@ como le acabo de [decir] @>.*
- T: *[She's young looking].*
- O: *Sí.
no sé cómo,
you know,
porque yo no lo puedo imani- maginar. (117-1A3: 345)*
- ‘O: My daughters, I tell them, ... I tell them, you know, **I don't know how** I speak, but they tell me that's why I look how I look, you know? A: Yeah. Like I just told you. T: She's young looking. O: Yes, **I don't know how**, you know, because I cannot imagine it.’

Besides (yo) *no sé*, the only other somehow noteworthy collocation with 1sg-*saber* is (yo) *sé* ‘I know’, although it shows a much lower frequency (n= 19). The collocation accounts for 14% (20/143) of all 1sg-*saber* utterances, 8.2% (20/243) of all *saber* items in present tense, and barely 2.5% (20/814) of all utterances with cognitive verbs. It tends to appear with expressed subject (15/20: 75%) in preverbal position (13/20: 65%), and it usually takes an objective complement in postverbal position (13/20:

65%). Said complement is introduced with *que* ‘that’ and shares the intonation with (*yo sé* most of the time (11/13: 84.6%). When *yo sé* introduces a clause, in addition to state knowledge of a fact, the speaker offers an explanation for something he/she has just said or is about to say, as in (175) and (176) below. When the complement is not a clause, it usually refers to a language (e.g. *español* ‘Spanish’, *los dos idiomas* ‘both languages) or proverbs.

(175) W: *y luego me fui con una de las agencias de .. Buick Oldsmobile, Cadillac.*

en todas esas (H) me ha ido levantando.

pero,

.. % ha apro- --

he querido aprovechar,

.. todo lo que ellos .. han tenido ..pa' ofrecerme.

B: *.. [Mhm].*

W: *[y por eso],*

eh digo yo y,

y yo sé que .. te vas adelantando,

% % te ayuda mucho,

en otras palabras.

B: *Mhm. (190-3B2: 105)*

‘W: Later I went to work to one of those car dealers, Buick Oldsmobile, Cadillac. In all of those places I have been promoted; but [to do so], I have wanted to take advantage of all they have had to offer. B: Mhm. W: That’s why I say that... and **I know** you make progress... In other words, t [going to college] helps you a lot. B: Mhm’.

(176) B: *Usted era alguacil mayor.*

...Cómo eran --

.. Cómo eran las elecciones?

que --

que= --

...Us- --

.. La gente salía a votar así --

...Yo sé que mi papá siempre ayudaba a los --

...iba,

y los pepal- - .. pepenaba a la gente que .. fuera a votar,

(H) cosas así.

(H) Se hacía eso aquí también?

Z: *Yeah.*

yo hacía mucho eso. (311-3B2: 15)

‘B: You used to be a Sheriff. How were the elections in that time? What, what.. Did yo--.. People went to vote just like that... **I know** that my dad always helped those.. he used to go and take people to the polls, things like that. Did that happen here too? Z: Yeah. I used to do that a lot.’

- (177) *Yo estoy hablándote aquí en mexicano a ti.*
It's too bad.. si ella no me entiende a mí.
porque si ella habla en inglés yo la voy a entender a ella.
porque yo sé los dos. (117-1A3: 869)
 'I'm here, talking to you in Mexican. It's too bad if she cannot understand me because, if she speaks in English, I'm going to understand her, because **I know** both languages.'
- (178) D: *Y no se acuerda usted de -- de refranes o,*
o= sayings no.
 S: *Me acuerdo [muchos].*
 D: *[dichos].*
 S: *sé muchos refranes pero,*
 D: *Pues así no más pa' [decirlos, no].*
 S: *[S=i se] a toca,*
Cuando se a toca la ocasión,
me acuerdo de ellos [no]? (102-3A2: 151)
 'D: Don't you remember refrains or sayings? no?. S: I remember a lot of them. D: sayings. S: **I know** a lot of sayings, but... D: Just to say them, no? S: If necessary, when there is a suitable situation, I remember them, no?

5.2.1.2. *(yo) creo* 'I think / I believe'

The second most frequent 1sg-cognitive collocation, *(yo) creo* 'I think / I believe' represents 13.6% (111/814) of the class, 25.2% (111/430) of all 1sg-cognitive, and 16.6% (111/669) of all present tokens with a cognitive verb. Unlike *(no) sé*, it occurs almost exclusively in affirmative utterances (108/111: 97.3%) with expressed subject always in preverbal position (92/111: 82.9%): *yo creo* 'I think / I believe'. Considering pronominal subjects in Spanish are only required in certain conditions, the predominance of *yo creo* over *Ø-creo* suggests speakers either use *yo creo* in contexts where the pronoun is needed or to introduce subjective nuances in the discourse. Both situations are illustrated in (179) below. The first item, line 1, seems unnecessary since there is no change of subject, nor contrast. On the contrary, in line 6 the presence of *yo* seems to mark a contrast between the interviewer and the interviewee's utterances.

- (179) 1. Z: ***Yo creo*** *que ni me enregistré,*
 2. *No -- no estoy seguro.*
 3 O: *No.*
 4 *tenía que haberse registrado.*

- 5 B: *Me [parece que sí].*
 6 Z: *[Yo creo que] [2sí2].*
 7 O: *[2No2],*
 8 *era [política].*
 9 Z: *[Pero anyway].*
 10 *no -- no me llamaron nunca. (311-4A1: 124)*
 ‘Z: **I think** that I didn’t register, not even that. I’m not... I’m not sure. O: No. You had to register [for the army draft]. B: I think so. Z: **I think so.** O: No, it was mandatory. Z: But anyway, I was never called for service duty.’

In more than half of the cases (58/108: 53.7%), *(yo) creo* takes a direct object complement, usually a clause (46/58: 79.3%) in postverbal position, and often sharing the intonation unit with it (38/46: 82.6%), as in (180) below. Based on these features, it can be argued that the construction constitutes a prefab because the verbal item is experiencing loss of morpho-syntactic variability, but it has not cancelled yet the slot for a direct object. Such a prefab may be described as *(yo) creo que* + CLAUSE.

- (180) L: *Porque no te- -- había explicación ni en -- en ellos.*
Y asina serían ellos, pero digo que lo aprendí mejor yo que ellos.
 E: *Sí.*
 L: *@@*
 E: *<@Sí @>.*
Cierto.
Pues,
creo que toda= su generación llegaron hasta el libro tres.
Eso era --
parecía ser lo normal. (219-1A1: 180)
 ‘Because they don’t... there was no explanation, not even for them. They might have been just like that, but I said I learned it [English] better than they did. E: Yes. L: (Laughs). E: Yes. That’s true. Well, **I think** all his generation studied until third grade. That seemed to be the usual.’

Regarding the prefab’s place in a continuum of subjectivity, the fact that its argument structure and its prosodic distribution are related to the lexical meaning of *creer* suggests is closer to a propositional construction than to an epistemic one: speakers use the *(yo) creo* to communicate the literal meaning of the verb: they believe that the information they are presenting is true. The claim seems to be supported by the predominance of tenses conveying factual meanings.

- (181) A: *Y nunca pensó aca- --
nunca ha hecho por acabar [su <X Masters X>]?*
Y: *[No. ya después],
.. ya no.
Ya no quise.
Ya no.
yo creo que ya no puedo estudiar.*
A: *[@@@@]*
Y: *[ya no está] mi mente pa' estudiar. (147-1A2: 471)*
'A: Have you ever thought about fin--- Have ever tried to finish your Master program? Y: No. Later... not anymore. I didn't want it anymore. **I think I cannot study anymore.** A: (Laughs). Y: My brain is not fit for studying anymore.'

Nevertheless, there are some cases in which development of subjectivity is observed because the speaker introduce a nuance of assertion into the discourse; the speaker's intention is not to merely express something she believes to be truth, but to present an argument or explanation for the event being discussed, argument that she expects the interlocutor to accept. In (182), for instance, *creo que* introduces an assertion that closes a whole criticism about modern youth; by doing so, the speaker is stating her opinion as a fact, so the interlocutor does not have an alternative but to accept it. A similar situation can be observed in (183), where the speaker first mentions that she has never been elected as the festivities' queen, and then she uses *yo creo que* to state that she will never be: rather than conveying that she believes such thing to be truth, she is asserting her point of view.

- (182) J: *Uno de los cambios que .. he notado yo es que la juventud de este día,
no se ocupan en ninguna cosa --
ahm,
... útil,
... y=,
y hm,
y tiene mucho --
mucho --
mucho tiempo en el cual no tiene ninguna cosa que hacer y,
.. y= --
y este,
están malgastando su tiempo.
me parece a mí.*

- [you know].
 C: [*<X De veras X>*].
 [Y=] --
 J: [Y] **creo que** esa es una de las razones porque hay tanto crimen y,
 y,
 y luego porque también no tienen --
 (H) no tienen --
 ahm,
 ahm,
 .. lo --
 ... El deseo de --
 de educarse. (144-4B2: 56)
 ‘J: One of the changes that I have noticed is that today’s youth are not interested in doing anything useful. They waste a lot of time doing nothing. They wasting their time, I think. You know? C: That’s true. J: And **I think** that is one of the reason why there is so much crime. Also, they do not have any desire to get an education.’
- (183) K: *ya se va a hacer cuarenta y siete <@ años que estoy @>*.
 A: @@@@
 K: *pero no me --*
no me chusan yo creo que ni --
.. ni me a van a chusar nunca.
 A: *Pero por qué?*
Porque es [mexicana]?
 K: [Porque] soy mexicana yo creo. (318-1A2: 165)
 ‘K: It’s almost forty seven years that I have lived here. A: (Laughs). K: But they do not choose me. **I don’t think** they will ever choose me. A: But why? Because you’re Mexican. K: Because I’m Mexican, I think.’

Slightly behind in token frequency, speech complements appear in 41.7%

(45/108) of all (yo) *creo* utterances. Formally, the main difference between this type of complements and the direct object one is the parenthetical relation with the verb, as illustrated in (184) below. Given the semantic nature of cognitive verbs, it could be argued, to be more precise, that the segment of the utterance to which the verb refers is not reporting a speech event but the speaker’s thought (see 3.3.9 above). Nevertheless, the construction allows speakers to comment on something they just mentioned or are going to mention, rather than just expressing a state of mind.

- (184) L: *Pero ya como de --*
.. de ocho años tenía yo,
yo creo,
ya comenzaron los libros a usar el --
en inglés. (219-1A1: 392)
 ‘L: But when around... when I was eight years old or so, **I think**, English books started to being used [at school].’

The most immediate consequence of the parenthetical relation between the verb and the information it refers to is that the former acquire the flexibility to appear in positions other than the beginning of the utterance, as in (184) above and (185) below. Such reduction of syntactic scope is important for the arguments of subjectivity and subjectification because it allows the verb to affect just one part of the utterance, namely, the one the speaker wants to present as doubtful, uncertain or inaccurate. For instance, in (184) above, by inserting *yo creo* in the middle of his utterance, the speakers is not stating something he believes to be true, but conveying his lack of certainty, thus implying he is not responsible for the accuracy of the information. Such epistemic function is even more relevant when the construction appears at the end of the utterance, as a hedge marker reducing the assertive tone of a statement. Said communicative intention is illustrated in (185) below.

- (185) U: *no está lejos de ahí de --*
Flagstaff?
 E: *No no está muy lejos.*
 U: *Ah=.*
 E: *<P No está muy lejos P>.*
 U: *two hundred y algo millas?*
[o menos]?
 E: *[Mhm].*
No me acuerdo cuántas millas [2pero2] –
 U: *[2two hundred2] and ninety yo creo,*
qué no? (214-1A1: 95)
 ‘U: Is it far from Flagstaff? E: No, it is not very far. U: Ah. E: It is not very far. U: Two hundred and something miles? Or is it less? E: Mhm. I don’t remember how many miles, but... U: Two hundred and ninety, **I think**, right?’

Given the syntactic, intonational and, especially, discursive implications associated with the type of complement introduced by the verb, it can be argued that two related constructions had emerged from (yo) *creo* + complement. The first one, the aforementioned prefab (yo) *creo que* + CLAUSE, leans toward the propositional pole of an objective-subjective continuum, although in some instances it shows certain degree of subjectivity; the second one, (yo) *creo*, is closer to subjective end of the continuum and it even shows some characteristics associated to discourse markers; namely, its loss of syntactic properties as a main verb and the fact that in some cases its use seems dispensable because speakers had already expressed their thought.

- (186) *O la autoridad que te- --
que tuvo la idioma inglés,
porque la idioma inglés entró,
creo,
y de una vez comenzó a prevalecer.
sobre todo,
you know. (144-4B2: 515)*
‘Or the authority the English language had, because the language came, **I think**, and then, immediately, it started to prevail about everything, you know?’
- (187) *y tuvieron -- they had that program por muchos años.
tuvieron bilingual. and ah --
y luego eh= --
hace como dos años,
yo creo,
o tres,
que lo quitaron aquí que sepa –
que sé yo. (117-1A2: 386)*
‘and they had... they had that program for many years. They had a bilingual one. Then, about two years ago, **I think**, or three, they cancelled it, as far as I know, as far as I know.’⁵²

Furthermore, the non-propositional value of (yo) *creo* is permeating cases in which the verb is followed by a conjunction introducing the adverbs *sí* and *no*. Thus, although

⁵² In the expression ‘*Que sepa yo*’, the verb is in the subjunctive form; in *que sé yo* it is in the indicative form. In both cases, the speakers is implying the information she is communicating is what she knows, but it is possible than something new has occurred and she has not realized.

the construction refers to a clause previously mentioned in the discourse, speakers use (yo) *creo que sí/no* ‘I think so / I don’t think so’ to convey agreement or disagreement, or to emphasize an opinion they have just stated, rather than to merely indicate they believe something to be true or false. The phenomenon is illustrated in (188) below. Although there are not enough data in the corpus to determine whether this construction is ongoing syntactic fixation or whether the occurrence of *yo* ‘I’ indicates change of subject, it is worth mentioning that speakers always include the pronoun when using *creo* to answer a question or to confirm an interlocutor’s comment, as in the example below.

- (188) Z: *Yo creo que ni me enregistré,*
No -- no estoy seguro.
 O: *No.*
tenía que haberse registrado.
 B: *Me [parece que sí].*
 Z: *[Yo creo que] [2sí2].*
 O: *[2No2],*
era [política].
 Z: *[Pero anyway].*
no -- no me llamaron nunca. (311-4A1: 123)
 ‘Z: I think that I didn’t register, not even that. I’m not... I’m not sure. O: No. You had to register [for the army draft]. B: I think so. Z: **I think so.** O: No, it was mandatory. Z: But anyway, they never drafted me.’

5.2.1.3. (yo) *me acuerdo* & (yo) *no me acuerdo* ‘I (don’t) remember’

Two 1sg-present-V constructions emerge from *acordarse* ‘to remember’ (n= 105), the third most frequent verb in its semantic class. Both constructions have practically equal token frequencies: (yo) *me acuerdo* ‘I remember’ appears 34 times, and (yo) *no me acuerdo* ‘I don’t remember’ occurs 32 times. Naturally, they account for similar percentages of all *acordarse* utterances (32.4% and 30.5%, respectively), and of all 1sg-present-V constructions with said verb (34/66: 51.5% and 32/66: 48.5%). I argue for the existence of two different constructions, rather than just a general one, in which the polarity would depend on the discourse’s context, because of the features each one of

them presents, the more definitive one being polarity; while *(yo) no me acuerdo* accounts for 11.4% (32/281) of all negative utterances in the class, *(yo) me acuerdo* does it for just 6.3% (34/533) of all affirmative utterances with a cognitive verb.

On the topic of subject, unexpressed ones are much more frequent with *(yo) no me acuerdo* (23/32: 72%) than with *(yo) me acuerdo* (17/34: 50%); for expressed subject, both collocations show a rather small preference for the pre-verbal position⁵³.

Unfortunately, the low occurrence of these collocations does not allow to provide a better description of their behavior on this topic.

(189) J: *Y le dijeron una palabra que no entendía y fue a --
a --
a ver a Rubén .. Murillo,
que le explicara qué era lo que quería decir la palabra esa.*

C: [yeah].

J: [me] acuerdo yo.

C: *Y no te acuerdas de la palabra?*

J: *No.*

C: [Yeah].

J: **[no] me acuerdo** qué era la palabra. (144-4A2: 174)

‘J: And [my dad] was told a word he didn’t know, so he went to see Ruben Murillo for him to explain what the meaning of that word was. C: Yeah. J: I remember. C: Do you remember the word? J: No. C: Yeah. J: **I don’t remember** what word was.

(190) T: *ah --
... Aquí hablaban mucho del año del ochenta y ocho,
que vino una nevada que acabó con los animales y venados y todo <X berrendos
X>.*

F: *O=h.*

T: *y luego=,*

.. En mil novecientos dieciocho estuvo muy frío.

*.. **Me acuerdo** que el a-*

-- el agua del Plan de la Noria se helaba. (270-1B2: 334)

‘T: Ah.. Here, they talked a lot about 88 [1998]; [they say] a blizzard came which killed all the animals, and deer, and all the elks. (elks = unclear). F: Oh. T: And then, 1918 was a very cold year. **I remember** that in, *Plan de la Noria*, water would freeze.

⁵³ *yo me acuerdo* (n= 10), *me acuerdo yo* (n= 7); *yo no me acuerdo* (n= 5), *no me acuerdo yo* (n= 4)

Both constructions tend to introduce a prepositional complement⁵⁴. For *(yo) me acuerdo*, said complements occur 64.7% (22/34) of the time, usually as a clause (13/22: 59.1%); for *(yo) no me acuerdo*, they appear 56.3% (18/32) of the time, mainly as a clause (12/18: 66.7%). These patterns suggest the emergence of a prefab still tied to the argument structure anticipated for the verb: *(yo) me acuerdo (de) + {SOMETHING}* ‘I remember SOMETHING’ and *(yo) no me acuerdo (de) {SOMETHING}* ‘I don’t remember + SOMETHING’. Said adherence to the lexical meaning of the verb would result on a literal function of the construction: to have or keep an idea in one’s mind.

(191) U: *lo poquito que estuve yo nos enseñaba las letras en inglés,*

E: *Mhm.*

U: *y los números,*

... de lo que me acuerdo yo.

Y luego cuando d-

después nos enseñaban un chart asina, con --

Agarraba una jarrita,

y ahí nos iba .. diciendo las palabras.

Me acuerdo yo que leí uno ahí y dije,

ah=, <VOX Mama see kitty VOX>,

y cosas de <Hx esas Hx>. (214:1A1: 962)

‘U: The little time I attend [school], [the teacher] taught us the English alphabet. E: Mhm. U: And the numbers, that’s what I remember. And then, they showed a chart like this, with... He took a little jar, and, one by one, he told us the words. **I remember** I read one there and I said: “Mom, see kitty.” And that sort of things’

(192) U: *.. Ese le dicen la --*

... a ver si me acuerdo.

...(2.0) Cómo le dicen a eso.

Tienen un nombre para [esa cosa].

E: *[Mhm].*

Sí,

es the clapper.

U: *Sí.*

*.. Pero **no me acuerdo** cómo le dice=n. (214-1B2: 293)*

‘U: That one is called the... let see if I remember. How is it called? They have a name for that thing. E: Mhm. It is the clapper. U: Right, but **I don’t remember** how they call it.’

⁵⁴ Although in all of the cases in which said complement is a clause, the preposition is dropped. Still, I consider them prepositional complements because they would not be replaced by an objective pronoun (e.g. *lo*) but a prepositional one (e.g. *de ello*).

Nevertheless, when the use of the construction is observed in context, interpretations other than the literal one emerge, as in (193) below, where the speaker seems to use *me acuerdo de eso* ‘I remember that’ not to indicate a memory, but the importance of it in the story being narrated. Also, by giving the impression that she is answering her own doubts or anticipating any possible questions regarding the truthfulness of what she is telling, the speaker uses the construction as a discourse particle to organize her story. The speaker’s attitude is observed as well in (194), where the memory introduced by *yo me acuerdo* emphasizes the speaker’s conviction regarding the persistence of discrimination against Mexican people.

- (193) *pues,*
me fui con mi hermana,
yo también atrás de ella,
.. y,
.. y nos estuvimos un --
me acuerdo de eso,
sí me acuerdo muy rebien,
nos estuvimos en una casa que está- --
.. Una casa que no vivía gente.
estaba vacía.
(H) fuimos y allá,
nos enbocamos,
áhi nos estuvimos jugando <F toda la plebe F>. (318-1A3: 53)
 ‘Well, I went with my sister, I followed her, and we were in, **I remember that**, yeah, **I remember it** very well, we were in a house that... a house where nobody was living, a vacant house. We went over there and we drank. All the kids, we played there for a long time.’

- (194) *B: [Entonces, no ha] cambiado mucho no?*
W: No.
no.
no.
B: Porque yo me acuerdo en Clayton que --
que mi mamá no podía --
.. ser maestra en la escuela ahí,
.. porque era mexicana. (190-3B2: 584)
 ‘B: So, things haven’t change, have they? W: No. No. No. B: Because I remember in Clayton my mom couldn’t be a teacher, in their school, because she was a Mexican.’

Rather than sporadic, non-literal interpretations can be observed in almost half of the occurrences of both constructions, particularly in those of *(yo) no me acuerdo*: by not being present in the speaker's mind, an event is implicitly introduced as not relevant, false or, even, non-existent. The strategy is particularly useful when speakers want to express disagreement towards what their interlocutor thinks or says because it gives them the ability of implicitly deny an event, without assuming any responsibility regarding its actual occurrence, and without being rude or impolite. For instance, in (195), by saying she does not remember whether kids were forbidden to speak Spanish when she was growing up, instead of straightforwardly denying said idea, the interviewee is rejecting the interviewer's assumption while mitigating her own stance about current discrimination against Spanish speakers, thus protecting her reputation in case someone else states the opposite.

(195) A: *Nunca le pegaron a alguno de .. los niños por hablar español?*

K: *No.*

... No no.

A: *Que XXX –*

K: *Nunca,*

no.

yo no me acuerdo que nunca nos dijieron eh --

.. you know,

cuando --

When I was growing up, que me --

(H) que me dijieran pues,

no hables en mexicano,

no pues.

ahora sí me han dicho. (318-1^a3: 269)

'A: Was any child ever hit for speaking in Spanish? K: No. No, no. A: That [unclear]. K: Never. No. **I don't remember** that they ever told us, eh, you know, when... when I was growing up, being t... well, being told: "don't speak in Mexican", no at all. Now, I have been told that.'

Similarly to what is observed for *(yo) no sé* 'I don't know' and *(yo) creo* 'I think', there are indications that *(yo) no me acuerdo* is experiencing loss of syntactic properties that may facilitate its subjectification. In half of the cases (16/32: 50%), the construction does not take a complement, either because what the speaker does not remember is

obvious from the context, or because there is not a real syntactic relation between the construction and the information it refers to. This is particularly clear when the construction occurs in a parenthetical position.

(196) A: *Nunca le platicaron cómo se conocie=ron o qué?*

P: *s- si le= --*

le --

Si me dijieron <@ no me= --
no me acuerdo @>.

A: *[@@@]*

P: *[no me acuerdo.*

hace ya] .. mucho tiempo.

no me puedo acordar.

‘A: Did they never tell you how did they meet or what? P: If they did... if they told me, I don’t... **I don’t remember.** A: (Laughs) P: **I don’t remember.** It has been a long time. I cannot remember.’

Although just a few, there are also some occurrences of (yo) *me acuerdo* without any complement and as a parenthetical element. Certainly, the literal meaning of the verb remains present in the construction, so it transmits a propositional value; however, the fact that the construction appears in its own intonation unit with final transitional intonation, and in the middle of two segments, also with final transitional intonation, suggests the construction is being used as a discourse particle to facilitate the flow of the discourse while emphasizing the importance of the event being narrated.

(197) M: *Ahí íbanos nosotros.*

en Coyote school en New Haven.

nos ganaba.

[@@]

B: *[@@@] <@ No,*

ella dice que tenían @> unas --

M: *[E=y] ((EXPRESIÓN DE ASENTIMIENTO))*

B: *[una] competición increíble,*
no?

M: *Y luego,*

una --

nos dieron spelling match.

Me acuerdo yo.

.. era beginning.

B: *Ahá.*

M: *que yo podía lo s- --*

u- --

a --

espelearlo. (76-1A1: 412)

‘M: We attended that school. Coyote School, in New Haven, they beat us. (Laughs). B: No, she says they had some... M: Aha. B: An incredible competition, right? M: En then, a... they gave us a spelling match. **I remember**, it was a beginning one. B: Aha. M: I could sp... spell it.’

There is also an instance in which (yo) *me acuerdo* is lexicalized as a temporal marker: *desde que yo me acuerdo* ‘for as long as I remember’. As seen in (198) below, the collocation indicates that an event has been happening since long time ago, without providing a specific date, thus emphasizing the continuity and relevance of said event.

(198) L: *pa' el estado de Utah,*
trabajó él siempre.

E: *Ahá.*

L: *Porque éle fue pastor de borregas.*
Ese era el trabajo de los hombres antes.

E: *Sí.*

L: *Y él --*

.. Ya te digo,

desde que yo me acuerdo,

.. hasta que ya estuvo viejito,

él fue de Utah. (219-1A1: 426)

‘L: in Utah, he worked there always. E: Aha. L: Because he was a shepherd. That was what men used to do in the past. E: Ok. L: And he... I’m telling you, **since I remember**, until he was a senior, he was from Utah.’

Finally, in addition to the occurrences of *acordarse* described above, speakers also recall information with the verb *recordar* ‘to remember’ both in its prescriptive form (n= 11) and in its non-normative form: *recordarse*⁵⁵ (n= 7). The syntactic innovation of the latter, however, does not seem to have any discourse or pragmatic implications yet.

(199) C: *... Y cuáles son algunas de las cosas que **recuerdas** de --*

.. algunas plantas medicinales que usaban,

o que todavía usas.

.. [aquí]. (4-2A2: 101)

‘C: And what are some of the things **you remember** about... some medicinal plants people used, or that you still use, here?’

⁵⁵ Unlike *acordarse* ‘to remember’, *recordar* is not a pronominal verb. By adding *se*, the verb literally means ‘to remember one self’. Nevertheless, speakers in the data sometimes extend the use of *se* from *acordarse* to *recordar*.

- (200) *Curaba de resfríos,
curaba de –
.. 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. (4-2A2: 170)*
‘[My aunt] healed colds, she treated... she healed different illnesses with different herbal medicines. But **I don’t remember** what herbs those might have been because I was... I still was little. But that woman treated a lot of people.’

5.2.2. 3sg Cognitive Constructions

As 1sg do, 3sg subjects, the second most frequent ones for cognitive verbs, (175/814: 21.5%) follow general patterns for the class: they mainly occur in present tense (119/175: 68%) and in utterances with single nucleus (145/175: 82.9%). Given their predominance, it may be expected that the confluence of said features would facilitate the emergence of constructions with different degrees of morpho-syntactic fixation and of subjectivity. Nevertheless, such is not the case due to the distribution of 3sg subject utterances among different verbs, with only one of them, *saber* ‘to know’, showing an important token frequency (81/175: 46.3%). The apparent strength of said verb is diluted when the cross-reference of the aforementioned features reveals that *sabe* ‘X knows’ has only 41 occurrences, which may account for 23.4% (41/175) of all 3sg subjects and 27.9% (41/147) of all 3sg-V utterances in the class, but barely represents 5% (41/814) of all cognitive utterances and 12.7% (41/324) of all *saber* occurrences. Despite low frequencies, however, results for subject realization, type of complement, and prosodic distribution suggest that two different constructions are emerging from 3sg-present utterances with *saber*.

5.2.2.1. *Quién sabe* ‘Who knows’

Given its morphosyntactic fixation, *quién sabe* ‘who knows’ may be the construction that, in the data, best represents what a prefab is. Even though it has a token frequency of just 13 occurrences, the generic reference of its fix subject (i.e. the interrogative pronoun *quién*) results in the construction having rhetorical and epistemic meanings. In (201), for instance, the speaker does not only ask the question *quién sabe qué tantas veces* ‘who knows how many times’ to himself rather than to his interlocutor, but he also utters the question to convey his regret, not to actually look for an answer. In other cases, *quién sabe* is, in a sense, an alternative to *(yo) no sé*, but with stronger epistemic and pragmatic meanings due to the absence of a specific subject-agent. That is to say, if with *(yo) no sé* ‘I don’t know’ speakers mitigate an assertion and, in this way, they distance themselves from its veracity and accuracy, with *quién sabe* the detachment is bigger because speakers do not even appear as the verb’s subject. Such subjectivity is attested in (202), where the speaker uses *quién sabe* ‘who knows’ to answer her interlocutor’s question while implying there is no way for her to know the answer, so she is not to blame for not providing the information.

- (201) J: *cómo quiero deciri.*
 ...No ten- --
 no --
 no teníamos --
 ...ahm,
 ...concepto.
 del --
 de --
 de --
 de la historia.
C: [Hm].
J: [no] no no,
 nunca no preocupaba yo. ((PREOCUPABA))
 ahora me pesa a mí.
 ***quién sabe** qué tantas veces,*
C: [Yeah].

J: [no] haberle preguntado a mi padre muchas cosas así [2you know2]?

C: [2Yeah2].

J: porque él se hubiera acordado mucho,
pues él nació en el setentidos, (144-4B2: 377)

‘J: What am I trying to say? We didn’t... we didn’t have... any concept of... of the history. C: Hm. J: No, no, no. It never worried me. Now, I regret it.

Who knows how much, [I regret] not having asked my father about that kind of things, you know? C: Yeah. J: Because he would have remembered a lot, as he was born in seventy two.’

(202) A: *qué año nació= usted?*

Y: Yo,
en veintiséis.

A: *En veintiséis.*
(H) mil novecientos veintiséis en –

Y: *Mhm.*

A: *en Denver.*

Y: *En Denver.*
[sí].

A: *[Ah=].*
Y cómo2 --
cómo era Denver entonces?

Y: *Pues **quién sabe**?*
no,

A: *No,*
no se acuerda.

Y: *No no.*
pues si yo estaba niña ve.
.. I don't remember.

It's a big city now. (147-1A2: 172)

‘A: In what year were you born? Y: Me? In twenty six. A: In twenty six, nineteen twenty six. Y: Mhm. A: In Denver. Y: In Denver, right. A: Ah. And how... how was Denver in that time? Y: Well, **who knows**, no? A: No, you don’t remember. Y: No, no. Well, I was little, see? I don’t remember. It’s a big city now.’

Quién sabe appears as a parenthetical unit, without any type of complement half of the time (7/13: 53.8%), which suggests it is acquiring the syntactic independence typical of a discourse marker. An example of such prosodic feature is shown in (203).

(203) P: *Qué puede hacer uno?*

A: ***Quién sabe.***
.. A la mejor sí se puede hacer algo.
a la mejor es la hora de hacerlo.

P: *Pos,*
no quieren <X los de la plebe X> oiga.
No más todo lo que [quieren] es cargar pistolas,

A: *[Hm=].*

P: *quieren eh,*
[las drogas] -- (88-1A3: 517)

‘P: There is nothing one can do. A: **Who knows**, maybe something can be done indeed. Maybe it’s time to do it. P: Well, kids do not want to do it, you know. All they want is to have guns. A: Hm. P: They want, eh, drugs.’

When it does have a complement, the prefab mostly introduces clauses (5/6: 38.5%) in postverbal position. Nevertheless, having a traditional argument structure has not inhibited the development of subjective meanings, especially if the subordinate verb is in the form of subjunctive or future of modality, as in (204) and (205), respectively.

(204) V: *Antes no le echaban nada.*

.. *Con el puro estiércol era suficiente.*

...[XX] –

C: *[Entonces] para el mejoramiento de la vida .. en el futuro,*
tiene que haber muchos cambios [crees tú]?

V: *[Munchos] cambios tiene que haber.*
munchos.

.. **Quién sabe** a qué lleguen.

.. *estos chiquitos que están naciendo ahora.*

.. *Porque aho-*

-- *está cambiando mucho todo.*

Todo todo.

...*De antonces a ahora.*

.. *ha cambiado mucho. (4-5B2: 276)*

‘V: In those years, they didn’t put anything; the manure was enough. C: So, to improve life, in the future, a lot of changes need to be done, do you think?’

V: A lot of changes, there need to be, a lot. **Who knows** how far they will get, these kids being born right now, because everything is changing a lot. Everything, everything. Since then to now, [everything] has changed a lot.’

(205) *hacía uno atole,*

.. *mataba marranos,*

.. *había --*

chicharrones,

hacía uno,

y,

.. *Carne de --*

.. *carne de --*

del --

de rancho.

Sí,

de los animales que criaba uno.

.. *Y ahora pues,*

todo viene mer- --

de la tienda.

.. *Todo viene --*

Quién sabe qué trairá.

.. pero de la tienda.

.. Todo ha cambiado mucho.

mucho ha cambiado todo. (4-5B2: 245)

‘People made atole, butchered pigs; there were... people make chicharrones and meat... meat... farm raised meat. Yeah, from the animal we raised ourselves. Now, well, everything comes from market... stores. Everything comes... **who knows** what is in it, from stores. Everything has changed a lot. A lot, everything has changed.’

Rarely present in the interviews, the prefab *sabe* ‘Ø knows’ is related to *quién sabe* in the sense that it always conveys a generic referent, although in this case the syntactic participant is always unexpressed. It may be the case, then, that it is a new construction emerging from *quién sabe*: since the subject always indicates a non-specific referent, its presence seems redundant. Unfortunately, there is not enough data in the corpus to prove such hypothesis.

Regarding its syntactic function, *sabe* always introduces some sort of complement in the same intonation unit. Nevertheless, the apparently traditional argument structure does not result in a propositional interpretation; on the contrary, speakers use the prefab to convey that the source of the information they are presenting is not important; what matters is what they think or do about a particular situation.

(206) E: <@ Pues eso @> --

<@ eso pasó @>.

Mhm. XX

U: Pobrecito.

E: ... Ya mero se acaba esta cinta,
y luego comenzamos con la otra.

U: Pero yo no salí muy mal esa vez [*sabe por qué*].

E: [(H)]

U: porque el viejito era muy bueno XXXX.

E: Aha.

U: Ya él me había ayudado con dinero,
y,

E: El [viejito e-] –

U: [<X Más trabajo X>].

E: era --

es es el marido? (214-1B3: 596)

‘E: That’s why. That’s why. Mhm. U: Poor thing. E: This tape is almost over; then we will start the other one. U: But I didn’t do very badly that time, **who knows why**. E: (H) Because the old man was very good XXX. E: Aha. He had already helped me with some money, and... E: The old man i... U: More work to do. E: was he... is he the husband.’

(207) U: *Yo no los quiero esos de México.*

E: <@ No @>?

U: .. *Y **sabe** por qué no los quiero.*

porque,

(H) unas hijas de la <X plaza X> de mi hermana están casadas ahí,

(H) y se han divorciado y luego,

la de le -- la Sara también.

(H) y luego,

oh ellos las han mantenido, (214-1B3: 388)

‘U: I don’t like people from Mexico. E: No? U: **Who knows** why I don’t like them. Because some of my sister’s daughters are married, and lived over there; and later they got a divorce and later... Sarah’s too. And then they have supported them.’

5.2.2.2. *X sabe* ‘X knows’

The second 3sg subject construction involving *saber* is *él-ella sabe* ‘he-she knows’. It presents a transitive structure consistent with the lexical meaning of the verb, and it shows variability in terms of subject expression, complement realization and polarity. Accordingly, it always transmits a propositional meaning: the knowledge that a specific person has or does not have. Based on such features, the few occurrences of the construction could be described as instances of a general schematic pattern: *X sabe Y* ‘X knows Y’.

(208) K: <VOX<@ yes you are @>.

your mama and me,

and grandpa we are all Spanish VOX>.

<VOX no.

you're not VOX>.

@@ <X quiso hacer X> --

se aferró que no <@ eran Spanish @>.

A: *[Hm=].*

K: *[@@@]*

Hm.

pues no la pudimos hacer entender,

dijo,

Oh, ella s- --

*ella **no sabe** que she's Spanish.*
como no habla más del inglés ve.

A: *Hm.*

K: *pues eh,*

A: *.. Pero,*

.. es lo que --

también eso pasa cuando <X ya [nomás] X> hablan inglés.

K: *[Sí].*

ves e- --

***Ella no sabe** que es mexicana.* (318-1A3: 675)

K: [We said,] yes, you are [Spaniard]; your grandma and me, and grandpa, we are all Spanish. [She said,] no, you're not. She clung to her idea that they were not Spanish. A: Hm. K: Well, we couldn't make her understand. She said... she s... **she doesn't know** that she's Spanish because she only speaks English, see? **She doesn't know** she is Mexican.'

(209) X: *Y el chico no.*

él,

e=h --

habla --

...No muy bien,

pero él habla.

A: *Mhm.*

X: *Más que el que --*

mayor.

el mayor,

sí no.

él,

***sabe** qué estás ha-*

-- diciéndole en mexican,

pero en inglés le responde pa'trás.

.. No lo puede hablar. (47-1A3: 95)

'X: And the youngest doesn't. He... eh, speaks... not that well, but he speaks [Spanish]. A: Mhm. X: Better than the... oldest one. The oldest, well, **he knows** you're speaking to him in Mexican, but he replies in English. He cannot speak it.

5.2.3. Other Prefabs and Collocations

Besides the prefabs and schemas described in 5.2.1 and 5.2.2 above, there are other constructions emerging from less frequent subject-tense combinations, as well as across subject and tense groups; such constructions range from prefabs with some degree of morphological variability (e.g. *saber* + infinitive) to more fixed units (*usted cree?* 'Do you think so?'). Even though none of them has a very high token frequency in relation to the whole class, they are worth mentioning because they represent a strong exemplar in

its particular group (e.g. subject, tense, or verb) or because they provide a good example of subjectivity and subjectification.

5.2.3.1. *SABER* ‘to know how to + infinitive’

The main characteristic of this prefab is the function of *saber* ‘to know’ as an auxiliary verb conveying modality; specifically, the ability to do an activity. The prefab represents just a small percentage of all *saber* items (24/324: 7.4%), but almost half of all modal constructions in the corpus (24/50: 48%). The collocation tends to appear in declarative and affirmative utterances (15/24: 62.5%), present tense (17/24: 70.8%), and with 3sg subjects (12/24: 50%), as shown in (210). Unfortunately, there are not enough data to identify predominant subject-tense combinations; the most frequent one (3sg-present) barely occurs 7 times. Due to the variability of the construction, it can be described as the schematic prefab: X SABER + INF, where the slot for the subject may be filled for any person, the verb can appear in any tense, and any infinitive may fill the corresponding slot.

Given the topic of the interviews, it is not surprising that a speech item is the main verb in half of the occurrences of the prefab (13/24: 54.2%), as shown in (210) and (211) below. Nevertheless, where the prefab’s modality becomes evident is in cases such as (212) and (213), in which the main verb belongs to the material class; a cognitive verb, then, indicates the subject’s ability to perform a non-mental activity.

- (210) A: .. *Y este=,*
 tiene= .. hijos?
 X: *Oh yeah.*
 [tiene a la] –
 A: *[hijos que=] entiendan el.. español.*
 que hablen español.
 X: *El --*
 uno de ellos es --
 e=h,

*entiende poquitito.
no mucho.
el Jakey.*

A: *Jakey.*

X: *.. Él le --
él le= --*

sabe dijir la= .. pledge of legience to the flag in mexicano.

A: *Oh,
el hijo de ella?
[Ahá].*

X: *[Ahá].
Mhm.*

En la escuela le enseñaron and he learned it.

.. and he knows it. (47-4B2: 103)

‘A: And uh, does she have kids? X: Oh yeah, she has... A: Kids that understand Spanish; that speak Spanish. X: The... one of them is... eh, understands a little, not much. Jakey. A: Jakey. X: **He knows how to tell** you the... pledge of allegiance to the flag in Mexican. A: Oh, her son? Aha. X: Aha. Mhm. They taught it to him at school and he learned it. And he knows it.

(211) *.. no podías saber que era gringa ella.
podías decir que era mexicana,
because tenía hasta el accent como nosotros.
.. Ves?*

*Por eso,
.. dice uno asina,
cuando aprende ya uno de más .. % tarde,
pues,
está poco .. dificultoso.
Ella sabía hablar más en mexicano,
.. que en inglés.*

Porque ella le gustaba. (117-1A2: 259)

‘You couldn’t tell she was gringa. You could say she was a Mexican, because she even had the accent like ours. See? That’s why, one speaks like that, when one learns as an adult, well, it is a little bit hard. **She knew how to speak** in Spanish more than in English, because she liked it.’

(212) L: *... Cuando mi papá se murió,
yo abrí todos esos agujeros.*

E: *@@ [XX]*

L: *[Y yo lo-] --
hice los marcos.
yo puse las ventanas.*

E: *[Usted mis-] –*

L: *[Todo.
Todo] es trabajo mío.*

E: *Usted misma.*

L: *.. Sí=,
sé cortar tabla.*

E: *Wo=w.*

@@@

L: *Yo como te digo,*

[no tengo miedo] de nada. (219-1A2: 580)

‘L: When my dad died, I did all those holes. E: (Laughs). L: I make the door frames. I installed the windows. E: By your.. L: Everything. All of this is my own work. E: By yourself. L: Yes. **I know how to saw** wood boards. E: Wow. L: I am telling you. I am not afraid of anything.’

(213) L: .. *Bueno,*

yo le digo ahora mero.

yo no --

No me avergüenzo,

E: *[Mhm].*

L: *[con] decirle a esta mi <X Ingrid, ésta es X>.*

E: *Mhm.*

L: *Ésa sabe manejar rancho,*

E: *[Mhm].*

L: *[Adám] no sabe.* (219-1A2: 536)

‘L: Well, I am telling you right now. I don’t... I am not ashamed. E: Mhm.

L: For example, Ingrid, this is her. E: Mhm. L: **She knows how to manage** a ranch. E: Mhm. L: Adam doesn’t know.’

5.2.3.2. *Se ME hace* ‘it seems to ME’

Like the prefab in the previous section, *se ME hace* ‘it seems to ME’ represents a small percentage of the whole cognitive class (22/814: 2.7%). Nevertheless, when observing its occurrence in relation to that of an impersonal constructions, the collocation accounts for a third of them (22/62: 35.5%). Furthermore, it shows a high degree of morphosyntatic fixation because it always appears in present tense and almost exclusively with the indirect pronoun *me* ‘to me’ (21/22: 95.5%).

In the data, speakers use the prefab mainly to introduce a clause (15/22: 68.2%); however, since, as nominalized units, clauses cannot be pluralized, it can be argued whether they are the verb’s subject or object. In any case, despite the propositional meaning such clause might have, by occurring in an impersonal form, especially with the clitic *se*, the prefab conveys a degree of subjectivity higher than that of other prefabs with a verb in a personal form.

- (214) E: (H) Claro yo le pago bien.
 Pago lo que él pida.
 U: Mhm.
 E: le pago cinco dólares la hora.
 [pero] es lo que él pide.
 U: [pues],
 Sabe usted que a mí **se me hace** que cinco dólares,
por cortar unos zacate es muy bueno=. (214-1B3: 435)
 'E: Of course I pay him well. I pay whatever he asks for. U: Mhm. E: I pay him five dollars per hour, that's how much he asks for. U: Well, you must know that **it seems to me** that five dollars for mowing the grass is a very good deal.'
- (215) P: Y,
 ... para mí era un= --
 una cosa tan feliz pa' conocerlo.
 .. porque era muy brillante el muchacho.
 ... y .. gradó.
 .. Y he is a= nuclear engineer.
 A: .. [Hm].
 P: [ahora] tiene **se me hace** que como .. veintitrés años.
 .. mexicano.
 ... Ése era eh --
 pa' mí era la cosa más sagrada pa' verlo.
 .. tú sabes,
 .. you know,
 .. mexicano,
 tú sabes. (88-1A3: 557)
 'P: And, to me it was a... a pleasure to know him, because the kid was very bright. And he graduated. And he is a nuclear engineer. A: Hm. P: Now, he is, **I think**, like twenty three years old. That was eh... for me, it was the holiest thing to see him. You know, you know. Mexican, you know.'

In the rest of the tokens, *se ME hace* appears as a parenthetical unit. In such cases, the epistemic meaning of the prefab is even more evident because speakers use it to reduce the assertive strength of what they have just said, either because they are not sure about the information or because they do not want to be held accountable for it.

- (216) A: y qué tal habla=n --
 qué % tal le parece usted que habla español.
 X: .. Qué tal?
 A: Mhm.
 .. Qué tan bien lo habla.
 X: .. No muy bie=n.
Se me hace a mí.
 yo no sé. (47-1A3: 22)

... How do they speak... how do you think you speak Spanish? X: How do I speak? A: Mhm. How well do you speak it? X: Not very well, **it seems to me**. I don't know.'

(217) Z: [2Ya2] *ha ido dos veces*.

O: *Allá 'onde de --
de Taos,
llevan a los [senior citizens].*

B: *[Oh sí. a los seniors].
yeah.*

O: *Y llevan a veinte=cinco,
se me hace,
el lunes,
este [lunes que] viene.*

B: *[de este] -- van a ir otra vez?*

O: *Van a [ir otra vez]. (311-4A1: 242)*

'Z: She already has gone twice. O: There, where to.. to Taos, they take the senior citizens over there. B: Oh, yeah, the senior citizens, yeah. O: They take twenty five of them, **I think**, on Monday. This next Monday. B: This... are they going there again? O: They are going again.'

5.2.3.3. *(tú) sabes & (usted) sabe* 'You know'

Even though 2sg subjects do not appear among the most frequent ones in the whole data, they are involved in two constructions emerging from the cognitive class: *(tú) sabes* 'you_(fam) know' and *(usted) sabe* 'you_(formal) know'. *(Tú) sabes* accounts for the majority of *saber* utterances with *tú* subjects (16/18: 88.9%) and almost a quarter of all cognitive utterances with said subject (16/64: 25%). The collocation shows some features that suggest loss of syntactic properties and development of discourse marker properties: it tends to take no complement (9/16: 56.25%), cases in which it always occurs with a pronominal subject in preverbal and parenthetical positions. Evidence for discourse marker status comes too from the tendency of the subject's referent to be a generic one rather than the actual hearer, and from the interactional and epistemic meaning the item conveys.

(218) P: *Y he is a= nuclear engineer.*

A: *.. [Hm].*

P: *[ahora] tiene se me hace que como .. veintitrés años.
.. mexicano.*

...Ése era eh --
 pa' mí era la cosa más sagrada pa' verlo. .
 . tú sabes,
 .. you know,
 .. mexicano,
 tú sabes. (88-1A3: 561)

'P: And he is a nuclear engineer. A: Hm. P: Now, he is, I think, like twenty three years old. That was eh... for me, it was the holiest thing to see him. You know, you know. Mexican, **you know**'.

(219) B: *Entonces usted y mi papá tienen el mismo cumpleaños.*

M: *Sí?*

B: *El quince [de marzo].*

M: *[Oh, sí]. Yo le mandaba .. birthday cards, y él también a mí.*

B: *Ah= [okay].*

M: *[tú sabes].*

B: *Mhm.* (76-1A1: 41)

'B: So you and my dad have the same birthday. M: Is that right? B: March 15th. M: Oh, yeah. I used to send him birthday cards and him to me. B: Ah, okay. M: **You know**. B: Mhm.'

When the verb takes a complement (7/16: 43.8%), it usually takes a direct object (6/7: 85.7%); it always appears with unexpressed subject and predominantly in interrogative utterances. Nevertheless, such a canonical transitive structure does not seem to result in a propositional use of the collocation; in all but one of the utterances, the collocation performs a rhetorical or an interactional function: the speaker "confirms" the knowledge of the hearer or checks the channel of communication before presenting information he or she considers to be important. In (220) and (221), for instance, it is clear that the speaker does not have any interest in knowing the interlocutor's answer, but wants to draw attention to what she is about to say. Furthermore, in cases such as (222), the speaker uses the same collocation and the same strategy to present important information even if the referent of *sabes* 'you know' is not really the hearer, but a participant in the story being narrated.

- (220) M: *estuve hasta que me casé en el treinta y nueve.*
Esto era en el treinta y --
en el treinta y ocho.
Trabajé hasta el treinta y [nueve] <X [en Trevero] X>,
 B: [Mhm].
 M: *y luego me casé.*
 B: *Hm=?*
Okay.
 M: *Y sabes que sé mejor yo=,*
.. el libro ocho que mi <@ familia @> que gradó del doce.
.. pa' hacere –
...hm=,
[arithmetic],
 B: [Oh sí]?
[2Ahá2].
 M: [2spelling2],
y todo eso?
.. Yo les gano. (76-1A1: 285)
 ‘M: I was there until I got married in thirty nine. This was in thirty... in thirty eight. I worked until thirty nine in Trevero [unclear]. B: Mhm. M: Then I got married. B: Hm? Okay. M: And **do you know** I know eighth grade better than my family who did until twelfth. Doing... hm... Arithmetic. B: Oh yeah. Aha. M: Spelling, and all of that? I am better than them.’
- (221) *Los chotas no hacen nada.*
Mira,
ahora mero tenemos un caso aquí en Española que pasó este invierno.
Una señora García que trabajaba en un ah .. comercio,
o algo asina.
.. Llegaron,
la sacaron,
la mataron,
la quemaron,
...y todo.
Y sabes cuánto le están dando al matador?
Siete años. (219-1B1: 599)
 ‘Cops don’t do anything. Look, right now, here in Española, we have a case that happened this winter. A Garcia woman who worked in a... store, or something like that; someone came into the store, took her away, killed her, and burned her and everything. And **do you know** how many years was the killer sentenced to? Seven years.’
- (222) U: *lo último que me dijo que --*
que me casara con ella,
.. Y entonces a ver si le enseñaba,
pero le dije yo,
(H) Yo me casara contigo,
le dije,
pero sabes lo que pasa?
(H) que tú perteneces a una religión y yo a otra.
Y tú no quieres venirte con la mía ni yo con la tuya.

E: .. Aha.

U: *De modos que,*

áhi -- ((AHÍ))

<X pues X> *áhi se quebró.* (214-1A1: 732)

‘U: Last thing she told me was that... that I should marry her and then I might teach her [to cook]; but I told her, “I could marry you”, I told her, “but **do you know** what the problem is? You have a religion and I have a different one. And you don’t want to convert to mine, nor do I want to convert to yours. E: Aha. So, well, that was the end of it.’

As for (*usted*) *sabe* (‘you know’), it accounts for the majority of *saber* utterances with *usted* subjects (12/13: 92.3%) and a little more than a fifth of all cognitive utterances with said subject (12/54: 22.2%). The collocation tends to appear with unexpressed subject (8/12: 66.7%), and it shows more variability than *tú sabes* and *¿sabes + complement?* because neither expressed subjects nor unexpressed ones are strongly associated with presence or absence of complements. Moreover, expressed subjects (4/12: 33.3%) do not show any preference for a particular position; they appear before the verb (2/12: 16.7%) as well as after it (2/12: 16.7%). Regarding type of complement, (*usted*) *sabe* takes direct objects most of the time (10/12: 83.3%), usually a subordinate clause (7/10: 70%).

The collocation is mainly used in affirmative utterances, either declarative (5/12: 41.7%) or interrogative (6/12: 50%) ones, as in (223) and (224) respectively. Even though not as predominant as in *tú* prefabs, rhetorical and interactional functions are also attested with (*usted*) *sabe* and *¿(usted) sabe?*. In (223), for instance, the speaker uses *¿sabe qué?* ‘you know what?’ as a device to emphasize the fact that he used to visit a sick old man. Because the speaker had gotten monetary help from the man in the past, it is important for him to mention those visits because they prove his gratitude. On the contrary, in (224) *ya sabe* ‘you already know’ seems to reduce the importance of the phrase *ya habían tratores* ‘there already were tractors’ because it presents the information

as something well known. With this strategy, the speaker is able to elaborate on his explanation without giving the impression that he is pretending to be smarter than the interlocutor.

(223) U: *porque el viejito era muy bueno XXXX.*

E: *Aha.*

U: *Ya él me había ayudado con dinero,*

y,

E: *El [viejito e-] –*

U: *[<X Más trabajo X>].*

E: *era --*

es es el marido?

U: *El marido.*

Pero se murió,

*.. y **sabe** qué .. XX iba yo a verlo cuando estaba enfermo?*

XX .. en el corazón. (214-1B3: 608)

‘U: Because the old man was very good XXX (unclear). E: Aha. U: He had already helped me with some money and... E: The old man wa... U: more work. E: Was he... is he the husband? U: The husband. But he died and **you know** what? XX I visit him when he was sick? A heart problem.’

(224) L: *Ya Ursula .. pagaba quien me hiciera el negocio.*

como cortar las alfalfas,

como .. barbechar.

E: *Ahá.*

L: *Porque ya no había caballos.*

*ya **sabe** ya habían -- .. tratores,*

como dicemos. (219-1A1: 470)

‘L: By that time, Ursula.. payed for someone to manage my business. How to harvest the alfalfa, how to plow. E: Aha. L: Because there was no longer horses, **you know**, there were tractors, as we say.

5.2.3.4. *Entiendo* ‘I understand’, ¿entiendes? ‘Do you understand? & (X) entiende ‘X understands’

Despite its low token frequency in relation to both the whole data (81/2134: 3.8%) and the cognitive class (81/814: 10%), *entender* ‘to understand’ produces two subject-tense combinations that may be developing prefab, and even discourse marker, status.

The first of such combinations is (yo) *entiendo* ‘I understand’, which accounts for a quarter of all instances of the verb (21/81: 25.9%), a quarter as well of all instances of *entender* in present tense (21/76: 27.6%) and, more importantly, practically all of 1sg-

present utterances with this verb (21/23: 91.3%). On the matter of subject, it tends to be unexpressed (16/21: 76.2%), followed by pronominal subjects in preverbal position (4/21: 19%).

The collocation usually takes objective complements (13/21: 61.9%) expressed by a subordinate clause (6/13: 46.2). Interestingly, it is in these utterances where the collocation seems to be developing epistemic meaning. Rather than expressing the cognitive achievement of an idea or state of mind (e.g. to understand a concept or a piece of art), (yo) *entiendo* conveys a meaning close to *creer* ‘to believe, *pensar* ‘to think’ or *juzgar* ‘to judge’, as in (225), where the collocation could be interpreted as the speaker’s lack of certainty regarding the information or, even, as a sort of evidential with which the speaker implies he did not acquire the information personally, but it is something of general knowledge. On the contrary, when the complement is not a clause, *entender* clearly conveys its lexical meaning, as observable in (227).

- (225) *en mi --*
% poquito .. concepto que tengo de los pueblos,
.. pa' acá pa' las montañas,
.. Trampas.
Yo creo es el más antiguo.
.. Y después,
... Truchas.
(H) porque en Trampas eh --
*yo **entiendo** que --*
que doce familias vinieron de Santa Fe y poblaron Trampas en el setes- --
mil setecientos .. cincuenta,
(H) mientras que aquí=,
doce familias se poblaron aquí en Truchas en el se- --
setecientos cincuenta y cuatro. (144-4B2: 214)

‘in my... little knowdge of the towns, mountain towns, Trampas, I think it is the oldest, and, after it, Truchas. Because in Trampas eh... **it is my understanding**, twelve families came from Santa Fe to settle Trampas in seventeen... seventeen fifty. Here, on the other hand, twelve families settled Truchas in se... seventeen fifty four.’

- (226) T: *Bueno,*
no más que iban a los cíbolos,
[y traiban] eh –
 F: *[O=h].*
 T: *mucha carne seca y los cueros también.*
 F: *O=h.*
 T: *los cueros.*
 F: *A dónde iban?*
 T: **Entiendo** *que iban pa' allá pa' aquel lado.*
pa' el --
pa' el rumbo del río de Pecos. (270-1B2: 511)
 ‘T: Well, when they went for buffalos, they used to bring eh... F: Oh. T: a lot of jerky and leathers too. F: Oh. T: Leathers. F: Where did they go? T: **I understand** they used to go over there, by the Pecos River’s area.’
- (227) B: *Usted prefiere .. ver .. la televisión en inglés o en español.*
 M: *Pues más inglés.*
 B: *[Mhm].*
 M: **[entiendo]** *más el inglés.*
 B: *Ahá.*
 M: *Del español hay palabras que no entiendo [muy bien]. (76-1A1: 641)*
 ‘B: Do you prefer to watch TV in English or in Spanish? M: Well, mostly English. B: Mhm. M: **I understand** English better. B: Aha. M: In Spanish, there are words I don’t understand very well.’

The collocation also appears without formal direct complement, although it is possible to identify what information the verb is referring to in the discourse. In such cases, there is an extension regarding the cognitive achievement or perception indicated by the lexical meaning of the verb; instead of being purely cognitive, it includes the sensorial aspect of the perception of an utterance. Thus, in (228), *yo le entiendo* (‘I understand what it says’), conveys both sensorial decoding (the understanding of the words) and cognitive understanding (the idea). Regarding this type of utterances, it is worth mentioning the addition of an indirect pronoun. When present, the verb tends to indicate that the speaker understands (sensorial and cognitively) what someone else says, as in (228) below. When the indirect pronoun is absent, *entiendo* ‘I understand’ indicates not only the speaker’s cognitive grasp of the information, but also his/her attitude towards

it. In (229), for instance, *entiendo* conveys the speaker's surprise regarding the events in the story he is narrating.

- (228) M: *Pos,*
sí sé escribir mejor inglés que hablarlo.
B: *Yeah.*
M: *Pues sí puedo hablarlo,*
pero=,
no puedo pronunciar algunas palabras que quiero.
B: *[yeah].*
M: *[y áhi] me quedo.*
áhi comienza el español.
B: *Yeah.*
[No=,
yo le entiendo].
M: *[Luego le sigo] en inglés. (76-1A1: 323)*
'M: Well, I know how to write English better than how to speak it. B: Yeah.
M: Because I can speak it, but I cannot pronounce some words I want to. B:
Yeah. So I stop there and start to speak Spanish. B: Yeah. **I know what you**
mean. M: Then I continue in English.'
- (229) H: *.. Ahora que estaba yo en el --*
que estaba yo en .. Alamogordo,
.. Agarraron unos cabrones,
.. con seventy --
.. Seventy tons,
F: *O=h.*
H: *de cocaine.*
Cómo hijo de la chingada le dije yo.
podían haber pasado seventy tons.
yo entiendo una pound,
two pounds.
F: *Yeah.*
H: *abajo del sobaco.*
F: *@@@*
H: *<@ pero @>,*
.. seventy tons? (156-1A1: 540)
'H: Just recently, while I was in ... I was in Alamogordo, they caught some
scumbags with seventy... seventy tons, F: Oh. H: of cocaine. How the fuck, I
said to him, they could have passed seventy tons? **I can understand** one
pound, two pounds. F: Yeah. H: Under the armpit. F: (Lauhgs) H: But
seventy pounds?'

The second collocation involving *entender* is *¿entiendes?* 'Do you understand?', an interrogative unit which accounts for a considerable percentage of all *entender* occurrences in the data (15/81: 18.5%) and for the vast majority of 2sg-*entender*

utterances in the data (15/17: 88.2%). It predominantly occurs with unexpressed subject (14/15: 93.3%), it tends to appear without any complement (13/15: 86.7%), and in parenthetical position (13/15: 86.7%), usually at the end of its utterance. Regarding its function in the discourse, the marker conveys a rhetorical question; rather than really asking for the interlocutor's answer, it points to information the speaker considers as obvious or easily understandable⁵⁶. By doing so, speakers implicitly push the hearer to agree with their opinion while checking if he or she is following the narration.

(230) F: *Mucha gente se ha mudado pa' Las Vegas con --
con los trabajos y toda XXX.*

H: *Las Vegas,
.. es <X bueno si no tienes vicios X>.
.. **Me entiendes?**
si tienes vicios la --
Forget about it.*

Más vale que te salgas de ahí. (156-1A1: 259)

'F: A lot of people have moved to Las Vegas because of the jobs and all of XX (unclear). H: Las Vegas, it is good if you don't have any vices. **Do you know what I mean?** If you have vices the... Forget about it. You better get out of there.'

(231) H: *Cómo hijo de la chingada le dije yo.
podían haber pasado seventy tons.
yo entiendo una pound,
two pounds.*

F: *Yeah.*

H: *abajo del sobaco.*

F: *@@@*

H: *<@ pero @>,
.. seventy tons?*

*.. Lo tenían --
lo tenían --
lo tenían escondido en --
.. en cinco plazas.
me entiendes?*

F: *Oh,
[yeah].* (156-1A1: 551)

'H: How the fuck, I said to him, could they have passed seventy tons [of cocaine]? I can understand one pound, two pounds. F: Yeah. H: Under the armpit. F: (Lauhgs) H: But seventy pounds?' They had it... they had it...

⁵⁶ In fact, in the remaining occurrences, *entiendes* is uttered as a rhetorical assertion; although it has not an interrogative intonation curve, it is not really asserting the information, but directing the hearer to agree with the speaker.

they had it hidden in... in five places. **Do you know what I mean?** F: Oh, yeah.

5.3. Summary

Two constructions emerge as the stronger exemplars in the cognitive class because of its token frequency and, more importantly, their prominence across the different features involved in their formal structure: *(yo) no sé* ‘I don’t know’ and *(yo) creo* ‘I think’. Even though both constructions follow general patterns for the class and the whole data in terms of subject, tense and type of predicate, the fact that each one of them occurs with a different verb, and with a specific polarity demonstrates that said general patterns do not apply homogeneously to each member in the class, thus probing the importance of examining the local context in which each verbal item occur in order to obtain a detailed and accurate account of how language usage and structure influence each other. In terms of subjectivity and subjectification, both constructions provide evidence for the assumption that development of epistemicity is related to deviance from the normative argument structure associated with the lexical meaning of the verb: constructions whose verb is losing, or has lost, syntactic properties are more likely to convey subjective nuances and, eventually, become discourse markers. Another feature that proved to play an essential role in the development of epistemicity was prosody: constructions that speakers use to express attitudes and opinions often appear in its own intonation unit and as parenthetical particles.

Results from the local analysis of cognitive utterances show as well that token frequency should not be the most important factor to consider when looking for prefabs and collocations, particularly those showing a high degree of subjectification. Instead, it is important to observe the occurrence of specific constructions in relation to the features

involved in their structural configuration. In the data, for instance, *quién sabe* ‘who knows’, *sabe* ‘Ø-knows’, *no me acuerdo* ‘I don’t remember’, *se me hace* ‘it seems to me’ and *tú sabes* ‘you know’ represent important percentages within their subject, tense or verb group, even if their token frequency is very low in comparison to the whole data. More importantly, said constructions show signs of lack of morpho-syntactic variability, and some of them seem to have entered into a subjectification process; they sometimes convey epistemic meanings and occur as parenthetical units.

Chapter 6. Speech Verbs

The present chapter presents and discusses results from the analysis of speech verbs in the data. As in chapter 4, the first part of the analysis focuses on general distributions of all of the tokens in the semantic class based on features such as subject, tense, complement, and type of utterance. These results serve later as the basis for a local analysis; one which focuses subject-tense-predicate combinations with specific verbs seeking to determine their degree of morpho-syntactic invariability, as well as their degree of subjectivity. The local analysis will discuss as well constructions arising from less frequent combinations seeking to demonstrate the importance of going beyond token frequency when examining the role of language use and interaction in the development of subjectivity and subjectification. The chapter concludes with a summary of the results.

6.1. Speech Verbs: General Patterns

The speech verbal class includes 11 verbs, whose occurrences sums up 1320 tokens. The class presents a very uneven distribution because together, just two lexical items, *decir* ‘to say’ (n= 853) and *hablar* ‘to speak’ (n= 375), account for the majority of the utterances (1228/1320: 93%). Furthermore, *decir* by itself represents more than half of the group (64.6%). The predominance of these two lexical items is even more evident considering than the next most frequent verb is *preguntar* ‘to ask’, with only 41 occurrences (3.1%). As it will be shown latter in this chapter, the emergence of *decir* as the most frequent verb both in its verbal class and in the whole data is related to its generic semantic meaning, which allows speaker to use it to perform a variety of speech acts: questions, demands, assertions, etc., thus easing the development of lexicalized and grammaticalized constructions. On the contrary, the high frequency of *hablar* seems to be

closely dependent on, and specific to, the topics narrated by the interviewees; that is to say, constructions observed in the data may not be as frequently used in other discourses. As in Chapter 5, I will first described and discussed the distribution of the utterances in terms of type of predicate, subject and tense. Next, I will focus on verb-specific constructions involving such general patterns to determine whether they share properties and functions as members of the same class or, on the contrary, they each behave differently despite apparently emerging from the same patterns. To give a comprehensive account of how speakers use speech utterances, I will discuss as well constructions that may be relevant in the specific groups they belong to (e.g. subject, tense and verb), and that show signs of subjectification, even if they are not particularly frequent in relation to the speech class as a whole.

6.1.1. Verbs of Speech: Predicate Type

Table 14 presents the most frequent types of predicate involving verbs of speech in the data. As shown, in most cases, speech verbs appear as the nucleus of simple predicates⁵⁷ (1124/1320: 85.2%), as in (232), result that closely follows patterns for the whole data (87%), and for the cognitive class (91.8%).

Table 14. Verbs of Speech by type of Predicate

Type of Predicate	Token Frequency and Percentage
V	1124 (85.2%)
Aux + INF _(speech)	91 (6.9%)
Aux + GER _(speech)	37 (2.8%)
V + INF _(speech)	31 (2.3%)
Other	37 (2.8%)
Total	1320 (100%)

⁵⁷ As described in 3.3, simple predicates include both simple and compound tenses.

(232) Y: *Tengo que comer algo.*
Por eso tengo que almorzar,
y comer,
y cen- --
XX,
pero no ceno.
porque no cenamos.
no.
tienes que.
dijo,
para que la tomes.
 @

A: ... *un vasito de leche o algo [así].*

Y: *[Asina] me dijo,*
just take a snack.
something to eat, (147-2A2: 316)

Y: I have to eat something (before taking medicine). That's why I have to eat breakfast, and eat lunch, and eat din- -- XX, but I don't eat dinner; because we don't eat dinner, no. You have to, he said, so you can take it (the medicine). A: A little bit of milk or something like that. Y: That's what he told me, "just take a snack or something to eat..."

The table highlights too an important difference between both semantic classes as well, namely, a larger percentage of speech items occur as main verbs in periphrasis with infinitive (6.9%), as in (233). Also, unlike cognitive verbs, there are not speech verbs in the slot of the auxiliary element. Within said group of periphrases, 28.9% (26/91) of the instances use the modal verb *poder* 'to can', (234), as the auxiliary. Although the percentage of occurrence for speech verbs is similar to the one in the cognitive class (7/25: 28%)⁵⁸, it is relevant that the majority of modal periphrases appear with speech verbs (26/33: 78.8%) rather than with a cognitive one.

(233) A: *Por qué no está bien?*
 R: *Porque, tenemos que= --*
.. que= ..
tener nuestra idomia. ((SIC))
No --
*No **debemos de olvidarla.***
Mhnh. ((NEGACIÓN))
No se debe de --

⁵⁸ The percentage is based on the periphrases in which the speech verb is the main one; it does not include those tokens (n= 24) in which *saber* 'to know' functions as an auxiliary.

de -- de olvidar.
A: .. *Pero para qué lo --*
como los chamacos dicen,
pa' qué sirve. (20-5A3: 89)
‘A: Why is it not good? R: Because we have to.. to ... have our language.
We mustn’t forget it. Mhmh. It must not be forgotten. But, what it... as the
kids say, what is it good for?’

(234) M: *Pos,*
sí sé escribir mejor inglés que hablarlo.
B: *Yeah.*
M: *Pues sí **puedo hablarlo,***
pero=,
no puedo pronunciar algunas palabras que quiero. (76-1A1: 315)
‘Well, I know how to write English better than speak it. B: Yeah. M: Well, I
can speak it, but I can’t pronounce some words that I want to say’

Periphrases with infinitive also include a group of utterances with *saber* ‘to know’ as the auxiliary verb (235). These tokens constitute less than one percent of the whole semantic class (11/1320: 0.8%), but more than a tenth (11/91: 12.1%) of the pattern Aux+INF_(speech). Furthermore, they represent a third of all utterances with *saber*+INF in the corpus (11/35: 31.4%). It seems, then, that, in this particular interviews, the auxiliary *saber* ‘to know’ shows a tendency to co-occur with verbs of speech to indicate a person’s linguistic ability.

(235) P: *Ella --*
su español no está muy muy muy buena.
.. Pero, .. su inglés sí.
*Sus chamacos no **saben hablar** español.*
A: .. [Hm]. P: [y el mío] *sí sabe.* (88-1A2: 486)
P: She... her Spanish is not so good. But her English it is good. Her kids
can’t speak Spanish. A: Hm. P: But mine, he can.’

There are two other types of predicates occurring more frequently in the speech class than in the cognitive one: periphrases with gerund, as in (236), and ‘V+INF’⁵⁹, as in (237) and (238). Even though the ‘Aux+GER_(speech)’ construction represents just 2.8%

⁵⁹ For this pattern, a verb of speech or cognition functions as the nominal complement of the main verb, which may belong to any semantic class.

(37/1320) of its class, its tokens account for most instances of this periphrasis in the data (37/41: 90.2%).

(236) K: *Ahora mismo notas que cada rato te digo something en inglés porque, ya le viene a uno de= -- yo no sé. a -- ... hay veces **que estoy platicando** con= somebody que no sabe en -- en % inglés.*

cuando menos acuerdo les he dicho en inglés, y luego tengo que dijir en mexicano.

@ @ @

A: @ @ @ @ @

K: *Pero,*

pero= mi idomia es mexicano, (318-1A3: 545)

‘K: Right now you can see that I often tell you something in English; that comes from... I don’t know. Sometimes **I am chatting** with somebody who doesn’t know En.. English; suddenly, I realized I just have just told them something in English. Then I have to say it again in Mexican (LAUGHS). A: (LAUGHS). K: But my language is Mexican.’

(237) A: *Perdió español de cuando era niña [o lo mejoró]?*

Y: *[No. no. pero=] --*

No lo practicábanos⁶⁰,

porque en la escuela,

*no nos **dejaban hablar** español para que aprendiéranos el inglés.* (147-1A2: 299)

‘A: Did you forget your Spanish when you were a child o did it get better? Y: No, no, but... We didn’t practice it because **we were not allowed to speak** Spanish in school so we would learn English.’

(238) Y: *Diferentes modos tenían los maestros pa' tra- -- hacernos a nosotros hablar inglés.*

No,

*porque nosotros **preferíanos hablar** español.* (147-1A2: 301)

‘Y: Teachers had different ways to tr- to make us speak English, no, because we **preferred to speak** Spanish.’

As for the occurrence of speech verbs in the pattern V + INF, their 31 tokens barely represent 1% (31/3120) of the semantic class, but the majority of the occurrences of this type of predicate in the data (31/40: 77.5%). Furthermore, despite its low token frequency, this grammatical pattern has produced the prefab *QUERER decir*, which not

⁶⁰ In New Mexican Spanish, instead of prescriptive *-amos* (e.g. *íbamos* ‘we used to go’), *-anos* is a common ending for first person plural of imperfect tense, both indicative (e.g. *ibanos*) and subjunctive (e.g. *fuéranos* ‘that we might go’).

always can be interpreted as the sum of the individual meaning of both verbs (i.e. to have a desire to say something), but as a whole unit whose closest translation to English would be ‘to mean’, as shown in (239) below.

- (239) ... *Mi papá en una ocasión,*
ah=,
se peleó con unos --
oh c- g- --
gringos que estaban áhi.
No sé por qué.
.. Y,
[...]
Y le dijeron una palabra que no entendía y fue a --
a --
a ver a Rubén .. Murillo,
*, que le explicara qué era lo que **quería decir** la palabra esa. (144-4A2: 167)*
 ‘Once, my dad had a fight with some gringos who were there. I don’t know why. And... and they said to him a word that he didn’t understand. So he went and asked Ruben Murillo what that **word meant**.’

To conclude this section, it is worth mentioning that only speech verbs appear as complement of prepositional verbs. Although uncommon in terms of frequency, this grammatical pattern is a good indicator that speech verbs as a class have more syntactic variability than cognitive verbs do.

- (240) *Ellos aprendieron a hablar el mexicano,*
*y nosotros **aprendemos a hablar** en inglés con ellos. (117-1A2: 137)*
 ‘They learned to speak Mexican and **we learn to speak** English with them’.

6.1.2. Subject and Tense

Table 15 shows the distribution of speech verbs according to the tense and subject of the utterance’s finite verb. Following global patterns, the most frequent tenses are present, preterit and imperfect. Nevertheless, important differences emerge when comparing results from the speech class to those from the cognitive class; the first one of which is the weight of each tense in each semantic class and in the corpus. Thus, although present tense is also the most frequent for utterances with speech verbs (569/1320:

43.1%), it is much less frequent than in the cognitive class (669/814: 82.2%); on the contrary, preterit utterances tend to occur with verbs of speech (394/423: 93.1%), and are relatively much more frequent (394/1320: 29.8%) than in the cognitive class (29/814: 3.6%). The imperfect tense group behaves similarly to the preterit one; it largely occurs with speech verbs (258/335: 77%) and, consequently, is more frequent in said semantic class (258/1320: 19.5%) than in the cognitive one (77/814: 9.5%).

Table 15. Verbs of Speech, distribution for Subject and Tense

	Present	Preterit	Imperfect	Other	Total
1sg	206	116	44	27	393 (29.8%)
2sg-tú	18	6	3	8	35 (2.6%)
2sg-Ud	42	15	12	11	80 (6.1%)
3sg	114	228	64	22	428 (32.4%)
1pl	26	6	27	10	69 (5.2%)
2pl	1	0	12	3	16 (1.2%)
3pl	150	23	96	16	285 (21.6%)
Other ⁶¹	12	0	0	2	13 (1%)
Total	569	394	258	99	1320 (100%)

Regarding subject distribution, speech utterances as a class follow global patterns too; the most frequent subjects being 1sg (393/1320: 29.8%), 3sg (428/1320: 32.4%) and 3pl (285/1320: 21.6%). However, important differences emerge once again when comparing the occurrence of said subjects in the speech class to that of the cognitive one. Thus, while 1sg subjects account for a larger percentage of cognitive (423/814: 51.9%) than of speech utterances (393/1320: 29.8%), the opposite happens for 3sg subjects. They have a larger participation in the speech class (428/1320: 32.4%) than in the cognitive one (168/814: 20.6%). Such a behavior is similar among 3pl subjects; they appear more frequently in the speech class (285/1320: 21.6%) than in the cognitive one (67/808:

⁶¹ It includes constructions without a verb in a personal form (e.g. *cómo decir*), as well as impersonal constructions in which the predicate's nucleus appears in a defective form. Although the form of said verbs coincide with that of 3s subjects, including them in the same category would affect results for subject expression and position.

8.3%). Other differences worth mentioning come from the use of both forms of 2sg subjects. While the percentage of occurrence for 2sg-*usted* subjects is very similar among utterances with speech (80/1320: 6.1%) and cognitive verbs (55/808: 6.8%), 2sg-*tú* subjects are less frequent in the former verbal class (35/1320: 2.7%) than in the latter one (70/808: 8.7%).

As a whole, 1sg subjects in the speech class show a high degree of structural uniformity for some of the features because one of the coded values accounts for the vast majority of all the utterances. Thus, 1sg subjects typically occur in declarative (381/393: 96.9%), affirmative (364/393: 92.6%) utterances, with the speech verb as the predicate's nucleus (318/393: 80.9%). Furthermore, even if there is not just one predominant category, the degree of variability may not be really very high because utterances tend to get grouped around just a couple of values; for instance, a large number of 1sg subject utterances, (206/393: 52.4%), occur in present tense. From the distribution of said features, it may seem that utterances with 1sg are very homogenous. Nevertheless, it is important to consider that such results may be skewed due to the high token frequency of *decir* (n= 853) and, in smaller proportion, *hablar* (n= 375). Examples of these typical exemplars are shown in (241) and (242) below.

- (241) O: [Y le digo,
no ha] <X ido X> al casino nuevo,
que abrieron la otra noche,
le **digo** yo [que] yo quiero ir a conocerlo.
B: [Hm].
(TSK) Ah=.
Mhm.
O: Ha llegado a ir usted pa' Las Vegas?
B: Sí.
unas .. dos veces.
O: Ah=.
B: yeah.
conozco.
pero n- --

A mí no me gusta --

.. gastar el dinero. (311-4A1: 272)

‘O: And I tell him he has not gone to the new casino, the one they just opened some nights ago. **I tell him** that I want to know it. B: Hm. Ah. Mhm. O: Have you ever gone to Las Vegas. B: Yes. Once or twice. O: Ah. B: Yeah, I know that city. But I don’t like to waste my money betting.’

(242) *P: Y muchos de estos lugares aquí en Albuquerque te dicen,*

.. No queremos que hablen español.

.. Pero nosotros,

.. Ha trabajado= otros lugares y,

preguntamos no,

eso n=- --

naide nos puede quitar eso.

*Ése es mi deber mío si **hablo** español,*

inglés,

.. japonés. (88-1A3: 220)

‘P: And in many of these places, here in Albuquerque, they tell you: “we don’t want you to speak Spanish”. But we... I have worked in other places and we ask whether we can or cannot, no? Nobody can take that right away from us. If **I speak** Spanish, English or Japanese, it is my duty to ask whether I may speak it [in my job].’

Furthermore, when examining how the subject interacts with other features, for instance, subject expression and position, important differences emerge. Considering that subject expression in Spanish is required just in very specific situations, it is natural to expect that unexpressed subjects will be the more frequent ones. The prediction is largely confirmed for 1sg, which is unexpressed, as shown in (243), most of the time (278/393: 70.7%). However, considering than one of the factors that trigger subject expression in Spanish is the possibility of referential ambiguity, and that such possibility is much lower for 1sg subjects than for 3sg, for instance, what is relevant is not really the predominance of unexpressed 1sg subjects, but the fact that a third of 1sg subjects are pronominal (115/393: 29.3%), and that they tend to appear after the verb (61/115: 53%) rather than before (52/115: 45.2%)⁶², as in (244) and (245) respectively.

⁶² There are a couple of instances with the pronominal subject expressed twice, one pronoun before the verb and another one after it.

- (243) *quítate el cute y,*
 .. No,
 Le digo,
 ya me voy pa' mi casa. (318-1A3: 160)
 ‘[He said] Take off your coat. No, **I tell him**, I’m going home.’
- (244) *Ellos sabían que me –*
 .. tenían que respetarme.
 .. y asina he sido siempre,
 *le **digo yo**.* (311-3B2: 311)
 ‘They know they.. they had to respect me, and I’ve been like that, I am
 telling you’.
- (245) ***Yo les digo,***
 Yo no nací pa' sentada. (219-1A2: 224)
 I tell them, I wasn't born [to be] sitting’.

Differences on patterns also emerge when focusing on how subjects in the whole data are distributed between both verbal classes. Since speech verbs are more frequent (1320/2128: 62%) than cognitive ones (808/2128: 38%), one would expect that for each subject, a larger percentage of tokens occur in the first semantic class than in the second one; such a behavior is followed by 3sg (428/596: 71.8%), 1pl (69/76: 90.8%) and 3pl (285/351: 81.2%) subjects in the data. Nevertheless, given the predominance of speech utterances in the data, it is relevant that 1sg subjects not only are almost equally distributed between speech (393/819: 48%) and cognitive verbs (423/819: 52%), but also show a slight tendency towards the latter class. Given the predominance of speech verbs in the data, said distribution is particularly important because it suggests some subjects are used more frequently in some verbal classes than in other ones.

Another interesting distribution comes from 2sg subjects. While *usted* subjects do appear more frequently with speech (80/135: 59.3%) than with cognitive verbs (55/135: 40.7%), *tú* subjects show the opposite tendency; they are more common in utterances with cognitive verbs (70/105: 66.7%) than with speech ones (35/105: 33.3%). Such a behavior suggests the contrast of *tú* vs *usted* may depend not only on pragmatic features

(e.g. politeness towards the interlocutor), but also on the existence of prefabs and collocations associated with one subject or the other.

Focusing on the relation between subject and tense, 1sg subjects appear mainly in present tense utterances (206/393: 52.4%) and then in preterit tense utterances (116/393: 29.5%); 3pl subjects also prefer present tense (150/285: 52.6%), but they are much more frequent in imperfect (96/285: 33.7%) than in preterit tense (23/285: 8.1%). Present is also the most frequent tense for 2sg-*tú* (18/35: 51.4%) and 2sg-*usted* (42/80: 52.5%); unlike the *tú* form, the *usted* one has a percentage of occurrence in preterit (15/80: 18.8%) close to the one in imperfect tense utterances (12/80: 15%). With a different distribution, most 3sg subjects appear in preterit (228/428: 53.3%) and then in present (114/428: 26.6%) and imperfect tenses (64/428: 15%), while 1pl subjects are more equally distributed in present (26/69: 37.7%) and in imperfect tense utterances (27/69: 39.1%).

From the perspective of tense, there are also important differences. Even though neither present, nor preterit nor imperfect utterances are evenly distributed between the three most frequent subjects in the verbal class, the disproportion is larger in some cases than in others. Thus, while preterit utterances usually take 3sg subjects (228/394: 57.9%) and, in a minor portion, 1sg (116/394: 29.4%) the percentage gap between subjects is not as large in present and imperfect tense utterances. For present tense, there is not a huge difference between 1sg subjects (206/569: 36.2%), the most frequent, and 3pl (150/569: 26.4%), the second one in frequency; let alone between 3pl and 3sg subjects (114/569: 20%). Although with a different order, imperfect tense utterances show a similar

percentage gap between the most frequent subjects: 3pl (96/258: 37.2%), 3sg (64/258: 24.8%) and 1sg (44/258: 17.1%).

Moving on to subject-tense combinations, only three of them have a percentage of occurrences larger than ten percent: 3sg-preterit (228/1320: 17.3%), 1sg-present (206/1320: 15.6%) and 3pl-present (150/1320: 11.4%). These three combinations are also the most frequent in the whole data after 1sg-present utterances with cognitive verbs⁶³, factor which may cause the emergence of prefabs and collocations from their in the discourse. As for imperfect tense, 3pl-imperfect (96/1320: 7.2%), the most frequent combination, accounts for less than 10 percent of the class, and it occupies only the sixth position in frequency after 1sg-preterit (116/1320: 8.8%) and 3sg-present (114/1320: 8.6%). Related to subject-tense combinations, it is worth mentioning that most 3sg-preterit (228/243: 93.8%), 1sg-preterit (116/127: 91.3%) and 3pl-imperfect utterances (96/116: 82.8%) in the data occur with a speech verb, which suggest the reference to past discourse is a major phenomenon in the interviews. Although general for the class, these results imply that, in the same discourse, speakers use each subject-tense combination differently; thus, each combination expresses particular communicative situations and, consequently, performs different functions. So far, patterns suggest that, when reporting another person's speech, speakers tend to use preterit utterances (e.g. *dijo* 'he/se said'), but when reporting speech from generic referents, they tend to use 3pl-imperfect utterances (*decían* 'they said'). Finally, when expressing their own speech, they prefer to do it in present tense (e.g. *digo* 'I say'). In the following subsections, I will discuss how

⁶³ 1sg-present-cognitive (382/2134: 17.9%), 3s-preterit-speech (228/2134: 10.7%), 1sg-present-speech (206/2134: 9.7%), 3p-present-speech (150/2134: 7%)

those patterns relate to speakers' need to contrast what different people say and, particularly, how speakers' expression of saying compare to the expression of thinking.

6.2. Speech Verbs: Local Patterns

In terms of tense, the two most frequent verbs in the class mirror general patterns for the class in terms of type predicate: predicates with simple nucleus represent 93% (794/853) of *decir* and 84% (315/375) of *hablar*. Distribution for tense shows a similar behavior; present tense, the most frequent in the class, accounts for 40% (341/853) of all *decir* (246) utterances, and for 50.4% (189/375) of *hablar* (247).

(246) A: (H) *Y no se acuerda usted de --
de refranes o,
o= sayings no.*

S: *Me acuerdo [muchos].
[...]*

A: *Los dice.*

S: *Los digo*. (102-3A2: 147)

'Do you remember any proverbs⁶⁴ or, or sayings? No? S: I remember a lot of them. A: Do you use them? S: **I do use them**.' (lit. Do you say them? I do say them.)

(247) E: *Esa era la siguiente pregunta.
(H) Es que --
usted habla el inglés también,
.. verdad?*

U: *Yo **hablo** (Ø).
sí,*

*muy --
muy regular.* (214-1A1: 676)

'E: That was my next question. That.. You also speak English, don't you? U: **I speak [it]**, yes. Very fairly.

When focusing on distribution for subject, however, the situation is different; although both verbs appear with the three most frequent subjects for the class (1sg, 3sg and 3pl), their order and percentages differ from each other. *Decir* occurs more frequently with 3sg (321/853: 37.6%), followed by 1sg (305/853: 35.8%), and 3pl

⁶⁴ In the original utterance, the interviewer utters 'sayings' immediately after *refranes* as a translation for the interviewee. I use 'proverbs' (Spanish *proverbio*) to avoid the repetition of 'sayings'.

(146/853: 17.1%). As for *hablar*, it leans toward 3pl (110/375: 29.3%), 3sg (87/375: 23.3%) and 1sg (66/375: 17.5%). Based on the dominance of said patterns, it is likely that the most frequent constructions for each verb will emerge from the convergence of those features, and, especially that they will differ in terms of form and function. For matter of consistency with section 5.2 in the previous chapter, I will discuss specific constructions based on their subject: 1sg, 3sg, and 3pl.

6.2.1. 1sg-Speech Constructions

As expected, given the patterns described in the opening of this section (6.2 above), constructions that show certain loss of morphosyntactic variability and convey a degree of subjectivity involve the form *digo* ‘I say’. In general, the item appears 120 times, accounting for 14.1% (120/853) of all *decir* utterances, 30.5% of all 1sg subjects in the class (120/393), and 21.1% (105/569) of all present tense utterances with a speech verb. Based on token frequency, one construction emerges as the predominant one: (yo) (LE) *digo* , ‘I say (to him/her)’ where the parentheses indicate the presence/absence of subject and indirect object pronouns is still variable, and the small caps for LE indicate that any pronoun in the paradigm may fill the slot. Beyond token frequency, the co-occurrence of *digo* with specific grammatical and lexical items suggests that three other constructions are emerging, either from the general 1sg-present-*decir* combination, or from the (yo) (LE) *digo* construction: *como (TE) digo* ‘As I say (to you)’, *por eso digo* ‘that’s why I say (so)’, and *ya te digo* ‘that’s what happened’.

6.2.1.1. (yo) (LE) *digo* ‘I say / I tell him-her’

This construction appears 97 times in the data, which makes it the most frequent of all utterances with *digo* . It predominantly appears in affirmative utterances (93/97:

95.9%) and with unexpressed subject (69/97: 71.1%), as in (248); when expressed, the pronoun follows the verb more frequently (17/97: 17.5%) than it precedes it (11/97: 11.3%), as in (249) and (250), respectively.

- (248) [2No hay nada shameful,
*le **digo2**,*
about our language.
A mí no me importa,
le digo,
quien me oiga a mí hablar en mexicano que porque –
aquella habla nomás en inglés.
Yo estoy hablándote aquí en mexicano a ti.
It's too bad .. si ella no me entiende a mí.
Porque si ella habla en inglés yo la voy a entender a ella.
Porque yo sé los dos. (117-1A3: 848)
 ‘There is nothing shameful, **I say** to her, about our language. I don’t care, I tell her, who may hear me speaking in Mexican because... she only speaks English. Here, I am speaking to you in Spanish. It’s too bad if she doesn’t understand me, because I am going to understand her, because I know both [languages].’

- (249) *Y fue esta señora a pie,*
hey Henry me dijo,
Quiere hacer un favore,
que está más XXX que la chingada,
y no vaya a empezar.
y qué pasa,
 digo yo.
Sabes tú,
dijo,
está mi hija muy enferma,
dijo. (156-1A1: 1115)
 ‘So this lady came walking all the way to here, “hey, Henry”, she told me. Do you want to do me a favor, it’s farther than hell, so don’t start complaining.” “What happens”, **I say**. “You know?” she said, “my daughter is very sick”, she said”.

- (250) L: *Pero los mestros de antes,*
no enseñaban más de a –
a darle fregazos a [uno],
como quien dice.
 E: [@@@] <@ Sí @>.
 Yo le digo mira,
yo,
en la edad que estoy ahora,
mis mestros míos,
hoy yo estoy de colegio.
 E: Sí.

L: *Ellos no sabían nada nada inglés (H) y ponían los libros en inglés,
 ..americanos,
 pero,
 no lo sabían,
 pues,
 qué iban a explicar?*

E: (H) *Entonces eh- --
 hablaban en español?*

L: *Puro español.* (219-1A1: 128)

‘L: But, in the past, teachers didn’t teach you anything but... they only knew how to beat you up, one could say. **I say to him:** “look, I, right now, at this age, I am in college, now, my teachers... E: Right. L: They didn’t know English at all, but books were written in English, American books, so they [i.e. teachers] didn’t know what they were supposed to explain. E: So they would speak Spanish? L: Only Spanish’.

Although syntactic factors (e.g. contrast or focalization) may still be the decisive factor for pronoun’s presence/absence in this construction, it is likely that they will eventually be less influential if any of the items labeled as (*yo*) *digo* increases its production and perception as a whole unit rather than as a subject-verb sequence. In fact, various scholars (Fernández Bernárdez 2002; Fuentes Rodríguez and Alcaide 1996) have claimed the status of *digo yo* ‘Say I’ as a separate unit; one functioning as a modal discourse marker to comment about or emphasize what has been just said has been documented by various scholars. The collocation barely appears in the interviews (n=17); situation that in certain way agrees with the comment of said scholars regarding the marginality of the marker. It is important to mention, though, that the construction always appears in its own intonation unit, as a parenthetical or closing unit.

(251) [*y por eso*],
eh digo yo y,
y yo sé que .. te vas adelantando,
% % te ayuda mucho,
en otras palabras. (190-3B2: 105)

‘and that’s why, eh **I mean**, and I know you get ahead, in other words, it helps you a lot.’

Most of the items include an indirect object pronoun (76/97: 78.4%), which implies the speaker is either addressing the interviewer or indicating the presence of an

interlocutor in the story he/she is narrating, as in (252). Structurally, the fact that the pronoun changes accordingly to the reference of said interlocutor causes the construction to have certain degree of morphological variability: it has an open slot to be filled by members of a specific morphological category. The construction, hence, is a good example of a prefab, which can be represented as (yo) (LE) *digo*, where the verbal item is invariable, (LE) indicates a variable indirect object pronoun, and (yo) a subject which may be tacit or explicit, and, if the latter, may precede or follow the verb. With said status, the prefab would be a realization of a general construction that could be described as (Sbj) DECIRLE [ALGO] ‘Sbj-say-[SOMETHING]’. The construction presents information to someone else, and it has a verb that can be realized in different tenses, and whose subject may be expressed or unexpressed.

(252) B: *A mí no me gusta --*

.. gastar el dinero.

O: *No=. no. [pues, yo le digo a él],*

B: *[Que soy muy <@ apretada @>].*

O: ***Le digo yo a él que no tiene que ir uno a gamble,***

B: *[No]? (311-4A1: 290)*

B: ‘I don’t like to spend waste money. O: No, no. Well, **I say** to him... B: I am very stingy. O: I say to him that one must not go to gamble. B: No?’

On the contrary, the lack of the indirect object pronoun reduces or eliminates the assumption that the utterance is addressed to someone in particular, which gives it a more generic scope, as shown in (253). In terms of structure, by not having an open slot, the construction has a higher degree of fixation than the prefab mentioned in the previous paragraph; consequently, it is likely that this subject-tense combination may become a collocation as its meaning grows dependable to the context in which it is used. If so, it can be represented as (yo) *digo* ‘I say’, where, as in the previous prefab, the subject may be unexpressed or expressed either before or after the verb:

(253) P: *En ese tiempo.*
Mi papá es cuando trabajaba en las minas.

A: *Mhm.*

P: *antonces eh= --*
lo que hicían no --
no nomás estoy devi- --
devinando. ((ADIVINANDO))
chanza .. que .. el veinticinco del mes,
digo,
en ese tiempo se ganaba buen dinero.
.. No más lo veían a él y dijían, ((DECÍAN))
Diosito,
aquí viene un hombre muy rico. (88-1A2: 415)

‘P: In those times, my dad worked in mines. A: Mhm. P: So, what they use to do, no? I am only guessing, maybe around the 25th of each month, **I mean**, back then you could make good money, people saw him and they use to say: God, here it comes a very rich man’.

Based on the structural differences and their implications, Fernández-Bernárdez (2002) and others argue that rather than variants of the same construction, the form with the pronoun (i.e. *yo le digo* ‘I tell him-her’) and the one without it (i.e. *yo digo* ‘I say’) constitute separate units. Thus, recognizing the existence of an interlocutor constraints the use of (yo) *LE digo* to dialogic units of discourse, while the absence of a specific recipient for the utterance allows (yo) *digo* ‘I say’ to be used in both dialogic and monologic discourses, although with a tendency towards the latter one.

On the topic of their function in the discourse, Fuentes (1990) claims that both constructions, despite their differences, have a “cohesive value” because they are frequently use to “continue and thread the discourse, after a pause or digression”. Such a function can be observed in *yo le digo a él* and *le digo yo* ‘I tell to him’ in (252) as well as in *digo* in (253) above.

Both constructions perform too other functions related to the communicative nature of the complement, if any, that they take. Thus, although, speech events is the most frequent complement for (yo) (LE) *digo* (50/76: 65.8%) and (yo) *digo* (13/21:

61.9%), they have a different purpose in the discourse and, consequently, pursue different communicative goals.

One of the most evident functions for (yo) (LE) *digo* is to report what the speaker said in a past event, but as if it were happening at the moment of the conversation. Thus, by combining the strategy of direct speech with the aspectual value of historical present, speakers manipulate the flow of the discourse to create the illusion that the event is happening at the moment and, in this way, emphasize the importance of what they say and, implicitly, their own participation in the event.

(254) (H) *Y luego cuando acabó con mi hermana le digo,*
usted,
le digo. (318-1A3: 143)

‘And then, when he was done beating my sister, **I tell him**: “well, you have no right to hit me because I do not live with you”, I tell him.

Nevertheless, the historical aspect of (yo) (LE) *digo* is not always very concrete because the reported speech comes from a series of similar events, not from a unique, specific and time-limited one. Because of that, the interlocutor cannot assume that the reported speech was uttered every time exactly as the speaker is expressing it, but that it expresses the speaker’s continuous opinion or belief concerning the topic under discussion. In (255), for instance, the speaker is telling the interviewer that she has had some health issues and, for that reason, her kids want her to rest as much as possible, but she refuses to do so by saying *Yo no nací pa’ sentada* ‘I was not born to be sat down’. The speaker, then, is not really quoting herself, but making a statement that applies to her life in general, not just to the specific event she is narrating: she is a very active person and cannot be without doing anything.

(255) *Yo les digo,*
Yo no nací pa’ sentada. (219-1A2:224)
I tell them, I wasn't born [to be] sat down'

Such a blurry boundary between actual and apparent direct speech makes evident that constructions develop gradually from previous ones and, consequently, they form a continuum in which they may share features and functions with those construction from which they arise. In (256), for instance, although the speaker reports something she said to her son in previous conversations, the habitual aspect of (yo) (LE) *digo* and the opinion expressed by the reported speech are more prominent than the temporal frame of the event being narrated; as a result, the speaker stance, her pride for her native language, becomes atemporal rather than constrained to a specific past conversation.

- (256) [2*No hay nada shameful,*
*le digo*2],
about our language.
A mí no me importa,
le digo,
quién me oiga a mí hablar en mexicano que porque – (117-1A3: 853)
 ‘There is nothing shameful, **I say** to him, about our language. I don’t care, **I say** to him, who may hear me when I speak Spanish because...’

Prosodically, (yo) (LE) *digo* and the speech it reports tend to appear in different intonation units (40/50: 80%), with the former often following the speech it supposedly introduces, as in (257) below, or even inserted in it, as in (259) above. The latter two positions are particularly relevant for the argument of this study because the displacement of the verbal construction to parenthetical positions makes clear that it does not form a syntactic unit with its complement and, as a result, it is not being used accordingly to its normative transitive structure.

- (257) K: *hace cuarenta y siete años que vivo aquí en Dolores.*
 A: *Mhm.*
 K: *Yeah.*
 A: *Mhm.*
 K: *Pa= --*
Pa' Dolores Day que celebran aquí,
.. les digo,
me debían de poner en una vez a mí de queen también porque --
 A: *@@[@@]*

K: [ponen] a todas las viejitas y uno que –
 A: Oh sí?
 K: Que jo- --
 que están aquí hace años y años,
 pues,
 ya se va a hacer cuarenta y siete <@ años que estoy @>.
 A: @@@@
 K: pero no me – no me chusan yo creo que ni --
 .. ni me van a chusar nunca. (318-1A2: 241)
 ‘K: It has been seven years that I have been living here, in Dolores. A: Mhm.
 K: Yeah. A: Mhm. For.. For Dolores’ Day celebration, **I tell** them that they
 should have me as the Queen, at least once, because... A: (Laughs). K: They
 invite all the old ladies, and one that... A: Is that right? K: One that has been
 here for years, well, it is almost forty seven years that I have lived here. A:
 @@@@. K: And yet they never choose me, I don’t think they will ever.’

An important piece of evidence supporting the claim that the meaning of a construction emerges from the discourse rather than being predetermined by its lexical meaning comes from the fact that few of the instances in which (yo) (LE) *digo* is complemented by a subordinate clause correspond to indirect speech, as one could think given the normative syntactic features of these utterances. In other words, although the prefab (yo) (LE) *digo* relates, structurally and semantically, to (yo) (LE) *dije*, the former is not merely the historical-present alternative to the latter; it rather has a different function in the discourse. In (258), for instance, the speaker uses *No le digo que ellos eran los gringos* ‘I’m telling you they were the Americans’ to refer something she mentioned in the conversation, not to quote herself and not even to remind the interlocutor what she said, but to remark the importance of such information in the context. The function of (yo) (LE) *digo*, then, is not to merely provide discourse continuity, but to emphasize a piece of information. Furthermore, the adverb *no* accompanying the prefab causes the utterance to be interpreted as a rhetorical one, which may be an indicator that a new prefab may eventually: *no* (LE) *digo*.

- (258) O: *Como que hablaban mucho inglés [ellos],*
 A: *[Entre] ellos?*
 O: *Entre ellos.*
 A: *[Pero –*
pero con] –
 O: *[y luego,*
y también] a nosotros nos hablaban en inglés.
.. Y en mexicano.
Pero nosotros,
.. yo y mi esposo hablábamos más en mexicano que en inglés.
.. yo y él.
 A: *N- --*
y con ellos les hablaban en español [<X o en mexicano X>*].*
 O: *[Eh,*
en mexicano].
En mexicano.
No le digo que,
ellos eran los gringos.
nosotros no.
@@[@@@@]
 A: *[@@@@] (117-1A3: 715)*
 ‘O: It seems they use to speak English a lot. A: To each other? O: To each other. A: But... but with... O: And later, they spoke English to us too, and Mexican [i.e. Spanish]. But we, my husband and I, would speak Mexican more often than English. He and I. A: N... And did you speak Spanish, or Mexican, to them? O: Eh, Mexican. **I am telling you** that the gringos were they, we were not. (LAUGHS) A: LAUGHS’.

The function of (yo) (LE) *digo* as an strategy to emphasize information is also present in (259), where the prefab introduces a clause expressing an assertion that the speaker has made in several occasions.

- (259) N: *(H) les digo yo que .. mejor trabajaba más antes la gente,*
pero vivía mejor oiga.
Mejore,
... y más saludable.
porque comían de la tierra.
Ahora no.
Pura tienda. (10-3A2: 303)
 ‘N: **I tell them** that people before worked harder that now, but they lived better, you know. Better and healthier, because they ate what they grow. Not anymore. People now buy everything at the store.’

Thus, despite the meaning of (yo) (LE) *digo* is still related to the literal meaning of *decir*, and although the construction is not always syntactically independent from the speech it presents, the prefab is developing context-related functions and semantic

nuances that suggest it has entered into a subjectification path and may eventually arise as a discourse marker.

6.2.1.2. *Por eso digo* 'that's why I say (so)'

Even though this construction has a low token frequency, barely representing 13.3% (16/120) of all utterances with *digo*, its formal features and its function in the discourse suggest it has developed into a prefab by itself. Furthermore, because it tends to occur with pronominal subject in postverbal position (13/16: 81.3%), it can be argued that it is in its way to become a collocation. Regarding its syntactic relation with its complement, the construction is equally distributed in hypotactic (8/16: 50%) and paratactic utterances (8/16: 50%), as shown in (260) and (261) below; in both cases it usually presents, and precedes, linguistic units with finite verbs (13/16: 81.3%).

- (260) W: *Y por eso digo yo que todavía aquí nosotros estamos así.*
'W: And **that's why I think** we're still in such a bad situation.'⁶⁵

- (261) [*Y por eso*] *digo yo que --*
no,
muchas veces nosotros,
por --
porque= --
nosotros no lo tuvimos,
y que --
ah,
queremos .. que tengan lo que nosotros no tuvimos.
.. queremos .. que tengan lo que nosotros no tuvimos.
por eso digo yo,
no sé si será (H) bien o no=. (190-3B1: 355)
'That's why **I say** that... No, a lot of time we... because we didn't have it [when growing up], and that... we want [our kids] to have what we didn't. We want them to have what we didn't have. That's why **I say**, I don't know whether that's good or not'

⁶⁵ The speaker is talking about how discrimination still affects Hispanics in her community.

Like (yo) (LE) *digo* ‘I say (to you)’, the construction is still related to the literal meaning of its verb; however, its function is not merely propositional, but textual and pragmatic as well. It does not merely declare a state of events, it manages discourse continuity while emphasizing the importance, and truth, of what the speaker is saying. By doing so, speakers do not only validate their role as the information source; they also present an argument as factual, therefore irrefutable. Such a function is illustrated by the previous examples, and by (262), where the emphatic aspect of the construction is evident in the fact that it occurs after the speaker’s statement.

- (262) W: [*antonces,*
el dinero] era muy escaso.
Y,
so,
ah,
en Christmas hacíanos --
nomás teníanos que comprar un presente para uno porque,
.. pues,
éranos muchos y ni --
pues,
mis papás ya no podían --
 B: [*Mhm*].
 W: [*<X pues X>*],
comprarnos a todos lo que que necesitaban.
Por eso we'd just --
cambiábanos nombres.
 B: *Usted --*
usted cree que eso le ayudó a apreciar lo que [tiene hoy día]?
 W: [*Sí.*
por eso] digo. (190-3B1: 380)

‘W: Money was very scarce. So, for Christmas we would... we had to buy a present just to one person because.. well, we were a big family, so my parents couldn’t... B: Mhm. W: well, get us what we needed. That’s why we’d just... exchanged names. B: You... do you think that helped you to appreciate what you have nowadays? W: Yes, that’s why **I say so**’⁶⁶

⁶⁶ This conversation continues the one shown in example (261) about how young people today do not appreciate what they have, and whether parents are responsible for said phenomenon.

6.2.1.3. *Ya te digo* ‘aprox. that’s what happened’

Another construction that, despite low token frequency, presents a good case for the argument of subjectification due to language usage is *ya te digo* ‘that’s what happened’. It just appears seven times in the interviews, but it always does it with unexpressed subject, with a 2sg indirect object pronoun, and in a relation of parataxis with the speech event it refers to. Said morphosyntactic invariability suggests the unit is emerging as an independent collocation. Furthermore, the constructions shows pragmatic and epistemic nuances because speakers use it not only to reintroduce information in the discourse, but also to explain a situation in the anecdote being told. In (264), for instance, *ya te digo* stresses what the speaker is going to say; he does not want the interviewer to ignore it because it is important for the narration. More importantly, even if the literal meaning of the prefab indicated the information has already been uttered in the discourse, it is possible that, in fact, it presents new details for what the speaker has told.

(263) *M: .. Y luego mi apá sembraba vegetables. ((SIC))*

F: Mhm.

*H: Eh alverjón de pizca,
frijol de pizca,
estas fregaderas,
entiendes?*

Y=,

.. Y ya te digo.

p- ocupaba mucha gente.

Toda la gente de Cuesta y de se Cierra XX,

a --

a arrancar el frijol [ahí]. (156-1A1: 856)

‘M: Also, my dad grew vegetables. F: Mhm. Peas, beans, all of these things.

Do you know what I mean? And, **that’s how things worked**, he would hire a lot of people. All the people from Cuesta and from Cierra, he would hire them to pick the beans.’

(264) *A: Hm.*

Habían niños gringos en el rincón?

K: Sí.

No=.

@ @

A: [No]?

K: [no].

*y ya te digo que,
que despachó a todos los eh gringos al –
.. al excusado y,
y a los mexicanos no.
.. Entonces vine a figurar,
yo dije pues eh,
seguro que por esto,
mi hijo .. es muy remiedoso,
era miedoso,
(H) y seguro que por eso no quiere ir, (318-1A2: 655)*

‘A: Hm. Were American kids sent to the corner as well? K: Yes. No.

LAUGHS. A: No? K: No. **Let me tell you.** He allowed all the American kids go to the bathroom, but not the Mexican kids. So I figured that... I say, “Well, surely this is why my kid is so frightened”, he was frightened; “surely that’s why he doesn’t want to go’.

6.2.1.4. *Como (TE) digo* ‘As I say / tell (YOU)’

This construction, the last of those involving *digo* ‘I say’, is another marginal unit in terms of token frequency, occurring merely 20 times in the interviews. However, as other non-frequent multiword sequences, it shows features suggesting it is acquiring prefab status, and also conveying non-propositional meanings. On the former issue, the construction tends to appear with unexpressed subject (16/20: 80%); when expressed, the subject always follows the verb. It also includes an indirect object pronoun in most cases (17/20: 85%), which suggests the speaker is addressing an interlocutor, or at least acknowledging his/her existence. An interesting phenomenon is that said interlocutor is usually expressed by the pronoun *te*, as shown in (265). The preference for such 2sg-informal pronoun is relevant because it goes against global patterns for the use of formal and informal 2sg subjects; patterns that suggest interviewers and interviewees are conscious about the *tú-usted* contrast and its social implications. Thus, since, as a subject, the *usted* form is more common than the *tú* one, it would be expected to find a similar phenomenon for the use of indirect pronouns. A possible answer for such an opposite

behavior may be that formal treatment is unidirectional; because, in general, interviewers are younger than the interviewee, *usted* forms are used when the former address the latter, but not in the opposite direction. Another plausible explanation is that the pronoun *te* is becoming morphologically invariable, as it has occurred to *le* in constructions such as *ándale / ándele*⁶⁷ (cf. (Company Company 2006a; Company Company 2006b). More data is necessary to explore any of these hypotheses.

(265) L: *[El radio ni] me gusta.*

E: *Ahá.*

.. *No le gustan las canciones en español.*

L: .. *No,*

pues,

como te digo,

yo estoy impuesta .. a estar muda .. en el telar,

bailando el día [entero].

E: *[Ahá].*

L: *Y no me gusta que me molesta naide.* (219-1A2: 668)

‘L: Radio, I don’t even like it. E: Aha. Don’t you like songs in Spanish? L: No. Well. **As I told you**, I am used to... be quiet... in my loom, working (lit. dancing) the whole day’. E: Aha. L: And I don’t like to be bother by anybody.’

Regarding its relation to other elements in the utterance, the construction appears as a parenthetical unit: it refers to a segment of speech, expressed by a clause or a sentence, but it does not form a syntactic unit with it. Such a syntactic behavior relates to the functions the unit plays in the discourse.

In a textual level, *como (TE) digo* functions as a mechanism of repetition because speakers use it to refer to something previously said. For Bustos (1996:116), the main role of this kind of units is to “guarantee proper levels of redundancy that allow for the right textual interpretation. [...] On the hand, they function as reminders of known information. On the other hand, by updating some concepts in particular, they allow

⁶⁷ *Anda* is the *tú* form for *andar* ‘lit. to walk around’ in present tense; *ande* is the *usted* form.

speakers to choose one concept or another among all of those which are available in the text⁶⁸. At the same time, the unit has different functions in the discourse because reactivating known information is not a gratuitous action; speakers have a communicative intention, even if unconsciously: maybe the information is relevant for the current topic of conversation, maybe it is the starting point of a new topic or maybe it was not exhausted because there was an interruption or digression in the discourse. In either case, *como (TE) digo* ‘As I say’ allows the speaker to resume the flow of information, to move into a new direction, or to emphasize the upcoming segment of speech. In (266), for instance, the speaker is telling the interviewer all the chores she used to do while growing up on a farm; she has listed several activities such as feeding horses early in the morning, making tortillas, doing the dishes, plowing the soil, etc.; she then uses *como te digo* to introduce a statement that both summarizes all those activities and explains why she is still a very active and strong woman. More importantly, the statement remarks how unique she is because she developed skills traditionally associated with men.

(266) L: [Yo] *manejé un arado.*
bueno,
lo que --
.. Como te digo,
como fue todo lo del rancho,
lo manejé yo. (219-1A2: 52)
 ‘I used a plow. Well, what... **As I told you**, everything in the ranch, I managed.’

The following fragment of a conversation illustrates another discourse function of *como (TE) digo*. In this case, the information that the speaker re-introduces works as an

⁶⁸“los mecanismos de repetición tienen como función primordial garantizar los niveles de redundancia exigidos a la hora de facilitar una correcta interpretación textual. Su existencia tiene dos grandes razones de ser. De un lado, porque sirven de recordatorio de contenidos ya conocidos. De otro, porque en el momento de actualizar determinados conceptos, permiten seleccionar unos u otros dentro del total que aparece en el texto.”

explanation of why she assures her English is better than most people's. Nevertheless, because the explanation is expressed with a verb in the conditional form, the statement is understood as a polite one, causing *como digo* to perform a mitigating function.

- (267) O: Y=,
 como digo,
 no lo aprendería a hablarlo tan fluently,
 .. you know?
A: Mhm.
O: *inglés.*
 But,
 nunca nunca nos atrasamos. (117-1A2: 127)
 O: 'And, **as I say**, I might not have learned to speak fluently, you
 know?' A: Mhm. O: English. But we never got behind in life.'

Regarding the status of *como (LE) digo* as a construction, Fernández-Bernárdez (2002) argues that it has a very low degree of lexicalization because the verb can appear in a subject-tense combination other than 1sg-present, as illustrated in (268) below. Thus, rather than a prefab with its own meaning, it is merely a realization of a more general construction '*como DECIRLE*'. Nevertheless, following Croft and Cruse's (2004) claim that a construction involves both form and meaning, it is undeniable that *como (TE) digo* express a different nuance than *como (LE) dicen* 'As they say', *como (LE) iba a decir* 'As I was going to say', and others. Thus, while in (268) below *como (LE) dicen* has an evidential interpretation because it refers to what other, non-specific, people said, even if the speaker use the expression to support his argument, in (267) *como te digo* shows an epistemic value because it introduces a statement with which the speaker is explaining her skills in a humble way. In addition to these pragmatic factors, the fact that *como (TE) digo* is more frequent than other related constructions with *decir* suggests that it is emerging as a separate prefab status by its own, and that the only variable features are the presence and reference of the indirect object pronoun.

(268) B: *What advice would you [leave them]?*

W: *[Pues yo],*

yo digo=,

.. que yo,

todo lo que tengo,

yo --

yo me ha tenido --

yo ha tenido que trabajar.

y no --

y no --

there's no --

.. como dicen (H) los americanos áhi.

There's no free meal.

.. Tienes que trabajar for whatever you get. (190-3B2: 28)

'B: What advice would you leave them? W: Well, I say that I... that everything I have, I have had... I have had to work for it. There's no...

There's no, as **Americans say**: "there's no free meal". You have to work for whatever you get'

6.2.2. 3sg-Speech Constructions

The major difference when studying the occurrence of 3sg subjects with speech verbs is that said person accounts for a larger percentage of all utterances with *decir* 'to say' (321/853: 37.6%), than of all utterances with *hablar* (87/375: 23.2%). Differences are even more evident when examining how such utterances are distributed for tense; while 3sg-*decir* tokens occur primarily in preterit tense (216/321: 67.3%), and then in present tense (68/321: 21.2%), 3sg-*hablar* utterances tend to appear in present tense (44/87: 50.6%) and then imperfect tense (27/87: 31%). Since both lexical items predominantly appear as the predicate's nucleus, these numbers suggest that the most frequent constructions in the class involve *dijo* 'he/she said', *dice* 'he/she says', *habla* 'he/she speaks', and *hablaba* 'he/she used to spoke'. Hence, when talking about another person's discourse, speakers seem to frequently do it as indirect speech, but also as direct speech; on the other hand, when talking about speech as an activity or ability, speakers refer to current states more often than to habitual past ones.

6.2.2.1. (X) *dijo* 'X said'

This construction occurs 128 times in the data; thus representing almost a third of all 3sg subjects (128/428: 29.9%) and half all 3sg-preterit utterances in the speech class (128/228: 56.1%), as well as 59.2% of all 3sg-preterit utterances with *decir* (128/216). In most cases, the construction occurs with unexpressed subject (105/128: 82%); when expressed, it is realized by a nominal phrase (14/128: 10.9%). These percentages suggest speakers often use *dijo* to refer what a specific person said, and, when doing so, they apply the system of their native language despite intense contact with English.

The construction largely refers to speech events, although without forming a syntactic unit with it. Said paratactic correlation is observed in 128 of the occurrences; indicating the major purpose of the construction in the interviews is to report direct speech. As it is well known, the formal characteristics of this strategy are the replication of the tense and the deictic elements used by the actual source of the quote, as well as the syntactic independence of the latter given the lack of any subordinating conjunction. Based on these structural features, as well as on their communicative implications, I argue that (X) *dijo* constitutes a prefab with a slot for subject (expressed or unexpressed) and whose main function is to report direct speech. Also, the prefab is independent from, although related to, to the construction emerging from contexts in which information is reported as indirect speech (section 6.2.2.2 below).

(269) B: *yo me acuerdo de que= mi mamá iba a .. dar ahm --
luz.
a mi hermanita.
y había algo de que s- --
mi abuelo **le dijo**,
.. hay luna llena,
tienes que salir,
ponerte unas llaves en la cintura,
y voltearte o algo así.*

Por qué hacían eso? (76-1B2: 271)

‘B: I remember that my mom was going to give birth to my little sister, and there was something about... and then my grandpa **told her**: “There is a full moon. You have to go out, tie these keys around your waist, and turn around” or something like that. Why did they do that?’

The syntactic independence of the construction allows it to appear not only before the information it reports, but also after, and even in the middle of, it. As a consequence of said freedom, the construction has the capability of introducing all sorts of linguistic units, for instance, clauses and sentences as in (270) below, adverbs that summarize a whole situation, as in (271), and nominal phrases, as in (272).

(270) *Y esa vez que fui la última vez,*

*.. **me dijo**,*

<VOX pues yo te diré,

he estado gastando mucho dinero en el doctor VOX>,

dijo,

yo no tengo dinero ahora dijo. (214-1B3:519)

‘And the last time I went over there, **she told me**: “I have been spending a lot of money on doctors”, she said, “I don’t have any money now”, she said”’.

(271) *A: Y qué hizo?*

K: Se quedó callada.

*no **dijo** nada.*

‘A: And what did she do? K: She remained silent. **She said** nothing.’

(272) *Y: [dijo] el doctor que tenía arthritis,*

and a bad disk.

and both the medical and the chiropractor,

.. and all of them say the same thing,

so,

.. Pero este doctor,

si no,

tienes que seguir un plan.

.. Definite plan,

***dijo.** (147-2A2: 291)*

‘Y: The doctor said I had arthritis [...]. But this other doctor, if you don’t, you have to follow a plan. “Definite plan”, **he said**’.

It is precisely the freedom to appear as a parenthetical unit what promotes the development of non-propositional nuances; because the verb’s scope may affect the whole quote or just part of it, speakers may use the construction to emphasize, explain or

clarify a particular element of what is being reported. For instance, in (272) above, the speaker uses *dijo* ‘he said’ to specify what kind of plan the doctor wanted her to follow, while in (273) below the construction focuses the attention on the phrase *qué garrero*, thus emphasizing a defect of the person about which the speaker is talking about.

(273) *le dijo de la puerta,*
le dijo.
qué garrero,
*le **dijo**,*
tienes aquí,
le dijo. (10-8A: 449)
 ‘he said to her from the door, “What a mess”, **he said**, “you have here”, he said to her.’

Another consequence resulting from its syntactic independence and parenthetical position is the use of two or more instances of the construction in the same reported quote to break it in two or more pieces, with a token of *dijo* accompanying each one, or at least most, of them. In (274), for instance, the speaker breaks the quote *Bueno, pa’ la tarde estás lista* ‘well, you must be ready in the afternoon’ in three pieces, intercalating two items of *me dijo* ‘he told me’. In a sense, then, the verb functions not only as a mechanism for direct speech, but also for discourse continuity: it allows the speaker to buy time while thinking what the next part of the quote is, and also reminds the interlocutor that the information being communicated is not the speaker’s, but someone else.

(274) *No,*
le dijo.
.. Primero quiero verla muerta que --
que --
que se case contigo.
... Sacó el muchacho la pistola y le dijo,
bueno,
*le **dijo**.*
.. Mató a la pobre muchacha.
y arrancó a huir. (311-3B2: 557)

“No”, she [a girl’s mother] said to him. “I rather have her dead that... married to you”. So the guy took his gun and said to her: “Okey”, he said. He killed that poor girl and fled’.

The use as a mechanism for discourse continuity, together with its syntactic independence may be the cause for another observed phenomenon for *X dijo*: speakers often use two instances of the construction not to refer to different pieces of the same report, but to actually report the same speech. In such cases, only one item was coded as having a complement, while the other was coded as having no complement. In (275), for instance, although the quote marked by <VOX VOX> has already been introduced by a token of *dijo*, it is followed by another token of the same item.

(275) *Y esa vez que fui la última vez,*
 *.. me **dijo**,*
 <VOX pues yo te diré,
 *he estado gastando mucho dinero en el doctor VOX>,
 dijo,*
 yo no tengo dinero ahora dijo. (214-1B3: 519)
 ‘And the last time I went over there, **she told me**: “I have been spending a lot of money on doctors”, **she said**, “I don’t have any money now”, she said’.

Because the direct speech function has already been performed by the first occurrence of *dijo*, the second one must have another functions, otherwise it will be superfluous. One possibility, of course, is that the construction is linking the first part of the quote with the second one (*yo no tengo dinero ahora*), nevertheless, such an answer only causes us to deal with the issue later, when we get to the third instance of *dijo* at the end of the quote. Also, when uttering the whole idea, the intonation indicates the first and second instances of *dijo* refer to the first part of the quote, and the third instance reports the third part. An alternative explanation, then, is that, given that one of the instances of *dijo* does not really have a complement, even if it is clear in the discourse to what information is related to, the independent *dijo* emphasizes the importance of a part of the speech, or merely indicate the end of a quote.

Once the verb appears without its own complement because it is sharing it with another item, the next step is to appear in an utterance with no complement at all, if the context allows the interlocutor to understand what the speaker is talking about. Thus, although the first entries of *decir* in DRAE define the verb as a transitive one, in oral conversation *decir* is often used without a direct object, even if it is obvious that someone said something to someone else. In (276), for instance, it is clear that what was said in *¿quién le dijo?* refers to the whole situation narrated by the speaker; nevertheless, the question does not include the pronoun *lo*. Even though this kind of utterances appear just twenty times in the data, by not requiring a grammatical element to form a perfectly functional unit, they are relevant because they provide evidence for Thompson's (2002) argument against the standard view of complements as subordinate clauses in a relation of hypotaxis with a main verb, and for a more appropriate treatment of complement-taking predicates as schemas consisting of an epistemic, evidential or evaluative fragment (in this case, *dijo*), and a clause (in this case, the non-expressed, but known information). More importantly, by showing that "transitive" verbs can be used by themselves also in Spanish, the construction *dijo* provides evidence for the study of complement-taking predicates and conversation across languages.

- (276) U: *[Es lo que] me dijo la muchacha aquella cuando me quería ir yo,
me dijo,
ahora me voy a comprar una máquina de esas,
pa' cuando vuelvas a hablarme.
que --
que --
que me digas una cosa que no me guste a mí,
yo te voy a <X grabar [durante] X> todo,*
E: *[Quién]?*
U: *me dijo.*
E: ***Quién le dijo?***
U: *<@ Aquella con la que me peleé el otro día,
con ella @>
se [acuerda]? (214-1B2: 755)*

‘U: That’s what that girl told me when I wanted to go. She said to me: “now I am going to buy one of those machines, for the next time you speak to me. If you say to me something I don’t like, I am going to record you.”’E: Who? U: She told me. W: **Who told you so?** U: That girl, the one I had a fight with the other day. Do you remember?’

6.2.2.2. (X) *dijo que* ‘X said that’ + [clause]

When reporting what another person said, speakers do not always apply the direct speech strategy, but also the indirect speech one. The construction represents 8.4% (36/428) of all 3sg subjects and 15.8% (36/228) of all 3sg-preterit utterances in the speech class, as well as 16.7% (36/216) of all utterances with *dijo*. In terms of structure, the main difference between this construction and the prefab discussed in section 6.2.2.1 is that it introduces a clause that functions as the verb’s direct object and which requires the presence of a subordinating conjunction or an interrogative pronoun. As anticipated by normative accounts of Spanish grammar, the clause always appears after the verbal construction, as shown in (277). Furthermore, sporadic utterances in which speakers introduce a report in English suggest that *que* plays an essential role as a linking element; the phenomenon is illustrated in (278). Due to the hypotactic relation it establishes with its complement, the construction lacks mobility, which, in turn, prevents the development of epistemic nuances.

(277) *y al fin vino el % --
 .. el bartender,
 y les preguntó que si qué querían,
 (H) y yo creo que el bartender tenía medio miedo a --
 a los .. cowboys,
 (H) y que le **dijo mi viejo** que querían dos cervezas.* (318-1A2: 349)

‘Finally, the bartender came and asked them what they wanted, and I think the bartender was afraid of the cowboys⁶⁹, and **my dad told him that they wanted two beers.**’

⁶⁹ The speaker is telling a story about American cowboys being upset because her father and uncle sat at the bar.

(278) W: *mi sobrino,*
*.. **dijo** que he had a fight in school today with a bunch of kids,*
you know.
we were standing in line,
dijo,
(H) y unos que vino y agarró un can opener,
de un can opener,
lo hizo puntiagudo de sharp,
y dizque,
he cut up some kid in his <X mouth X>. (190-3B2: 872)
 ‘W: My nephew, **he said** that he had a fight in school [...], he said, and
 some kid came and grabbed a can opener. He got it to be sharpened, and then
 he cut up some kid in his mouth.’

Another important difference related to subject expression; although unexpressed subject is still the most frequent one (21/36: 58.3%), it is much less frequent than it is in the prefab for direct speech. Furthermore, when expressed, the percentages of nominal phrases (8/36: 22.2%) and pronouns (7/36: 19.4%) are very similar, which suggests the construction shows variability not only in terms of subject expression, but also in terms of subject realization.

Considering formal features and the function it has in the discourse, the construction discussed in this section constitutes a schematic prefab that may be described as (X) *dijo que* + [clause], where an invariable verbal item has an open slot for subject and another for a clause reporting speech. The subject shows variability in terms of expression and realization, and the clause shows variability in terms of expression. As for the adjective ‘schematic’, it seeks to indicate that the construction maintains the syntactic properties of the verb and so it follows the normative argument structure of the latter. Said schematic prefab, on the one hand, relates to the independent “direct speech” prefab in the sense that both of them report another’s person speech⁷⁰, and, on the other

⁷⁰ Given its innovative syntactic and prosodic patterns, it is plausible that the direct speech prefab may have arisen from the indirect prefab as an independent construction.

hand, it represents a realization of a more schematic (di)transitive construction with *decir*: (SBJ) *DECIR* + complement.

Related to this construction, there are a few utterances (n= 11) in which the verb's direct object is expressed not by a clause, but by a morphological element such as object and relative pronouns. As expected, such elements appear in their normative position, which may be seen as an obstacle for the development of epistemicity. Nevertheless, this may not be the case because, being clitics, pronouns get attached to the verb, co-occurrence that may eventually result on both of them being interpreted as a whole unit rather than as a morpho-syntactic sequence. Said unit could develop independence from the event the pronoun refers to, thus appearing in a parenthetical position expressing meanings other than the propositional one, particularly if another word becomes part of the verb-complement unit. In (279), for instance, it is clear that the pronoun *lo* 'it' refers "No quiero que ni uno" 'I don't want any of you', quote that in fact is part of a larger reported speech. By including *asina* 'like that', however, the speaker's intention is to convince her interlocutor what she is repeating exactly what someone else said, and it constitutes a strong argument for her case. Thus, although formally *lo* is the verb's direct object, the construction is closer in status and function to the the prefab *X dijo* than to *X dijo que* + clause.

- (279) A: .. *No quiero que ni uno,*
 .. y ***lo dijo*** *asina.*
 Nunca se me va a olvidar.
 Les dijo,
 ustedes pat- --
 patas rajadas,
 que los hablen eh --
 que hablen eh= --
 español.
A: *Qué no quie- --*
 No quería que hablaran español?

P: *Yeah.* (88-1A3: 141)

‘A: I don’t want any of you, and **he said so**, I will never forget it, I don’t want that any of you, Indians, speaks Spanish. A: He didn’t w-- he didn’t want them to speak Spanish? P: *Yeah.*’

6.2.2.3. (X) (LE) *dice* ‘X says (to him-her)’

The construction accounts for more than half of all 3sg-present utterances involving the verb *DECIR*⁷¹ (38/68: 55.9%) and particularly the item *dice* (38/65: 58.5%); the construction represents as well an important percentage of all *DECIR* utterances in present tense (59/342: 17.3%). In most cases, the construction appears with unexpressed subject (21/38: 55.3%), thus implying subject continuity. Nevertheless, when expressed the subject shows variability: it occurs with nominal phrases (9/38: 23.7%) and different kinds of pronouns. Given said phenomenon, it is very possible that the construction experiences fewer innovative uses than those constructions with less morphosyntactic variability.

(280) X: *la chiquita de la Elena.*
ésta sí quiere aprender.
[...]
ella sí,
*va y **le dice** a la Elenita cosas en mexicano.*
A: *Mhm.*
X: *que le enseñan áhi en la escuela.*
.. Y luego me telefonea pa' acá,
*y me **dice** what's this,*
mom?
y ya le digo.
*.. Y **ella le dice** a la= --*
a la chiquita.
y ella sí quiere.
sí.
Ella quiere aprender no más que --
(H) pos,
no lo hablan allá. (47-1A2: 335)

‘X: Elena’s youngest kid, she does want to learn [Spanish]. She does indeed. She goes and **tell her mom** things in Spanish. A: Mhm. Words she learns at school. So my daughter, she calls me, and **tells me**: what’s this, mom? So I

⁷¹Other 3s-present utterances include those in which *decir* is part of a periphrasis or the nominal complement of another verb.

tell her, and **she tells** the little one. She wants to [learn Spanish], but they don't speak it at her home.'

In terms of complement, it shows a tendency to introduce a segment of speech, usually clauses with which it does not establish a main verb-subordinate clause correlation. As it happens with other constructions, the syntactic independence allows the construction to appear in parenthetical positions (14/38: 36.8%).

- (281) Z: *Qué pasa,*
le dije.
.. Pues,
no quiere prender mi troca,
dice.
.. eh,
.. y le dije yo,
ponle la cadena,
le dije.
... Yo --
yo te llevo. (311-3B2: 432)
'Z: "What's going on", I said to him. "Well, my truck won't start", **he says**.
"Eh", I told him, "hook this chain", I told him, "I'll pull you."

Based on such formal features, the construction constitutes a prefab described as (X) *dice* 'X says', where (X) indicates an animate or inanimate subject with specific referent, and the absence of an explicit slot for the complement indicates that the prefab often appears without it. Regarding its function, formal and prosodic patterns suggest the prefab is a mechanism for direct speech; a prefab related to (X) *dijo* but with extra meanings conveyed by the aspect of historical present. In (282), for instance, the speaker alternates both prefabs when telling a story; by doing so, she reduces the distance, both temporal and affective, between the event and her interlocutor, which may result on the latter getting more involved in the narration.

- (282) K: *y había lugare,*
pero adrede se hicieron pa' que no cupieran ellos,
y mi viejo no era dejado.
you know él l- --
mi cuñado sí era muy remiedoso.

(H) *Dizque le **dice** mi --*
.. Vámonos,
Vámonos.
aquí nos van a pegar estos .. gringos nos pegan.
oh no,
nos vamos.
*Dizque le **dijo**,*
nosotros semos ciudadanos de los Estados Unidos,
y aquí nos vamos a estar.

A: *Ahá.*

K: *.. No.*
y nos van a vender cerveza también,
*dizque le **dijo**. (318-1A2: 328)*

‘K: And there was room in the bar, but they wouldn’t let them to sit. But my dad never let anybody to bully him, you know. His brother-in-law, he was very cowardly, so reportedly **he tells** my dad, “Let’s go; let’s go. We’re going to get beat by these gringos... they will beat us.” “Oh. No! We’re not leaving”. **He** supposedly **said** to him: “we’re American citizens, and we are going to stay here”. K: Aha. A: “Not only that. They will serve us beer too”, he apparently said to him.’

Present tense, however, does not only introduce one-time past events as if currently happening, it also presents an idea as a well-known, atemporal fact; something that exists rather than something than someone said. Furthermore, by using (X) *dice* to introduce an idea, instead of simply expressing it, the prefab conveys a nuance of evidentiality: speakers make clear that they are not the source of information and, consequently, cannot be held accountable for any issue that it may arise from the statement. In (283), for instance, the statement *el vino y las mujeres son la perdición del hombre* ‘women and wine are a man’s ruin’ may be polemical, especially because the interviewer (E) is a female, and the interviewee (U) is a male. Thus, to avoid any conflict, the latter inserts, *dice la canción* ‘the song says’, indicating he is not the author of what he says, even if, by expressing it, he seems to agree with the idea.

(283) U: (H) *Que las mu- --*
las mujeres --
y y el vino y las mujeres,
dice la canción,
que es la perdición del hombre.
 E: *Y qué cree usted?*

U: *Pues yo creo que así e=s,
 porque e --
 esas Sugars me tiene a mí muy= --
 @@ loqui=to. (214-1A1: 296)*

‘U: Wom... women and and wine and women, **says the song**, are a man’s ruin. E: And what do you think. U: Well, I think it’s true because those Sugars⁷² made me all crazy”.

To conclude the discussion of this construction, it is worth mentioning that there are some instances (n= 19) in which the verbs introduces direct objects; nevertheless, those complements are expressed by clauses (n= 8) just as frequently as by other units (n= 9) suggesting a present tense alternative to the schematic prefab (X) *dijo que* + clause has not yet emerged. An example of the construction taking a clausal complement is shown in (284). When expressed by a morphological element, the direct object is variable: it can be a nominal phrase or a noun, as in (285), an object pronoun, a relative or a demonstrative, such a variability and the low token frequency complicate the emergence of any non-schematic construction.

(284) K: *Pero,
 pero= mi idomia es mexicano,
 y yo --
 .. Yo no voy a hablar en inglés no más porque me --
 .. **me dice alguien** que hable en inglés. (318-1A3: 544)*

‘K: But my language is Mexican and I... I’m not going to change just because **someone told me** to speak in English’.

(285) X: *la chiquita de la Elena.
 ésta sí quiere aprender.
 [...]
 va y **le dice** a la Elenita cosas en mexicano. (47-1A2: 328)*

‘X: Elena’s youngest kid, she does want to learn. [...]. **She** [i.e. Elena] goes and **says** things in Mexican to Elenita.’

⁷² “Sugar” is how the interviewee calls elderly-women at the nursing home.

6.2.3. 3pl-Speech Constructions

The predominance of *decir* and *hablar* in the speech class is particularly observed in utterances with 3pl subject: 3pl-*decir* represents 51.2% (146/285) and 3pl-*hablar* 38.6% (110/285) of the group; with the former subject-verb combination accounting for 17% (146/853) of all *decir* occurrences in the data, and the latter representing 29.3% (110/375) of all utterances with *hablar*.

Agreeing with general patterns for the class, most 3pl utterances with *decir* have said verb as the predicate's nucleus (144/146: 98.6%), and they are distributed between present (76/146: 52.1%) and imperfect tense (55/146: 37.7%), with preterit having a small representation (11/146: 7.5%). Nevertheless, 3pl-*decir* utterances distinguish themselves from other subject-verb combinations in the data because their subject predominantly refers to generic entities (123/146: 84.2%); in fact, it is precisely this feature which promote the emergence of prefabs conveying epistemic nuances.

As for 3pl *hablar* utterances, they are distributed too between present (60/110: 54.6%) and imperfect tense (31/110: 28.2%), but, unlike 3pl *decir*, the subject's referent is usually a specific entity (89/110: 80.9%). The interaction of this feature with other ones (e.g. variability of subject's realization) causes that no independent prefabs emerge from this particular subject-verb combination.

6.2.3.1. *DECIRLE(s) X* 'to call it-them X'

In comparison to other constructions with *decir*; *DECIRLE(s) X* 'to call it/them X' has a low token frequency (n= 45), and some degree of variability regarding tense. Nevertheless, it includes some features that differentiate it from other constructions in the

data; namely, the generic reference of its subject and the presence of a third person indirect object pronoun.

- (286) U: [*Rastros*] *les dicen*,
 o les dicen huellas.
 [no sé]. (214-1B2: 143)
 Traces, **they call them**, or **they call them** footprints. I don't know.'

In most cases, the construction appears in present tense (33/45: 73.3%), as in (286) above; occurrences that account for a significant portion of all *dicen* utterances in the data (33/75: 44%). Alternatively, it occurs in imperfect tense (12/45: 26.7%); combination that, although rather infrequent, represents 22.2% (12/54) of all utterances with *decían*. An example is shown below:

- (287) A: *Las raíces cómo --*
 cómo se llaman en inglés las raíces.
 N: *Las raíces coloradas? ((SIC))*
 A: *Sí*.
 N: *Yo no sé en inglés cómo se llamará*.
 A: *Son [dulces]?*
 N: *[X] --*
 Sí son dulces.
 coloradas.
 A: *Ajá*.
 N: *Y hay otras blancas que se dan <X ahí X>*.
 *les **decían** tanapes*.
 .. unas [blancas].
 A: *[Tanapes]*.
 N: *tanapes*. (10-3A2: 105)
 'A: Root, how... how are *ráices* called in English? N: Those red roots? A:
 Yes. N: I don't know how they are called in English. A: Are they sweet? N:
 Yes, they are sweet. Red. A: Aha. N: And there are also whites ones that
 grow over there. **People used to call them** turnips. A: Turnips. N: Turnips.'

In terms of structural and semantic features, the imperfect tense variant always appear with unexpressed subject, as in (287) above, while in the present one the subject is occasionally expressed (4/33: 12.2%) by an indefinite pronoun or phrase such as *unos* 'some' or *otros* 'other people', as in (288) below.

(288) U: *Yo digo que son rastros.*

E: *[Y dijo huellas].*

U: *[Porque les] dicen cuando pintan,*
(H) unas les dicen huellas a esas,

E: *Hm.*

U: *Y otros les **dicen** rastros. (214-1B2: 158)*

‘U: I call them traces. E: You also said footprints. U: Because they call them that when they are printed. Some call them footprints, E: Hm. U: and **other people call them** traces.’

Although the complement appears slightly more frequently before the verb (19/33: 57.6%) than after it (14/33: 42.4%), no particular pattern seems to emerge for the feature. No matter the order, though, the verb and the noun it introduces tend to occur in the same intonation unit (30/33: 90.9%); distribution that inhibits the occurrence of the construction as a parenthetical unit.

The purpose of *DECIRLE X* is to indicate how other non-specific people call a particular object, food or place. In a broad sense, the construction works as some sort of copula, with the complement being an attribute rather than a direct object⁷³; rather than expressing a speech event, it establishes an equivalency between two elements: an entity and a name. Certainly, using *decir* to provide the name that other people used to refer to an object or place, rather than *ser* ‘to be’ or even *llamarse* ‘to be named’, may be a mere matter of preference; nevertheless, as it is often the case when a situation can be expressed in different ways, it may pursue as well a particular communicate intention on behalf of the speaker; because interviewees recognize the existence of other linguistic communities, they are able to present information without being responsible for its accuracy.

⁷³ As a third entry, the *Real Academia Española*’s dictionary defines *decir* as *nombrar* or *llamar* ‘to name’, and categorize the verb as a transitive one. Nevertheless, substituting the “direct object” with the pronoun “lo” seems to effect negatively the acceptability of the clause (e.g. *le dicen abeja* / *?se lo dicen*). In fact, the *Corpus del Español Actual* does not include any cases with a direct object pronoun substituting the name of an entity.

Based on the interplay of the discussed factors, *DECIRLE X (a Y)* can be categorized as an evidential prefab consisting of a realization of *decir* and a slot for a required third person, indirect object pronoun, which is co-referential to the entity being named. Regarding the verb, in theory, it may be realized in any imperfective tense; in reality, though, *dicen* emerges as the strongest exemplar. Finally, because its meaning and function are rooted in the lexical meaning of its verb, the generic nature of the subject's referent, and the imperfective aspect of the verb's tense, the prefab can be seen as part of a network of constructions conveying evidentiality.

6.2.3.2. *Dicen (que)* 'it is said' & *decían (que)* 'it was said'

Other evidential prefabs emerging due to the use of third person subjects with generic referents are *dicen* and *decían*. Although closely related in the sense that they communicate other people's ideas, the reason to treat them as separate constructions is the variability in terms of subject expression, even if small, shown by the variant in present tense but not by the one in imperfect one tense.

Dicen appears barely 28 times in the data, frequency that seems low in relation to the whole class, but that it is relevant because it represents 37.3% (28/75) of all utterances with *dicen* in the interviews. Furthermore, the construction clearly shows structural and functional features that indicates its prefab status. First, it always appears in declarative utterances. Second, as aforementioned, its subject is mostly unexpressed (23/28: 82.1%). Third, speakers use the prefab to state an idea that supports, explains, or illustrates the situation they are talking about. However, because the source of information is a non-specific referent, the information needs to be understood as hearsay; thus, no matter whether speakers are judging the quality of an object, as in (289),

exalting their own qualities, as in (290), or presenting information that the interlocutor may find hard to believe, as in (291), they are merely repeating what others have said so they cannot be held accountable for the veracity of said information.

- (289) *P: It's a beautiful --
es libro muy bonito.
eh dicen,
.. Es cuando el --
.. El gringo vino para= Nuevo México, (88-1A3: 77)*
'P: It's a beautiful... it's a beautiful book, **people say**. It's from when the ... Gringos arrived to New Mexico'.

- (290) *L: [Pero] no me muero.
Tengo las ve- --
vidas de los gatos.
E: @@@
[las nueve vidas].
L: [Es que los gatos] tienen siete vidas dicen.
E: [2Aha2].
L: [2ya2] yo me pasé de las siete. (219-1B1: 117)*
'L: But I haven't died. I have cats' lives. E: (LAUGHS). Nine lives. L: Cats have seven lives, **people say**'.

- (291) *R: [You have] to call back.
E: (H) Bueno.
[no sé] --
U: [Y ahora] que --
que van a hacer unas máquinas que --
.. que --
que --
hasta van a agarrar el nombre de uno cuando hablen dicen. (214-1B2: 707)*
'R: You have to call back. E: Well, I don't know. U: And now, they are going to build machines that... that.. that will even get your name when you call; **they say**.'

Regarding the relation between the prefab and the information it presents, results show a tendency towards parataxis (17/28: 60.7%). As can be observed in the three examples above, the prefab's syntactic independence gives it flexibility to appear after the statement it refers to, thus conveying evidentiality while indicating the end of the utterance. In the remaining utterances, the stated information functions as the verb's direct object and, as such, it is introduced by a subordinating conjunction; because of

such syntactic relation, the prefab lacks the ability to appear in parenthetical positions. Furthermore, having or not having said flexibility may result in the emergence of two constructions rather than one, as it was discussed for *dijo* ‘X said’ and *X dijo que* ‘X said that’; unfortunately, there are not enough data to confirm the hypothesis.

- (292) T: *Sí,*
well dicen que ya se murieron,
 (H) *Hay muchos áhi en Gallup. (270-1B3: 374)*
 ‘T: Right. Well, **people say** they already died. There are a lot of them over there in Gallup’.

6.2.3.3. *HABLAR [ESPAÑOL]* ‘to speak Spanish’ & *HABLAR en [ESPAÑOL]* ‘to speak in Spanish’

Despite its high token frequency in the data, *hablar* mostly refers to two situations, both of them directly linked to its literal meaning: the ability and activity of communicating with other people, as in (293) and (294); and the ability to speak a language, as in (295) and (296).

- (293) B: (H) *Se hacía eso aquí también?*
 Z: *Yeah.*
yo hacía mucho eso. ((MUCHO))
 B: *Eso?*
 Z: *Mhm.*
 B: *.. Sacaba la gente a votar.*
 Z: *... Yeah.*
...(2.0) Yo iba a las casas,
 B: *... Mhm.*
 Z: *...(2.0) Les **hablaba** a la gente <@ todo el día @>.*
@ @ @ @ [@ @]
 B: *[Y no] querían ir a votar?*
o qué era? (311-3B2: 40)
 ‘B: Did that use to be done here too? Z: Yeah. I did that frequently. B: Same thing? Z: Mhm. B: I used to get people to vote. Z: Yeah. I used to go to their houses. B: Mhm. Z: **I talked to** people there all day. (LAUGHS). B: They didn’t want to vote or why was that?’

- (294) U: *pues ha- --*
hay m- --
m- --
hay mixed [.. mexicanas] y= gringas,
E: *[Mhm].*
U: *y las que me hablan e- --*
en= inglés, yo les respondo inglés,
las que me hablan español,
yo les respondo español.
E: *.. Parece que **habla** más con las mujeres que con los hombres.*
U: *Hm=.*
Sí,
porque esas me dan más plática.
(H) y luego yo les hago historias y todo,
y= le --
le .. hago reír y todo junto.
E: @@@@ (214-1A1: 819)
‘U: Well, it’s mixed. There are Mexicans mixed with Americans. E: Mhm.
U: And to those who talk to me in English, I talk back to them in English. To the ones who talk to me in Spanish, I talk them back in Spanish. E: It seems that **you speak** with women more than with men. U: Hm. I do, because they chat with me more. Then I tell them stories and everything. I make them laugh and all of that’.
- (295) D: *Y en la escuela=,*
les enseñaban inglés?
o cómo estuvo.
S: *Ya cuando fuimos a la escuela,*
pues
*era muy trabajoso pa’ los otros maestros .. que **hablaban** español.*
... pa’ comprenderlo.
D: *Mhm.*
S: *.. porque=,*
pues,
Todo era en inglés ya en ese tiempo.
Ya la escuela --
.. todos los enseñaban a hablar en inglés. ((NOS))
no tuve más que aprender a hablar inglés. (102-1A1: 51)
‘At school, did they teach you English? Or how did you learn it? S: When we attend school, it was very hard to understand what other teachers said, the ones who **spoke Spanish**. D: Mhm. S: Because by then everything was spoken in English. At school, they taught everybody to speak English. So I don’t have any other option but learn English’.
- (296) K: *Pero,*
pero= mi idomia es mexicano,
y yo --
.. Yo no voy a hablar en inglés no más porque me --
.. me dice alguien que hable en inglés.
*(THROAT) más que lo **tenga que hablar** a escondidas,*
si pasa la ley <@ el gobierno @>. (318-1A3: 556)

‘But, but Mexican is my language, and I... I’m not going to speak in English just because someone tells me so. If I have to, I’ll hide to speak it [Spanish]; if the government passes such law.’

The fact that the function of *hablar* emerges from its literal meaning and not from the specific context in which it appears is related to the high degree of variability for some of its features. Thus, while for tense most of the utterances involve present (188/375: 50.1%) and imperfect utterances (108/375: 28.8%), in other features, the distribution is more disperse; such a phenomenon may inhibit the emergence of constructions other than schematic ones. For instance, although in most cases *hablar* functions as the nucleus of a simple predicate (287/375: 76.5%), it also appears as the main verb in a variety of periphrases, as the subject or complement of another verb, as an adverb, and even as the complement of a preposition. In said syntactic contexts, *decir* and other frequent verbs in the data rarely appear.

- (297) A: ... *Y por qué será eso?*
Su papá le habla en español?
[o] --
R: [Todos] *le hablan en inglés.*
todos le [2hablan en inglés2].
A: [2Pues, ahí está2].
R: *Mhm.*
A: ***Hay que hablarle .. más en [español].*** (20-5A3: 265)
‘A: And why is that, I wonder. Does his dad speak Spanish to her? or... R: Everybody speak English. Everybody speak English to her. A: Well, that’s why. A: **They have to speak** Spanish to her more often’.
- (298) O: *Cuando estaban los muchachos aquí con nosotros en la ca- --*
en la cocina,
o en el cuarto donde estuviéramos,
*y pa' que ellos supieran qué **estábamos hablando.***
(H) a veces te hablábamos inglés.
A: *Mhm.*
O: *Pero los míos casi nunca tuvimos que --*
que --
que pensar qué estábamos hablan- --
que ellos --
ellos no sabían qué estábamos hablando porque,
(H) ellos entendieron todo. (117-1A3: 752)

‘O: When the kids were with us, here in the kitchen or in the bedroom, wherever we were, so they would know what we **were talking** about, we spoke English. Sometimes we speak English. A: Mhm, O: But my kids, we almost never had to worry that they couldn’t understand what we were talking about, because they could understand everything [in Spanish]’.

(299) S: *Me siento igual de hablar inglés que español porque,
.. pues, depende con quién.
pero **me gusta más hablar** español.*

D: *Mhm.*

S: *Es más bonito.
más --*

... *El modo de expresarte es más bonito pa’ mí.* (101-1A1: 133)

‘S: I feel the same speaking English or Spanish because... Well, it depends on whom I am talking to. But **I like speaking Spanish** better. D: Mhm. S: It’s prettier. More... For me, the way you express yourself [in Spanish] is prettier.’

(300) X: *Pero la que sabía más,
.. era la Anita.
Porque ésa se crió con mi mama.*

A: *A=h.*

X: *y mi mama no le gustaba que hablaran inglés.
A mí **no me hagas hablar** a güirigüiri aquí,
ni güirigüí.*

A: *@@@@*

X: *a mí me van a hablar mexicano.
y se enojaba con la Anita.* (47-4A2: 38)

‘X: The one who knew the most was Anita, because she was raised in my mom’s house. A: Ah. X: And my mom didn’t like that they spoke English.

“**Don’t make speak** güirigüiri. A: (LAUGHS). You have to speak Spanish to me. So they would get mad at Anita’.

(301) X: *La mamá de la An- --
de la= Cindy,*

A: *Mhm.*

X: *la que estaba aquí?*

A: *Mhm.*

X: *ésa fue la que se creció un --
que= se crió con ella.*

.. *y **llegaba hablándole** en inglés,*

y la agarraba de las orejas ella,

<VOX no no no oye.

aquí me vienes a hablar n=o güirigüiri.

aquí me vas a hablar en mexicano. (77-1A2: 233)

‘X: Anita’s mom... Cindy’s mom. A: Mhm. X: The one who was here? A: Mhm. She grew up... she got raised at her house [grandma’s]. So **she would speak** English to grandma, but grandma would grab her by the ears: “No, no, and no. Listen. Here you will not speak güiriri. Here you are going to speak to me in Mexican’.

- (302) A: *qué lengua aprendió a hablar primero.*
y cuándo aprendió [a] --
 O: *[El] mexicano.*
 A: *M=hm.*
 O: *ésa fue la primer lengua que yo aprendí.*
 A: *Y cuándo aprendió a [hablar inglés]?*
 O: *[(H)] Ah=,*
pos,
*.. **aprendí a hablar** inglés --*
% % antes de entrar a la escuela.
Porque teníamos muchos ah --
muchos ahm --
ahm --
.. Neighbors.
 A: *Mhm.*
 O: *Aquí.*
ah=.
que vivían muy cerquita de nosotros. (117-1A2: 112)
 ‘A: What language did you learn to speak first, and when did you learn to...?
 O: Mexican. A: Mhm, O: That was the first language I learned. A: And when
 did you learn to speak English. O: Ah, well. I **learned to speak** English ...
 before attending school because we had a lot of, ahm, neighbors, A: Mhm. O:
 who lived very close to us.

Such a morpho-syntactic variability is not particular to the behavior of *hablar* utterances as a whole group; it is also observable when focusing on how features correlate to each other. For instance, although present tense accounts for half of the utterances with *hablar* (188/375: 50.1%), and it appears in the three most frequent subject-tense combinations for said verb –3pl-present (60/375: 16%), 3sg-present (44/375: 11.7%), and 1sg-present (36/375: 9.6%)—, none of those combinations shows strong patterns regarding subject expression, subject realization, type and realization of complement, or co-occurrence of other items. The combination 3pl-present, for instance, represents almost a third of its tense group (60/188: 31.9%), it tends to appear in utterances with simple predicate’s nucleus (47/60: 78.3%), and unexpressed subjects (22/47: 46.8%). These result suggest that *hablan* ‘they speak’ is likely to emerge as a prefab or collocation. Nevertheless, because together all expressed subjects are more frequent (25/47: 53.2%), even if slightly, than unexpressed ones, and because they are

realized by a variety of linguistic units (e.g. personal and relative pronouns, nominal phrases, etc.), (sbj) *hablan* is more variable than it seems at first sight. Furthermore, the high occurrence of expressed subjects suggests that the contrast between \emptyset -*hablan* and (sbj) *hablan* largely depends on the flow of information (i.e. new vs known subjects) and does not have implications such as evidentiality or epistemicity. This latter assumption is confirmed by the fact that, in many cases (35/47: 74.5%), the subject of *hablan* refers to specific entities; so the utterances convey indeed a speech event. Other frequent subject-tense combinations show a similar behavior, so it is unlikely that prefabs or collocations will emerge with a verbal item as their fixed element.

(303) (H) *Los que no responden en i- --*

en i- --

en es- --

en español,

que no quieren .. hablar español,

son los de Johnny.

(H) *Los de Johnny **hablan** .. muy poquito español.* (20-1A1: 578)

‘The ones who do not answer in Eng.. in Spanish, who don’t want to speak Spanish, are Johnny’s. Johnny’s kids **speak** very little Spanish.’

In terms of complement type, *hablar* utterances show low variability: most of the time (181/375: 48.3%), they are complemented by a direct object which is usually realized by a noun or nominal phrase (133/181: 73.8) referring to the name of a language. Considering that the pattern persists across different subjects and tenses, it is plausible to propose the existence of a schematic prefab *HABLAR [español]* ‘to speak Spanish’, in which the verb can appear in different tenses and with different subjects, and the slot [Spanish] can be filled by the name of any language, real or imaginary.

(304) *Cuando estaban los muchachos aquí con nosotros en la ca- --*

en la cocina,

o en el cuarto donde estuviéramos,

y pa' que ellos supieran qué estábamos hablando.

(H) *a veces te **hablábamos** inglés.* (117-1A3: 753)

‘When the kids were with us here in the kitchen, or in any other room, for them not being able to understand what we were saying, sometimes we used to speak English.’

Adjuncts are the next type of complement in terms of frequency (118/375: 31.4%); it tends to be expressed with a nominal phrase (101/118: 85.6%) which, interestingly, often refers to a language (87/101: 86.1%) but it is preceded by the preposition *en* ‘in’. The phenomenon is observed no matter what the subject or tense of *hablar* is, nor the type of predicate in which the verb appears; said variability suggests the emergence of the schematic construction *HABLAR en* [español], in which the verb varies in form, the preposition *en* ‘in’ is required and practically invariable, and [español] is the name of a language or a nominalized element referring to it.

- (305) O: *Tenían su accent and everything.*
XX y nosotros.
so -- .. *sabían que ellos .. eran gringos,*
y ***estaban hablando en mexicano.*** (117-1A2: 200)
‘O: They had their accent and everything. (Unclear) and we did too. So they knew they were gringos, and that **they were speaking in Mexican.**’

It seems, then, that there are two schematic constructions used to express similar, if not the same, situations; so, using one or the other would make no difference beyond personal preferences. For instance, in (306) below the speaker uses the noun *español* as a direct object in the two first occurrences of *HABLAR* but as an adjunct ‘*en español*’ in the third one.

- (306) S: *No más pa’ comunicar con el maestro.*
D: *Claro.*
S: *porque,*
.. *era como un pecado en ese tiempo,*
.. ***hablar español*** no?
Había --
D: *En la escuela?*
S: *Había discriminación en --*
<X por causa de que X> la gente ***habla español.***
Querían que aprendieran el inglés más no?
*Pues si **hablaban en español** en veces les pellizcaban.*

D: [Ahá=].

S: [les] *reprendían o* – (102-1A1: 65, 69 & 71)

‘S: Only to talk to the teacher. D: Right. S: Because, back then, it was like a sin to **speak Spanish**, right? D: At school? S: There was discrimination at... because people **spoke Spanish**. They want them to learn more English, no? If kids **spoke Spanish**, they were punished. D: Aha. S: They were reprehended.’

Nevertheless, it will naive to assume that a structural change does not imply semantic, cognitive, or pragmatic differences. Even if the ideas are very similar, the inclusion of *en* indicates a conceptualization of the event that does not exist in the variant without the preposition. Thus, using the name of the language as an argument rather than as an adjunct implies it is cognitively more affected by the action than when used as an adjunct; it can be said that speaking a language conveys an inherent characteristic, while speaking in a language conveys an skill, something the subject can do, but that it is not a defining feature of who he or she is.

Regarding the language’s name used in these two constructions, besides the expected *español/Spanish*, *inglés/English* and *mexicano/Mexican* items, there is a couple of instances in which the speaker uses *güirigüiri*, meaning babbling; not even a language but just unintelligible sounds. Even if scarce in the data, these instances represent an idiosyncratic use of the verb; the construction, then, constitutes an idiom similar to *hablar en chino* ‘to speak in Greek’ lit. ‘to speak in Chinese’, *hablar borucas* ‘to babble’ *hablar en castellano* or *hablar en cristiano* ‘to speak clearly’, all of them implying lack of understanding, even lack of recognition, of the interlocutor’s language and, implicitly, ideas. Interestingly, though, while most of these idioms use the *hablar en* + language pattern, *güirigüiri* appears preceded by the preposition *a* or without any preposition.

(307) X: *ésa fue la que se creció un --*
que= se crió con ella.
.. y llegaba hablándole en inglés,
y la agarraba de las orejas ella,
<VOX no no no oye.

*aquí me vienes a **hablar** n=ó güirigüiri.*
aquí me vas a hablar en mexicano.

A: @@@@

X: *no quiero no,*
el güirigüiri. (47-1A2: 236)

‘X: She was the one who grew up a... the one who was raised by her. And she [the interviewee’s daughter] would come home speaking in English and she [the interviewee’s mom] would grab her from the ears while telling her: “here you won’t babble; here you have to speak in Mexican.”’

(308) X: *Porque ésa se crió con mi mama.*

A: *A=h.*

X: *y mi mama no le gustaba que hablaran inglés.*
*A mí no me **hagas hablar** a güirigüiri aquí,*
ni güirigüí.

A: @@@@

X: *a mí me van a **hablar** mexicano.*
y se enojaba con la Anita.
(SNIFF) y la Anita quería enseñarla inglés,
y ella no. (47-4A2: 41)

‘X: because she was raised in my mom’s house. A: Ah. X: And my mom didn’t like that they spoke English. “**Don’t make speak** güirigüiri. A: (LAUGHS). You have to **speak** Mexican to me. So they would get mad at Anita. And Anita wanted to teach her English, but she would refuse.’

Another relevant aspect regarding the use of these two schemas in the interviews is its weight within each subject group. Although direct object is, overall, the most frequent type of complement for *hablar*, in the 1sg-present tense combination the adjunct *en* + language is almost as frequent, with fourteen and thirteen tokens respectively. Results such as these suggest that speakers, when talking about their linguistic activity, do not want to state firmly that they indeed speak the language (e.g. *hablar español*); so, instead they convey that they speak “in” that language. As mentioned before, the fact that the language is not a participant in the event being communicated may entail a certain cognitive distance between said language and the speaker. More data is necessary for a more comprehensive analysis of the phenomenon.

6.3. Summary

Even though the speech class has a very high token frequency, and includes the two most frequent items, *decir* and *hablar*, in the examined corpus, it produces fewer prefabs and collocations than the cognitive class. Moreover, constructions likely to develop morpho-syntactic fixation and discourse marker-related functions occur mainly with *decir*, while only schematic constructions, rooted in the lexical meaning of the verb, emerge from *hablar*. A phenomenon of such nature proves that token frequency by itself is not enough for a sequence of words to enter a grammaticalization, let alone subjectification, path: it is necessary that other factors interact in the process

Constructions emerging from utterances with *decir* include *(yo) (LE) digo* ‘I say (to HIM/HER)’, *digo* ‘I mean’, *como (te) digo* ‘as I said (to you)’, and *por eso digo* ‘that’s why I say so’. All of them constitute examples of collocations because they show some degree of structural fixation, and, more importantly, they are often used not just to convey propositional values, but also to regulate the flow of information, and to introduce speakers’ attitudes towards what is being narrated.

Also, there are constructions that have developed prefab status because they consist of elements with a certain degree of fixation and with one or more open slots to be filled by specific categories; for instance *(X) dice (que)* ‘X says (that)’, *(X) dijo* ‘X said’, *dicen* ‘they say’ and *decirle(s) X* ‘they call it/them X’. Although some of these prefabs are developing textual and discourse marker related functions, they cannot be categorized yet as collocations conveying epistemicity because the nuances they introduce in the discourse still depend largely on the lexical meaning of the verb.

For *hablar*, only two constructions were identified: *HABLAR [ESPAÑOL]* ‘to speak [Spanish]’, and *HABLAR en [ESPAÑOL]* ‘to speak in [Spanish]’. Both of them are very variable in terms of subject and tense, tend to follow the structural framework related to the lexical meaning of the verb, and convey a propositional value. Nevertheless, because their complement involves the name of a couple of languages, the construction may reduce its degree or variability, eventually producing prefabs with non-propositional meanings.

Chapter 7. Conclusions

This final chapter summarizes the most relevant results from the analysis and discusses their implications for the study of subjectification in Spanish and across languages. General patterns identified and discussed in section 4.4 suggest that, when considered as a whole, cognitive verbs seem more propitious to the expression of subjectivity, and to subjectification, because the speaker functions as the utterance's subject. Thus, by appearing as the main participant of the event, the speaker has a privileged stance to present his or her point of view. On the contrary, most utterances with speech verbs are distributed between three subjects: 1sg, 3sg, and 3pl, which suggests speakers consider other people's events as relevant as their own. In terms of tense, a similar situation is observed for both semantic classes; while cognitive verbs tend to appear in present tense utterances, speech utterances are distributed between present, preterit and imperfect tense utterances, thus suggesting a larger repertoire of communicative events in terms of who say what, and when they say it. When cross-referencing both features, differences become even more evident: while the combination 1sg-present is the most frequent for the cognitive class, 3sg-preterit is the most used in the speech class. General patterns of this sort raise two questions: a) whether patterns for the same class apply equally to all its members, particularly those showing a high token frequency; and b) how the standpoint from which speakers construct events using other semantic classes (e.g. material, relational, existential) and the expression of subjectivity affect each other. While the first question relates to the analysis of this dissertation, the second one provides avenues for future research.

Results discussed in chapters 5 and 6 show that, as useful as a general patterns may be to provide a picture of what speakers do in conversation, an accurate description, one providing more elements for the study of structural patterns and subjectivity across languages, can only be achieved through the analysis of the relation between the occurrence of certain constructions in particular contexts and the communicative intentions pursued by speakers. Table 16 summarizes the constructions identified through such local analysis, and their characteristics.

Table 16. Summary of Constructions with Cognitive and Speech Verbs

Construction	Invariability	Syntactic Independence	Prosodic Independence	Subjectivity	Type of Construction
(yo) <i>no sé</i>	++	-/+	-/+	++	prefab - collocation
(yo) <i>creo</i>	++	+	+	++	prefab - collocation
(yo) (no) <i>me acuerdo</i>	-/+	-/+	-/+	-/+	prefab
<i>quién sabe</i>	++	+	+	++	prefab - collocation
<i>X sabe</i>	-/+	-	-	-	schematic prefab
<i>SABER + INF</i>	-	-	-	+ (modality)	prefab
<i>se me hace</i>	++	-/+	-/+	+	prefab
(tú) <i>sabes</i>	+	++	++	++	prefab - collocation
(Ud.) <i>sabe</i>	+	-/+	-/+	-/+	prefab
(yo) <i>entiendo</i>	+	-/+	-	-/+	prefab
<i>¿entiendes?</i>	++	++	++	++	collocation
(yo) (le) <i>digo</i>	+	-/+	-/+	-/+	prefab - collocation
<i>por eso digo</i>	++	-/+	-/+	+ (textual)	prefab
<i>ya te digo</i>	++	++	++	++	collocation
<i>como (te) digo</i>	+	++	++	++	collocation
<i>X dijo</i>	+	-/+	-/+	+	prefab
<i>X dijo que</i>	-/+	-	-	-	schematic prefab
<i>X (le) dice que</i>	-/+	-	-	-	schematic prefab
<i>le dicen X</i>	+	-	-	+	prefab
<i>le decían X</i>	+	-	-	+	prefab
<i>dicen que</i>	+	-	-	-	prefab
<i>HABLAR {español}</i>	-	-	-	-	schema
<i>HABLAR en {español}</i>	-	-	-	-	schema

For cognitive verbs (Chapter 5), the local analysis demonstrates that the predominance of present tense and 1sg subjects in the class are largely due to the high frequency of two constructions, both of which usually convey epistemic and pragmatic

meanings: *(yo) no sé* ‘I don’t know’ and *(yo) creo* ‘I think’. The study reveals too the need of paying attention to specific subject-tense combinations for individual verbs: constructions that do not surface in a general analysis show structural invariability and lack of syntactic properties, and often constitute the strongest exemplars within their particular subject, tense or verb group. More importantly, they have entered into a subjectification: they often convey epistemic meanings and occurs as parenthetical units. In the corpus, some of said constructions include *quién sabe* ‘who knows’, *sabe* ‘Ø-knows’, *no me acuerdo* ‘I don’t remember’, *se me hace* ‘it seems to me’ and *tú sabes* ‘you know’.

A similar situation is observed for the speech class; on the one hand, the relevance of 1sg and 3sg subjects, as well as that of present and preterit tense, results from the frequent occurrence of the constructions *(yo) (LE) digo* and *(X) dijo* ‘X said’). On the other hand, less frequent constructions such as *le(s) dicen X* ‘they call it/them X’ and *dicen* ‘they say’ reveal themselves as good candidates to experience subjectification due to their structural features and the nuances they introduce in the discourse.

The analysis of the data demonstrates too that different prefabs can emerge not only from the same verb, but also from the same construction. In addition, it shows that not only morphosyntactic features, but also prosodic ones provide a set of operational measures to determine whether a particular construction constitutes a prefab, a collocation, an idiom or a schema.

The analysis also provides evidence regarding the role of established prefabs in the development of new ones. Thus, the loss of morphosyntactic variability of *digo* ‘I say’ allows the prefab to move to sentence-final, and even to internal positions, in which

it no longer functions as a mere device to report speech, but it conveys epistemic and pragmatic nuances such as assertion, opinion and modesty. The phenomenon is also observed in the emergence of *le(s) decían X* ‘they called it/them X’ and *le(s) dicen X* ‘they call it/them X’, constructions syntactically, semantically, and pragmatically related to each other.

Local analysis reveals as well that high token frequency of a lexical item is not enough to provide a fertile soil for new constructions; a combination of features is required for a form-meaning pair to start detaching itself from the propositional function imposed by the verb’s literal meaning. Evidence for this claim comes from the patterning of *HABLAR* utterances; despite its high token frequency in the corpus, the verb is predominantly used in contexts indicating what language, or in what language, people speak. In addition, the verb shows a high degree of variability in terms of subject expression and subject realization, and it barely appears in an intonation unit by itself, let alone in parenthetical positions. It is important to mention, though, that speakers use the verb in two related, yet different constructions; sometimes, they use a transitive pattern with the name of a language as the verb’s direct object (e.g. *ella habla español*), while other times they use an intransitive pattern with the name of the language as an adjunct (e.g. *ella habla en español*). Although minimal, the formal difference may implicate distinct conceptualizations of an event, with the latter one marking certain degree of distance between the subject and a language.

Finally, it is important to stress that verbs which are conventionally said to belong to the same verbal class behave very differently from one another, and that speakers use them to perform a range of non-propositional functions. These phenomena raise two

questions that require further research: 1) should these verbs still be considered members of the same verbal semantic class, and, 2) should they still be considered part of the category of “verbs” rather than as, say, pragmatic or metalinguistic constructions?

Appendix. DuBois et al (1993). List of Symbols and their Meaning

Symbol	Meaning
,	Continuing transitional continuity: intonation contour whose transitional continuity is regularly understood as continuing, in a given language.
.	Final transitional continuity: an intonation contour understood as final.
-	Truncated word: where the speaker has left the end of the projected word unuttered. "This symbol is not used to mark words that have been pronounced in an abbreviated fashion as part of an informal speech style." (Du Bois <i>et al</i> 1993: 48)
--	Truncated intonation unit: The speaker breaks off the intonation unit before completing its projected contour. Does not represent an incomplete unit as understood by normative grammar.
=	Lengthening: the preceding segment is lengthened prosodically, to a degree greater than what is expected on the basis of accent and lexical stress patterns.
?	Appeal: a class of intonation contours whose transitional continuity is regularly understood as an appeal. "Appeal" here refers to when a speaker, in producing an utterance, seeks a validating response from a listener.
%	A paralinguistic glottal stop or glottal constriction. This notation is not used in positions where glottal stop is phonologically predictable.
X	Speech that is not audible enough to allow a reasonable guess as to what was said. One X is used for each syllable of indecipherable speech. It is usually possible to make at least a rough estimate of how many syllables were uttered, even when one cannot make out what the words are.
<X X>	Uncertain Hearing: used to mark portions of the text that are not clearly audible to the transcriber, to such an extent that there is some doubt as to what words were spoken. The words so enclosed represent the transcriber's best guess as to what was said, but their accuracy is not assured.
@	Laughter (each pulse) (time)
<VOX VOX>	Its use is warranted where there is some actual shift in the quality of the stretch of quoted speech, as when the quoting speaker imitates some mannerism of the quoted speaker.
()	Vocal noises (COUGH) (THROAT) (SWALLOW) (SNIFF) (YAWN), etc.
[]	Speech overlap. The bracket is placed within the word at a point corresponding to the overlap. When a bracket is written inside a word, no space should be inserted whether the bracket is indexed or not-since any space would break up the word and cause it to appear as two separate words.
..	Brief pause: a pause of less than 0.2 second, perceived more as a break in rhythm.
...	Medium Pause: a pause of medium length-one that is noticeable but not very long, about half a second in duration (0.3 to 0.6 seconds long).
...(N)	Long Pause: relatively long pauses (.7 seconds or longer).
(H)	Audible inhalation
(Hx)	Audible exhalation

References

- AIJÓN OLIVA, MIGUEL ÁNGEL.; and MARÍA JOSÉ SERRANO. 2010. El hablante en su discurso: expresión y omisión del sujeto de *creo*. *Oralia* 13.7–38.
- ALVAR, MANUEL. 2000. *El español en el sur de Estados Unidos : estudios, encuestas, textos*. Alcalá, España: La Goleta Ediciones ; Universidad de Alcalá.
- ANDERSON, J. 1993. *Rules of the mind*. Lawrence Erlbaum Associates. Hillsdale, NJ.
- ASHBY, WILLIAM J. 1988. The syntax, pragmatics, and sociolinguistics of left-and right-dislocations in French. *Lingua* 75.203–229.
- BARLOW, MICHAEL. 2002. MonoConc Pro 2.0. *Houston, TX: Athelstan Publications*.
- BECKNER, CLAY.; and JOAN BYBEE. 2010. A Usage-Based Account of Constituency and Reanalysis. *Language Learning* 59.27–46.
- BENTIVOGLIO, PAOLA. 1992. Linguistic correlations between subjects of one-argument verbs and subjects of more-than-one-argument verbs in spoken Spanish. *Romance Languages and Modern Linguistic Theory: Selected papers from the XX Linguistic Symposium on Romance Languages, University of Ottawa, April 10 14, 1990*, 91:11. John Benjamins Publishing.
- BENTIVOGLIO, PAOLA. 1993. Full NPs in spoken Spanish: A discourse profile. *Linguistic Perspectives on Romance Languages: Selected Papers from the XXI Linguistic Symposium on Romance Languages, Santa Barbara, February 21 24, 1991*, 103:211. John Benjamins Publishing.
- BENTIVOGLIO, PAOLA.; and ELIZABETH G. WEBER. 1999. El perfil discursivo del verbo saber en el español hablado en Venezuela. *Estudios de lingüística hispánica: homenaje a María Vaquero*, 90–109. Universidad de Puerto Rico.
- BENVENISTE, EMILE. 1971. *Problems in general linguistics* 1. University of Miami Press. Coral Gables, Florida
- BERNAL-ENRÍQUEZ, YSAURA. 2000. Factores socio-históricos en la pérdida del español del suroeste de los Estados Unidos y sus implicaciones para la revitalización. *Research on Spanish in the United States: Linguistic issues and challenges*. 121–136.
- BILLS, GARLAND. 1996. Language shift, linguistic variation, and teaching Spanish to native speakers in the United States. *La enseñanza del español a hispanohablantes: Praxis y teoría*, ed. by M.C. Colombi and F.X. Alarcón, 263–282. Boston: Houghton Mifflin.
- BILLS, GARLAND D. 1997. New Mexican Spanish: demise of the earliest European variety in the United States. *American Speech* 72.154–171.

- BILLS, GARLAND D.; and NEDDY A VIGIL. 1999. Ashes to ashes: The historical basis for dialect variation in New Mexican Spanish. *Romance Philology* 53.43–68.
- BILLS, GARLAND D.; and NEDDY A. VIGIL. 2000. The continuity of change: Nahuatlismos in New Mexican Spanish. *Research on Spanish in the United States: Linguistic issues and challenges*. 137–153.
- BILLS, GARLAND D.; and NEDDY A VIGIL. 2008. *The Spanish language of New Mexico and southern Colorado: a linguistic atlas*. Albuquerque: University of New Mexico.
- BRÉAL, MICHEL. 1964. *Semantics: Studies in the science of meaning*. W. Heinemann.
- BRIZ GÓMEZ, ANTONIO. 1993. Los conectores pragmáticos en español coloquial (II): su papel metadiscursivo. *Español Actual* 59.39–56.
- BROWN, ESTHER L. 2005. Syllable-initial/s/in Traditional New Mexican Spanish: linguistic factors favoring reduction ahina. *Southwest Journal of Linguistics* 24.13–31.
- BRUNER, JEROME S.; JACQUELINE J GOODNOW.; and A GEORGE. 1956. Austin. A Study of Thinking. *New York: John Wiley & Sons, Inc* 14.330.
- BÜHLER, KARL. 1990. *Theory of language: The representational function of language*. Vol. 25. John Benjamins Publishing.
- BUSTOS GISBERT, JOSÉ M. 1996. *La construcción de textos en español*. Salamanca: Universidad de Salamanca.
- BYBEE, JOAN. 1985. *Morphology a study of the relation between meaning and form*. Amsterdam; Philadelphia, PA.: John Benjamins.
<http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=376350>.
- BYBEE, JOAN. 1998. The emergent lexicon. *CLS 34: the panels*, 421-435. Chicago: Chicago Linguistics Society
- BYBEE, JOAN. 2000. Lexicalization of sound change and alternating environments. (Ed.) Michael Broe and Janet B Pierrehumbert. *Papers in laboratory phonology V: Acquisition and the lexicon* 5.250–268.
- BYBEE, JOAN. 2001. *Phonology and Language Use*. 1st ed. Vol. 2. Cambridge UK [u.a.]: Cambridge University Press.
- BYBEE, JOAN. 2006. From Usage to Grammar: The Mind's Response to Repetition. *Language* 82.711.

- BYBEE, JOAN. 2007. *Frequency of Use And the Organization of Language*. Oxford University Press.
- BYBEE, JOAN. 2010. *Language, Usage and Cognition*. 1st ed. Cambridge University Press.
- BYBEE, JOAN.; and CLAY BECKNER. 2009. Usage-Based Theory. *The Oxford Handbook of Linguistic Analysis*, ed. by Bernd Heine and Heiko Narrog. Oxford University Press.
- BYBEE, JOAN.; and JOANNE SCHEIBMAN. 1999. The effect of usage on degrees of constituency: the reduction of don't in English. *Linguistics* 37.575.
- BYBEE, JOAN.; and RENA TORRES CACOULOS. 2009. The role of prefabs in grammaticization: how the particular and the general interact. *Distribution and Historical Change*, ed. by Roberta L. Corrigan, Edith Moravcsik, Hamid Ouali, and Kathleen Wheatley, 1:187–217. Amsterdam: John Benjamins.
- CACOULOS, RENA TORRES. 2000. *Grammaticization, synchronic variation, and language contact: A study of Spanish progressive-ndo constructions*. Vol. 52. John Benjamins Publishing.
- CHAFE, WALLACE L. 1979. The Flow of Thought and the Flow of Language in Discourse and Syntax. *Syntax and Semantics Ann Arbor, Mich.* 12.159–181.
- CHAFE, WALLACE L. 1980. The deployment of consciousness in the construction of narrative.
- CHAFE, WALLACE L. 1987. Cognitive constraints on information flow. *Coherence and grounding in discourse* 11.21–51.
- CHAFE, WALLACE L. 1993. Prosodic and Functional Units of. *Talking data: Transcription and coding in discourse research*.33.
- CHAFE, WALLACE L. 1994. *Discourse, consciousness, and time : the flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- CHAFE, WALLACE L. 1998. Language and the flow of thought. *The new psychology of language cognitive and functional approaches*, I.: Mahwah, N.J.: L. Erlbaum.
- CLARK, EVE V. 1978. Discovering what words can do. *Papers from the parasession on the lexicon*.34–57.
- CLARK, HERBERT H. 1992. *Arenas of language use*. University of Chicago Press.
- COBOS, RUBÉN. 1983. *A dictionary of New Mexico and southern Colorado Spanish*. Sante Fe: Museum of New Mexico Press.

- COMPANY COMPANY, CONCEPCIÓN. 2004. Gramaticalización por subjetivización como la prescindibilidad de la sintaxis. *Nueva Revista de Filología Hispánica* 52.1–27.
- COMPANY COMPANY, CONCEPCIÓN. 2006a. Subjectification of Verbs into Discourse Markers: Semantic-Pragmatic Change Only? *Belgian Journal of Linguistics* 20.97–121.
- COMPANY COMPANY, CONCEPCIÓN. 2006b. Zero in syntax, ten in pragmatics, or subjectification as syntactic cancellation. *Subjectification: Various paths to subjectivity*, ed. by Angeliki Athanasiadou, Costas Canakis, and Bert Cornillie, 31:375–398. Berlin: Walter de Gruyter.
- CORNILLIE, BERT. 2004. The shift from lexical to subjective readings in Spanish prometer “promise” and amenazar “threaten”. A corpus-based account. *Pragmatics* 14.1–30.
- CORNILLIE, BERT. 2005a. Agentivity and subjectivity with Spanish “prometer” and “amenazar”. A study of constructional and diatopical variation. *Revista Internacional de Lingüística Iberoamericana*. 171–196.
- CORNILLIE, BERT. 2005b. On modal grounding, reference points, and subjectification: The case of the Spanish epistemic modals. *Annual Review of Cognitive Linguistics* 3.56–77.
- CROFT, WILLIAM. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford / New York: Oxford University Press.
- CRUSE, D. A.; and WILLIAM CROFT. 2004. *Cognitive Linguistics*. Cambridge University Press.
- DAVIES, MARK. 2002. Corpus del Español: 100 million words, 1200s-1900s Davies, Mark. (2002-) Corpus del Español: 100 million words, 1200s-1900s. Database. *Corpus del Español*. <http://www.corpusdelespanol.org>.
- DU BOIS, JOHN W. 2003. Discourse and Grammar. *The new psychology of language cognitive and functional approaches to language structure*, ed. by Michael Tomasello, 2:47–87. Mahwah, N.J.: L. Erlbaum.
- DU BOIS, JOHN W.; STEPHAN SCHUETZE-COBURN.; SUSANNA CUMMING.; and DANAE PAOLINO. 1993. Outline of discourse transcription. *Talking data: Transcription and coding in discourse research*, ed. by Jane A Edwards and Martin D Lampert, 45–90. Hillsdale, NJ: Lawrence Erlbaum Associates.
- DUMONT, JENNY. 2006. Full NPs as Subjects. *9th Hispanic Linguistics Symposium*, ed. by Nuria Sagarra and Almeida Jacqueline Toribio, 286–296. Somerville, MA: Cascadilla Proceedings Project.

- EDMUNDS, PAUL. 2005. Usage of the Spanish preposition *en* in monolingual and contact varieties. *Selected Proceedings of the 7th Hispanic Linguistics Symposium*, 22–30.
- EDMUNDS, PAUL. 2006. Buen---buena gente: Repair in the Spanish of the Southwest. *Selected proceedings of the 8th Hispanic linguistics symposium*, ed. by Timothy Face and Carol Klee, 204–213. Somerville, MA: Cascadilla Proceedings Project.
- ELMAN, JEFF.; and ELIZABETH BATES. 1997. Response to letters. *Science*.
- ERMAN, BRITT.; and BEATRICE WARREN. 2000. The idiom principle and the open choice principle. *Text-Interdisciplinary Journal for the Study of Discourse* 20.29–62.
- ESPINOSA, AURELIO MACEDONIO. 1909. *Studies in New Mexican Spanish: Part I: Phonology*. University of Chicago.
- ESPINOSA, AURELIO MACEDONIO. 1911. *The spanish Language in new Mexico and southern colorado*. New Mexican Printing Company.
- FERNÁNDEZ BERNÁRDEZ, CRISTINA. 2002. *Expresiones metalingüísticas con el verbo decir*. [A Coruña]: Universidade da Coruña, Servicio de Publicacións.
- FILLMORE, CHARLES.; PAUL KAY.; and MARY CATHERINE O’CONNOR. 1988. Regularity and idiomaticity in grammatical constructions. *Language*.501–538.
- FINEGAN, EDWARD. 1995. Subjectivity and subjectivisation: an introduction. *Subjectivity and subjectivisation: Linguistic perspectives*.1–15.
- FORD, CECILIA E.; BARBARA A. FOX.; and SANDRA A. THOMPSON. 2003. Social Interaction and Grammar. *The new psychology of language. Cognitive and Functional Approaches to Language Structure*. Mahwah, N.J.; London: Lawrence Erlbaum Associates.
- FUENTES RODRÍGUEZ, CATALINA 1990. Procedimientos intradiscursivos: decir y los explicativos. *Sociolingüística andaluza*, ed. by Palet Plaja, Ma Teresa, 5:103–123.
- FUENTES RODRÍGUEZ, CATALINA.; and ESPERANZA R ALCAIDE. 1996. *La expresión de la modalidad en el habla de Sevilla*. Sevilla: Ayuntamiento de Sevilla.
- GARACHANA CAMARERO, MAR. 2008. En los límites de la gramaticalización. La evolución de encima (de que) como marcador del discurso. *Revista de filología española* 88.7–36.
- GOLDBERG, ADELE E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press.

- GÓMEZ TORREGO, LEONARDO. 1999. Los verbos auxiliares. Las perífrasis verbales de infinitivo. *Gramática Descriptiva de la Lengua Española*, ed. by Ignacio Bosque and Violeta Demonte, 2:3323–3389. Madrid: Espasa Calpe.
- GONZALES, MARIA DOLORES. 2005. Todavía decimos 'Nosotros [Los] Mexicanos': construction of identity labels among Nuevo Mexicanos. *Southwest Journal of Linguistics* 24.65–78.
- GOODWIN, CHARLES. 1979. The Interactive Construction of a Sentence in Natural Conversation. *Everyday language: Studies in ethnomethodology*. 97–121.
- GOODWIN, CHARLES. 1980. Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning. *Sociological inquiry* 50.272–302.
- GOODWIN, CHARLES. 1981. *Conversational Organization: Interaction between speakers and hearers*. Academic Press.
- HAIMAN, JOHN. 1997. Repetition and identity. *Lingua* 100.57–70.
- HALLIDAY, MICHAEL AK. 1994. Functional grammar. *London: Edward Arnold*.
- HILLS, ELIJAH CLARENCE. 1906. New-Mexican Spanish. *Publications of the Modern Language Association of America*. 706–753.
- HOPPER, PAUL J. 1987. Emergent Grammar. *Berkeley Linguistics Society* 13.139–157.
- HOPPER, PAUL J. 1998. Emergent Grammar. *The New Psychology of Language: Cognitive and Functional approaches to Language Structure*, ed. by Michael Tomasello, 1:155–176. Mahwah, N.J: L. Erlbaum.
- JENKINS, DEVIN L. 2003. Bilingual verb constructions in Southwest Spanish. *The Bilingual Review/La Revista Bilingüe* 27.195–204.
- KAY, PAUL.; and CHARLES FILLMORE. 1999. Grammatical Constructions and Linguistic Generalizations: The What's X Doing Y? Construction. *Language* 75.1–33.
- LABOV, WILLIAM. 1973. The boundaries of words and their meanings. New ways of analyzing variation in English, ed. by C.-J. Bailey and R. Shuy.
- LAKOFF, GEORGE. 1987. *Women, fire, and dangerous things*. Chicago: University of Chicago Press.
- LAKOFF, GEORGE. 1990. The Invariance Hypothesis: is abstract reason based on image-schemas? *Cognitive Linguistics (includes Cognitive Linguistic Bibliography)* 1.39–74.
- LANGACKER, RONALD W. 1987. *Foundations of cognitive grammar*. Stanford, Calif.: Stanford University Press.

- LANGACKER, RONALD W. 1990. Subjectification. *Cognitive Linguistics (includes Cognitive Linguistic Bibliography)* 1.5–38.
- LANGACKER, RONALD W. 1991. *Foundations of cognitive grammar, vol. 2: Descriptive application*. Stanford, Calif.: Stanford University Press.
- LEVELT, WILLEM JM. 1989. *Speaking: From intention to articulation*. ACL. MIT Press Series in Natural-Language Processing. MIT Press, Cambridge, Massachusetts.
- LIPSKI, JOHN M. 1994. *Latin American Spanish*. Longman Pub Group.
- LYONS, JOHN. 1982. Deixis and subjectivity: Loquor, ergo sum? *Speech, place, and acton: studies in deixis related topics*, ed. by Robert Jarvella and Wolfgang Klein, 101–124. New York: Wiley.
- LYONS, JOHN. 1994. Subjecthood and subjectivity. *Subjecthood and subjectivity: The status of the subject in linguistic theory*. 9–17.
- MALDONADO, RICARDO. 1999. *A media voz: Problemas conceptuales del clítico se*. Universidad Nacional Autónoma de México, Instituto de Investigaciones Filológicas.
- MALDONADO, RICARDO. 2007. Gramamatical Voice in Cognitive Grammar. *Geeraerts, D. & Dirk*.
- MALT, BARBARA C.; and EDWARD E. SMITH. 1984. Correlated properties in natural categories. *Journal of Verbal Learning and Verbal Behavior* 23.250–269. doi:10.1016/S0022-5371(84)90170-1.
- MATTHEWS, PETER. 1981. *Syntax*. Cambridge: Cambridge University Press.
- MEDIN, DOUGLAS L.; and MARGUERITE M. SCHAFFER. 1978. Context theory of classification learning. *Psychological review* 85.207.
- NOSOFSKY, ROBERT M. 1988. Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14.54–65. doi:10.1037/0278-7393.14.1.54.
- NUNBERG, GEOFFREY.; IVAN A SAG.; and THOMAS WASOW. 1994. Idioms. *Language* 70.491–538.
- OLBERTZ, HELLA. 2007. Dizque in Mexican Spanish: the subjectification of reportative meaning. *Rivista di linguistica* 19.151–172.
- ONO, TSUYOSHI.; and SANDRA A. THOMPSON. 1995. WHAT CAN CONVERSATION TELL US ABOUT SYNTAX?“. *Alternative linguistics: Descriptive and theoretical modes*, 213–271. Amsterdam: John Benjamins Publishing.

- PIERREHUMBERT, JANET B. 2001. Exemplar dynamics: word frequency, lenition, and contrast. *Frequency and the emergence of linguistic structure* 45.137–157.
- POSNER, MICHAEL I.; and STEVEN W. KEELE. 1968. On the genesis of abstract ideas. *Journal of experimental psychology* 77.353.
- REAL ACADEMIA ESPAÑOLA. 2014. *Diccionario de la Lengua Española*. 23a ed. Consultado en <http://www.rae.es>.
- SACKS, H. 1972. On the analyzability of stories by children. *Directions in sociolinguistics: The ethnography of communication*, ed. by J.J. Gumperz and D. Hymes, 245–325. New York: Holt, Rinehart & Winston.
- SACKS, H. 1974. An analysis of the course of a joke's telling in conversation. *Explorations in the Ethnography of Speaking*, ed. by R. Bauman and J. Sherzer, 337–353. Cambridge: Cambridge University Press.
- SACKS, H. 1987. On the preference for agreement and contiguity in sequences in conversation. *Talk and Social Organisation*, ed. by Graham Button and John R. E. Lee, 54–69. Philadelphia: Multilingual Matters.
- SACKS, HARVEY.; EMANUEL A. SCHEGLOFF.; and GAIL JEFFERSON. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language*. 696–735.
- SANZ, ISRAEL.; and DANIEL J. VILLA. 2011. The genesis of Traditional New Mexican Spanish: The emergence of a unique dialect in the Americas. *Studies in Hispanic and Lusophone Linguistics* 4.417–442.
- SCHEIBMAN, JOANNE. 2001. Local patterns of subjectivity in person and verb type in American English conversation. *Frequency and the Emergence of Linguistic Structure*, ed. by Joan L. Bybee and Paul Hopper, 61–89. John Benjamins Publishing.
- SCHEIBMAN, JOANNE. 2002. *Point of view and grammar: structural patterns of subjectivity in American English conversation*. Philadelphia, PA: John Benjamins Publishing Company.
- SCHIFFRIN, DEBORAH. 1988. *Discourse markers*. Cambridge University Press.
- SILVEIRA, AGRIPINO. 2007. Patterns of subjectivity in spoken Brazilian Portuguese. *Ensaio provindos do corpus oral do Português de Fortaleza*, ed. by C.L. Campelo, 219–264. Fortaleza, Brazil: Educe.
- SINCLAIR, JOHN. 1991. *Corpus, concordance, collocation*. Oxford University Press.
- SPAULDING, ROBERT K. 1926. *History and syntax of the progressive constructions in Spanish*. University of California Press.

- SQUARTINI, MARIO. 1998. *Verbal periphrases in Romance: aspect, actionality, and grammaticalization*. Vol. 21. Walter de Gruyter.
- TAO, H. 2001. Discovering the usual with corpora: The case of remember. *Corpus linguistics in North America*, ed. by R.C. Simpson and J.M. Swales, 116–144. Ann Arbor: University of Michigan Press.
- TAO, HONGYIN. 2003. A usage-based approach to argument structure: “Remember” and “Forget” in spoken English. *International journal of corpus linguistics* 8.75–95.
- THOMPSON, SANDRA A. 2002. “Object complements” and conversation: Towards a realistic account. *Studies in language* 26.125–163.
- THOMPSON, SANDRA A.; and ANTHONY MULAC. 1991. A quantitative perspective on the grammaticization of epistemic parentheticals in English. *Approaches to grammaticalization* 2.313–329.
- THOMPSON, SANDRA.; and PAUL HOPPER. 2001. Transitivity, clause structure, and argument structure: Evidence from conversation. *Frequency and the emergence of linguistic structure*, ed. by Joan L Bybee and Paul J Hopper. Amsterdam; [Great Britain]: J. Benjamins.
<http://public.eblib.com/EBLPublic/PublicView.do?ptiID=622579>.
- TOMASELLO, MICHAEL. 1998. *The new psychology of language cognitive and functional approaches to language structure*. Mahwah, N.J.: L. Erlbaum.
- TOMASELLO, MICHAEL. 2003a. Introduction: Some Surprises for Psychologists. *The New Psychology of Language: Cognitive and Functional Approaches To Language Structure*, ed. by Michael Tomasello, 2:1–14. Lawrence Erlbaum Associates.
- TOMASELLO, MICHAEL (ed.) 2003b. Social Interaction and Grammar. *The new psychology of language*, 2.: Mahwah, N.J.; London: Lawrence Erlbaum Associates.
- TORRES CACOULOS, RENA. 2001. From lexical to grammatical to social meaning. *Language in Society* 30.443–478.
- TORRES CACOULOS, RENA.; and JESSI ELANA AARON. 2003a. Bare English-origin nouns in Spanish: Rates, constraints, and discourse functions. *Language Variation and Change* 15.289–328.
- TORRES CACOULOS, RENA.; and JESSI ELANA AARON. 2003b. Determiner variation with English-origin nouns in New Mexican Spanish: Borrowing bare forms. *University of Pennsylvania Working Papers in Linguistics (PWPL). Selected Papers from NWAV*, 2.: 9.

- TORRES CACOULOS, RENA TORRES.; and SCOTT A. SCHWENTER. 2005. Towards an operational notion of subjectification. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, 31.:
- TORRES CACOULOS, RENA.; and CATHERINE E. TRAVIS. 2010. Variable *yo* expression in New Mexico: English influence. *Spanish of the US Southwest: A language in transition*.185–206.
- TORRES CACOULOS, RENA.; and CATHERINE E. TRAVIS. 2011. Testing convergence via code-switching: priming and the structure of variable subject expression. *International Journal of Bilingualism* 15.241–267.
- TORRES CACOULOS, RENA.; and NEDDY A. VIGIL. 2003. Sustantivos de origen inglés en discurso español:¿ cuál es su gramática? *Rio Bravo. A Journal of the Borderlands* 2.69–84.
- TORRES CACOULOS, RENA.; and NEDDY A VIGIL. 2004. Mi dad hablaba mexicano: ¿Préstamos lingüísticos o alternancias de códigos? *Actas del XIII Congreso Internacional, Asociación de Lingüística y Filología de América Latina*, ed. by Víctor M Sánchez Corrales, 455–462. San José, Costa Rica.
- TRAUGOTT, ELIZABETH C. 1989. On the rise of epistemic meanings in English: an example of subjectification in semantic change. *Language* 65.31–55.
- TRAUGOTT, ELIZABETH C. 1995. Subjectification in grammaticalization. *Subjectivity and subjectivisation*.31–54.
- TRAUGOTT, ELIZABETH CLOSS. 1997. Subjectification and the development of epistemic meaning: The case of promise and threaten. *Modality in Germanic languages: Historical and comparative perspectives*.185–210.
- TRAUGOTT, ELIZABETH CLOSS. 1999. The rhetoric of counter-expectation in semantic change: a study in subjectification Elizabeth Closs Traugott. *Historical semantics and cognition* 13.177.
- TRAUGOTT, ELIZABETH CLOSS. 2003. From subjectification to intersubjectification. *Motives for language change*, ed. by Raymond Hickey, 124–140.
- TRAVIS, CATHERINE E. 2005. El verbo decir en el español conversacional: un marcador epistémico. Paper presented at the XIV Congreso ALFAL, Monterrey, Mexico.
- TRAVIS, CATHERINE E. 2006a. Subjetivización de construcciones: los verbos “cognitivos” en el español conversacional. *Memorias del VIII Encuentro Internacional de Lingüística en el Noroeste*, ed. by R.M. Ortiz Ciscomani, 2:85–109. Hermosillo, Sonora, México: Unison. <https://researchers.anu.edu.au/publications/71404>.
- TRAVIS, CATHERINE E. 2006b. Dizque: a Colombian evidentiality strategy. *Linguistics* 44.1269–1297.

- TRAVIS, CATHERINE E. 2007. Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language variation and change* 19.101–135.
- VALDIVIA, VÍCTOR A. 2010a. Paragoge en el español de Nuevo México. Paper presented at the 9º Coloquio de Lingüística, Escuela Nacional de Antropología e Historia (ENAH). Mexico City, Mexico.
- VALDIVIA, VÍCTOR A. 2010b. Tequila, ron o dizque brandy: the development of Spanish dizque into an epistemic marker. Paper presented at the 9th High Desert Linguistics Society Conference., University of New Mexico. Albuquerque, NM.
- VALDIVIA, VÍCTOR A. 2011. Patrones de Subjetividad en el español de Nuevo México. Paper presented at the 23rd Conference on Spanish in the US and 8th Spanish in Contact with other Language, Sacramento, California.
- VALDIVIA, VÍCTOR A. 2013. “*Seguro que no había trabajo, yo creo. Yo no sé.* Verbs of cognition in spoken New Mexican Spanish.” Paper presented at the LII Conference of the Linguistic Association of the Southwest (LASSO), New Brunswick, NJ.
- VALDIVIA, VÍCTOR A. 2014. Determining prefab status in New Mexican Spanish, the case of decir. (Ed.) Benjamin Anible, Keiko Beers, Laura Hirrel, and Deborah Wager. *Proceedings of the High Desert Linguistics Society Conference* 10.143–153.
- VALDIVIA, VÍCTOR A. 2015. Conversación y subjetividad: marcadores discursivos en entrevistas a hablantes de español nuevomexicano. Paper presented at the Encuentro Académico de Egresados de la Maestría en Lingüística Aplicada, Guadalajara, Mexico.
- VERHAGEN, ARIE. 1995. Subjectification, syntax, and communication. *Subjectivity and subjectivisation: linguistic perspectives*. 103–128.
- VERHAGEN, ARIE. 2005. *Constructions of Intersubjectivity: Discourse, Syntax, and Cognition*. New York. Oxford University Press
- VIGIL, NEDDY A.; and GARLAND F. BILLS. 2004. Dialect Shift in New Mexican Spanish: A Turkey by Any Other Name... *Romance Philology* 57.323–350.
- VILLA, DANIEL. 2010. Y nos vamos p’atrás: Back to an analysis of a supposed “calque.” *Spanish of the U.S. Southwest: A language in transition*. Madrid: Iberoamericana/Vervuert.
- WEBER, ELIZABETH.; and PAOLA BENTIVOGLIO. 1991. Verbs of cognition in spoken Spanish: A discourse profile. *Discourse Pragmatics and the Verb: Evidence from Romance*, ed. by Suzanne Fleischman and Linda Waugh, 194–213. London: Routledge.

- WILSON, DAMIAN VERGARA. 2012. The intersection of identity, gender, and attitudes toward maintenance among beginning Spanish as a heritage language students. *Southwest Journal of Linguistics* 31.177–199.
- WILSON, DAMIÁN VERGARA. 2013. One construction, two source languages: Hacer with an English infinitive in bilingual discourse. *Proceedings from the 6th International Workshop on Spanish Sociolinguistics*, 123–134.
- WILSON, DAMIÁN VERGARA.; and JENNY DUMONT. 2015. The emergent grammar of bilinguals: The Spanish verb hacer “do” with a bare English infinitive. *International Journal of Bilingualism* 19.444–458.
- WILSON, DAMIAN VERGARA.; and RICARDO MARTÍNEZ. 2011. Diversity in Definition: Integrating History and Student Attitudes in Understanding Heritage Learners of Spanish in New Mexico 8. *Heritage Language Journal*.115–132.
- WRAY, ALISON. 2005. *Formulaic language and the lexicon*. Cambridge. Cambridge University Press.