

REFORMULATION AND ITS MARKERS: A STUDY OF CLASSIFICATION, SEMANTIC  
RELATION, AND FUNCTION

by

J. GRAHAM JOHNSON

(Under the Direction of Sarah E. Blackwell)

This thesis investigates three principle topics related to reformulation markers (RMs) in Spanish: the first is the classification of these markers based on their function in spoken discourse; the second is a proposed relation of entailment between the speech segment being reformulated and the reformulation; the third is the syntactic position of markers within discourse. An extensive literature review sets-up the proposed framework for RMs, based on four categories: explanation, recapitulation, rectification, and distancing. Using oral interview data taken from the Madrid corpus of PRESEEA, this framework, as well as the relationship of pragmatic entailment, are tested through qualitative discourse analysis. I conclude that most RMs are found in utterance-medial position, while further research is needed to corroborate findings based on classification and entailment relationships.

INDEX WORDS: reformulation, reformulation markers, discourse markers, classification, semantic entailment, pragmatics, discourse analysis, Spanish

REFORMULATION AND ITS MARKERS: A STUDY OF CLASSIFICATION, SEMANTIC  
RELATION, AND FUNCTION

by

J. GRAHAM JOHNSON

BA, Berry College, 2017

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment  
of the Requirements for the Degree

MASTER OF ARTS

ATHENS, GEORGIA

2019

© 2019

J. GRAHAM JOHNSON

All Rights Reserved

REFORMULATION AND ITS MARKERS: A STUDY OF CLASSIFICATION, SEMANTIC  
RELATION, AND FUNCTION

by

J. GRAHAM JOHNSON

Major Professor:	Sarah E. Blackwell
Committee:	Gary K Baker
	L. Chad Howe

Electronic Version Approved:

Suzanne Barbour  
Dean of the Graduate School  
The University of Georgia  
August 2019

## ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my committee head, Dr. Sarah Blackwell, for the encouragement and patience she provided me. I am extremely grateful her invaluable advice, constructive criticism, and helpful contributions throughout the duration of this project. I would also like to extend my sincere thanks to the other members of my committee, Drs. Chad Howe and Gary Baker, for their support and valuable feedback during this process. Lastly, but certainly not least, I would like my friends and colleagues for their unwavering support this past semester.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	iv
LIST OF TABLES.....	vii
LIST OF FIGURES.....	viii
CHAPTER	
1 INTRODUCTION.....	1
1.1 Problem Statement.....	1
1.2 Terminology and Basic Concepts.....	3
1.3 Research Questions and Outline.....	10
2 LITERATURE REVIEW.....	12
2.1 Studies on Reformulation.....	12
2.2 Paraphrastic Reformulation.....	15
2.3 Non-paraphrastic Reformulation.....	20
3 METHODOLOGY.....	27
3.1 Corpus and Data Collection.....	27
3.2 Initial PRESEEA Data Analysis.....	28
4 RESULTS AND ANALYSIS.....	31
4.1 <i>O sea (que)</i> .....	36
4.2 <i>Vamos</i> .....	39
4.3 <i>De todas formas</i> .....	42
4.4 <i>Es decir</i> .....	43

4.5 <i>Más bien</i> .....	45
4.6 <i>Quiero decir</i> .....	47
4.7 Summary and Conclusion.....	48
5 CONCLUSIONS .....	50
REFERENCES .....	54

## LIST OF TABLES

	Page
Table 1: The total number of tokens extracted from the data set .....	28
Table 2: The total number of RMs extracted based on gender.....	29
Table 3: The total number of RMs extracted based on level of education.....	30
Table 4: The total number of tokens based on syntactic position .....	31
Table 5: The total number of tokens based on reformulative function .....	33
Table 6: The number of tokens with and without an entailment relationship .....	34
Table 7: A summary of the characteristics of <i>o sea</i> .....	36
Table 8: A summary of the characteristics of <i>vamos</i> .....	40
Table 9: A summary of the characteristics of <i>de todas formas</i> .....	42
Table 10: A summary of the characteristics of <i>es decir</i> .....	44
Table 11: A summary of the characteristics of <i>más bien</i> .....	46
Table 12: A summary of the characteristics of <i>quiero decir</i> .....	47
Table 13: A summary of the characteristics analyzed.....	49

## LIST OF FIGURES

	Page
Figure 1: The classification framework for my analysis .....	25
Figure 2: The percentage of each marker in the data set .....	28

# CHAPTER 1

## INTRODUCTION

### 1.1 Problem Statement

Within the last twenty years, interest in the study of discourse markers (DMs) has increased greatly, as more linguists recognize their importance both in the general study of language, and more specifically in the study of semantics and pragmatics. According to Fraser and Malamud-Makowski (1996), these markers, sometimes also called connectives or discourse particles, communicate a relationship, not between elements within a given utterance, but between utterances in discourse. Fischer (2006) notes that, although this linguistic subfield has drawn a large degree of attention in recent years, there is very little agreement about what constitutes a DM, how these expressions are defined, and how they are classified. She observes that, due to varying methodologies, terminologies, and proposed functions, it can be difficult for researchers to “find the bits and pieces that constitute an original model of meanings and functions” (2006:1).

For the purpose of the present study, discourse markers are defined based on the definitions provided by Portolés (2001) and Crible (2017). On the one hand, Portolés explains that DMs are “unidades lingüísticas invariables, no ejercen una función sintáctica en el marco de la predicación oracional y poseen un cometido coincidente en el discurso: el de guiar, de acuerdo con sus distintas propiedades morfosintácticas, semánticas y pragmáticas, las inferencias que se realizan en la comunicación” (2001:25-26). On the other hand, Crible defines them as a “grammatically heterogeneous, multifunctional type of pragmatic markers”, which “function on

a metadiscursive level as procedural cues to situate the host unit in a co-built representation of on-going discourse” (2017:106). Based on these two definitions, I define DMs as multifunctional linguistic expressions with procedural meaning, which operate on a metadiscursive level by guiding the inferences promoted by the discourse surrounding them. The following are examples of DMs in English and Spanish, which highlight important information about their basic nature.

- (1) bueno, aunque nací en Murcia, yo: siempre me he considerado alicantina. **o sea** de Valencia nada y de Cataluña menos. (Schwenter 1996:859)
- (2) She believes that to educate and reassure people, the campaign must be “glocal”, **in other words**, it will need Europe-wide coordination but must be able to tap into the concerns of the man or woman in the street. (Murillo 2016b)

Both (1) and (2) are examples of a category of DMs known as *reformulation markers* (RMs). In example (1) the speaker uses the DM *o sea* to introduce an utterance that expands and explicitly states her feelings towards her regional identity. In (2), ‘*in other words*’ is used to introduce an explanation of the previous utterance and more specifically, the use and meaning of the word ‘*glocal*’.

Murillo (2016a) defines reformulation as

a prominent discourse phenomenon, both in oral and in written discourse, which is intrinsically related to communication processes: speakers or writers often feel the need to offer a second formulation or reinterpretation of a first discourse member in order to expand, adjust, specify, clarify, define, correct or modify different aspects. (2016a:1)

In other words, a speaker corrects or reformulates what he has already said in order to be better understood and try to better communicate his intended meaning. For this thesis, I analyze six reformulation markers in Spanish: *o sea*, *es decir*, *quiero decir*, *más bien*, *de todas formas*, and *vamos*, which were the markers found in the data taken from the Madrid corpus of the PRESEEA database. While several of these markers have received notable attention by researchers, some aspects of their use and classification still remain unclear and contested.

Reformulation can involve the act of introducing a new formulation in the discourse, due to the speaker's belief that the previous utterance was not sufficiently explanatory. For instance, in (3a) and (3b), the DMs indicate that the following segment is a reformulation.

- (3) a. Nombró a los dos; **a saber**, a Alicia y a Juan. (Portolés 2001:120)  
b. Tiene mucho dinero; **es más**, es una de las mujeres más ricas que conozco.

(3a) uses the less common RM, *a saber*, to indicate that what follows is a reformulation. More specifically, the second segment specifies the meaning of the first segment, *nombró a los dos*, with more information—the identities of *los dos*, Alicia and Juan. In (3b), the RM *es más* is used to indicate that what follows provides more explanation and context to the first segment; in addition to the hearer's knowing that the subject of the discourse has a lot of money, she now also knows that the subject referent is one of the richest women the speaker is acquainted with, which both emphasizes and adds more information to support the previous segment, seeming to justify the speaker's use of *mucho*.

The objective of this thesis is to determine which DMs can be classified as reformulation markers. Additionally, I will investigate a broadly defined relationship of entailment between discourse segments connected by RMs. While semantic entailment is a semantic relation between two propositions in which the truth of one statement or segment ensures the truth of another, the relationship proposed in this thesis is not semantic, but pragmatic. The differences between these two kinds of discourse relation are discussed in Chapter 3. This entailment relationship is seen in example (3b) above, where the second segment, *es una de las mujeres más ricas que conozco*, entails the sense of the first segment, *tiene mucho dinero*. In other words, based on the utterance in (3b), we understand that being “one of the richest women he knows” entails that “she has a lot of money”.

The next section defines terms and basic concepts of the study, followed by a justification for the study, my research questions, and an outline for the rest of the thesis.

## 1.2. Terminology and Basic Concepts

Firstly, I will discuss features of DMs, and subsequently, I will focus on defining reformulation markers in greater detail. For this study, I adopt the term *discourse marker* (as opposed to *discourse particle*, *pragmatic marker*, *discourse connective*), which was defined earlier in this chapter. I also adopt Fraser's (1999) notion of *discourse segments*. According to Fraser, DMs "impose a relationship between some aspect of the discourse segment they are a part of, call it S2, and some aspect of a prior discourse segment, call it S1" (1999:938). For this study, Fraser's notion of discourse segment will be used to refer to the portion of the speaker's utterance preceding an RM (S1), which the speaker reformulates in some way in the discourse sequence following the RM.

### 1.2.1 Characteristics of DMs

The earliest mentions of DMs as objects of linguistic study can be seen in the works of Labov and Fanshel (1977), Levinson (1983), and Zwicky (1985) (all cited in Fraser 1999), although Labov and Fanshel and Levinson do not enter into any kind of in-depth discussion on the subject. Zwicky is perhaps the first to use the label 'discourse markers', writing:

Within the great collection of things that have been labeled 'particles', we find at least one grammatically significant class of items, in English and in languages generally. These have been variously termed 'discourse particles' and 'interjections'; here I will call them 'discourse markers' ... On the grounds of distribution, prosody, and meaning, discourse markers can be seen to form a class. But like the 'particles' discussed ... they are independent words rather than clitics... (Zwicky 1985:303, cited in Fraser 1999:933)

Zwicky notes several characteristics of DMs, such as their prosodic independence from surrounding discourse and their typically utterance-initial placement, and acknowledges that they should not be categorized together with other function words—although he does not propose a separate class distinction as do other authors. However, one of the most thorough and

comprehensive early efforts to study DMs was carried out by Schiffrin (1987). Her research findings align with Zwicky, although she uses a different framework—that of coherence-based planes. She posits that DMs must appear utterance-initially. Fraser (1990) also finds evidence of DMs in utterance-initial position, as seen in (4a) and (4b).

- (4) a. A: Then we stopped for a while. (Fraser 1990:389)  
B: **And** then what happened?  
b. I tried to get there. I failed. **And** then I tried to call.  
c. John ran away from home **and** Mary did the same later on.

Using (4) as an example, Fraser does, however, acknowledge the difficulty of determining what constitutes an utterance, and thus determining what is considered to be ‘utterance-initial’, as well as admitting the occasional presence of markers in utterance-medial and final position, as seen in (4c). However, not all researchers share the purely-sentence-initial-view.

In Schwenter’s (1996) work on *o sea*, the relative prosodic independence of DMs, referred to by Zwickey (1985), is corroborated, although he suggests that, with respect to *o sea* at least, DMs are relatively flexible when it comes to their syntactic position—be it initial, medial, or final. Crible (2017) and Aijmer and Simon-Vandenberg (2011) agree that DMs can occur in all of these positions in discourse. For description of syntactic position, I use the terms *utterance* and *turn* interchangeably (i.e., turn-initial position vs. utterance-initial position). Turns occur in conversational discourse and are marked by a change of speaker; utterances can occur in the same turn, as in (4b) and (c) above. For that reason, utterance-initial, unless otherwise specified, refers to turn-initial discourse in dialogical conversation.

The discussion of DM ‘core meaning’ is also prevalent in the literature. Schourup (1985) and Schiffrin (1987) propose a specific core content meaning for each DM, which survives despite the context in which they find themselves. Fraser (1996) disagrees with this notion, adhering to the idea that DMs do not create meaning in an utterance, but instead exhibit *procedural* meaning. Specifically, Fraser maintains that, while sentence propositions “have representational meaning,

in virtue of which they denote concepts, the discourse markers have procedural meaning and specify how the sentence of which they are a part is related to the preceding discourse” (1996:170). This idea is postulated as well by Blakemore (1987, cited in Fraser 1996).

### 1.2.2 Defining Reformulation

Various researchers have investigated reformulation in discourse, both in English (e.g., Cuenca 2003; Cuenca & Bach 2007; Milgros del Saz Rubio 2007; Murillo 2004, 2016b) and in Spanish (e.g., Cuenca 2003; Cuenca & Bach 2007; Garcés Gómez 2003, 2010; Murillo 2016a; Pons 2013, 2017; Portolés 2001; Núñez & González 2016). Portolés (2001:109) describes the action of formulation as “una actividad intencional que requiere un esfuerzo, ya que es preciso resolver diversas dificultades”. Furthermore, he explains that it can be difficult to achieve optimal communication, which in turn results in a new formulation. He provides the following as an example of reformulation.

(5) Tiene muy buena voluntad, **o sea**, es muy trabajadora. (Portolés 2001:141)

The speaker in (5) reformulates S1, *tiene muy buena voluntad*, with the explanation in S2 that the subject is a very hard-worker.

Cuenca (2003) defines reformulation as “a discourse function by which the speaker re-elaborates an idea in order to be more specific and facilitate the hearer’s understanding of the original or in order to extend the information previously given” (2003:1071). Additionally, Cuenca and Bach (2007) note that reformulation ensures discursive cohesion and facilitates discursive progress by reducing possible defects in discourse. In other words, if a speaker does not believe her first utterance to be sufficiently informative, she makes an effort to correct or change it. Cuenca (2003) uses the following example to demonstrate reformulation.

(6) En esta página observamos la presencia de 62 unidades verbales: 51 se hallan en nuestro centenar y otros once verbos no, pues son específicos de la conversación

grabada. **O sea**, un 82% de los verbos que se emplean en esa página corresponde a nuestra lista seleccionada (Cuenca 2003:1071)

In (6), the speaker restates the information in S1, that 51/62 verbs are found in the speaker's list, with a percentage (82%). She provides the same information in a different, but semantically equivalent form.

Murillo (2004), working within the framework of Relevance Theory (Sperber & Wilson 1995), explains that “reformulations are representations of utterances which they resemble” (2004:2061), “where the resemblance involves the sharing of logical and contextual implications” (Blakemore 1996:338, cited in Murillo 2004). She provides the following example of reformulation in English.

(7) Carver chairs—**that is**, any style with arms—are much more comfortable, but they do take up a lot more room. (Murillo 2004:2062)

In (7), the speaker introduces an explanation of S1, *carver chairs*, with the English marker *that is*, and then explains that carver chairs are “any style with arms.”

While other researchers have also provided definitions of reformulation, in this thesis I adopt Cuenca's (2003) definition as the basis for my classification framework (see Chapter 2) and analysis (see Chapter 4).

### 1.3 Research Questions

As mentioned above, there are strikingly low levels of consensus regarding the classification and features of DMs and the definition of reformulation. Some researchers argue that a definition of reformulation similar to the one put forth by Murillo (2016b) (see above) is not sufficiently descriptive, or that it is too broad a category. One of the more notable and more recent debates in this field is that between Pons (2013, 2017) and Murillo (2016a). Pons proposes a single type of reformulation, while Murillo defends the position adopted by the majority of linguists, that of

two different types of reformulation. Through an extensive review of the literature, I aim to provide a more concise classification framework of reformulation by identifying the criteria that have been used previously to identify and classify RMs, and to investigate the questions raised by other investigators. Therefore, my first set of research questions is:

1. How can reformulation sub functions best be defined, and what Spanish RMs have been classified in which sub-function(s)? Can certain RMs introduce more than one sub-function?

The literature to date has not clearly explained the relationship between discourse segments or utterances that RMs connect. Most studies focus on identifying and describing reformulation functions; in this study, however, in addition to investigating functions and use, we propose an entailment relationship between discourse segments connected by a RM. Example (4) illustrates this proposed relationship in English.

- (8) She's a child, **I mean** she's just three years old.  
Three years old | - child

The segment “three years old” entails “child”, while the reverse does not hold true (i.e., “child” does not entail “three years old”). The reformulation marker *I mean* acts to direct the hearer to a more specifically rephrased second segment. I hope that an investigation of data from a corpus of spoken Peninsular Spanish (see Chapter 3) will help to answer my second set of questions:

2. How and to what extent does Segment 2 entail Segment 1? How and to what extent are these two segments different?

Early studies on DMs show that the majority of DMs occur syntactically in utterance-initial position. Researchers, such as Aijmer and Simon-Vandenberg (2011) and Crible (2017), find that DMs can appear flexibly in any position. However, to my knowledge, no work has been

done solely on the syntactic position of RMs in discourse. It follows logically that RMs would most likely occur in utterance-medial position, since speakers tend to reformulate their own discourse, as seen in (9).

- (9) El polo semántico es la significación contextual de una expresión, **es decir**, la conceptualización detallada que constituye nuestra comprensión total de la expresión en contexto (Cuenca 2003:1077)

The speaker in (9) describes the *polo semántico* (semantic pole) as the contextual meaning of an expression. She then uses *es decir* to introduce an explanation of the previous description.

Therefore, my third research question is:

3. In what syntactic position are RMs most commonly found? Are they as likely to appear in turn-initial and turn-final position as they are in turn-medial position?

#### **1.4 Outline of the Thesis**

This thesis is organized as follows. Chapter 2 is a review of the literature on reformulation and RMs. In that chapter, I aim to answer the first research question about the classification and identification of RMs. Chapter 3 presents my methodology. To answer my second and third questions, I analyze data from the Madrid corpus of the PRESEEA database (2014). These data consist of interviews with native speakers of Peninsular Madrilenian Spanish. The fourth chapter presents an analysis of the data and discussion of the results. Chapter five concludes the thesis and offers some closing thoughts, as well as ideas for future study.

CHAPTER 2  
LITERATURE REVIEW

### 2.1 Introduction

Several studies have been done on the use of reformulation, both in Spanish and in English. As previously discussed, little consensus exists among researchers regarding classification and identification of markers and how to define the notion of reformulation. In this chapter, I review and evaluate various studies related to reformulation, highlighting proposed definitions of this discourse function. This chapter is organized around the commonly accepted division between paraphrastic reformulation, and non-paraphrastic reformulation, with subsections based on sub-functions of paraphrastic reformulation and non-paraphrastic reformulation. The first section deals with paraphrastic reformulation, which conveys the semantic content already expressed in the previous utterance, albeit with a certain degree of change; it can involve a restatement or clarification of content, as in (10):

- (10) Pero son claras dos cosas: por una parte, que los niveles de lengua se interrelacionan en el uso de una determinada lengua; **es decir**, que el análisis en niveles es únicamente metodológico (Cuenca 2003:1072)

The second section examines functions of non-paraphrastic reformulation, which indicates a dissociation of semantic content between discourse segments. It conveys a change of perspective through distancing from or reconsideration of propositional content, as in (11):

- (11) usted ha dirigido también programas en televisión. ¿Se aprende mucho de un medio hacia otro, **quiero decir**, se aprende dirigiendo un programa en televisión para luego verterlo al cine? (Garcés Gómez 2011:94)

Returning to Portolés's definition of reformulation, "una actividad intencional que requiere un esfuerzo, ya que es preciso resolver diversas dificultades" (2001:109), it is understood that optimal communication is difficult to achieve, creating the need for reformulation. Garcés Gómez (2003, 2010) offers yet another definition of reformulation. She defines it as "un proceso de vuelta sobre un miembro anterior del discurso, explícito o implícito, para formularlo de una manera distinta" (2003:114) and points out the ability of speakers to choose the formulations and linguistic contents that seem most appropriate in a given context (2010:88). Consequently, we can surmise that when a speaker feels she has been insufficiently explanatory, she can reformulate previous discourse to make herself more informative by, for instance, correcting or rewording her previous utterance.

One view of reformulation is that it is "una relación de equivalencia semántica gradual entre dos segmentos [S1] y [S2], o bien en una predicación de identidad vehiculada por marcadores cuando esta relación de equivalencia es débil o no es interpretable a partir de los significados proposicionales" (Polanco Martínez 2016:17). There are several crucial parts to this definition, which enable us to better understand the nature of reformulation, the first being that of a *gradual* relationship between segments. We can understand Polanco Martínez's idea of gradation as depending of the level of semantic equivalence existing between S1 and S2, as well as on the amount of required inference on the part of the hearer, both of which change depending on speaker intentions and hearer understanding. Polanco Martínez adds that "cuanto mayor sea la semejanza semántica entre ambas expresiones, menor será la necesidad de señalar una relación de equivalencia; por el contrario, cuanto mayor sea el peso inferencial en el establecimiento de la equivalencia, mayor será la necesidad de explicitar dicha relación" (2016:17). This second part of his definition is crucial in that it is based on the *semantic* relationship that exists between related segments and can be analyzed based on the significant common elements between the

two utterances. Cuenca and Bach (2007) maintain that reformulation involves a pragmatic relationship, not a semantic one. They explain that the propositional contents of each segment are not semantically equivalent, given that they do not contain the same exact propositional content, but a reformulated version. Their data are taken from the written versions of oral presentations of linguistics papers in English, Spanish, and Catalan. An example of this written discourse is seen in (12):

- (12) Al considerarlos como cosas diferentes, la primera consecuencia es que se «contabilizan» aparte, haciendo sospechosos los porcentajes (con independencia de que ya lo sean por sí mismos) y sin hacerse jamás las tres preguntas científicas obligatorias que proceden; **es decir**, a) ¿son dos cosas distintas?; b) ¿en qué consiste su diferencia semántica, si la hay?; c) ¿cuáles pueden ser las causas de la mayor abundancia del uso de ‘probabilidad’?  
(Cuenca & Bach 2007:158)

The speaker in (12) reformulates S1, *las tres preguntas*, by specifying what the three questions are, expressed in S2. We can clearly see that the content of S1 is not the same as the content of S2.

Various researchers have proposed models for the reformulation relationship between two segments. These models provide a very basic understanding of how reformulation works. Their usefulness comes from being able to talk in general terms about the reformulative function.

Cuenca and Bach (2007) propose a model based on equivalence operation.

- (13) Formulation A = Formulation A' (Cuenca & Bach 2007:152)

Their model indicates that each formulation is meant to be equivalent in some way (A) but presented in different forms (A'). They offer the following example to illustrate the model.

- (14) I briefly describe below some of the evidence that is thought to favor the autonomy of language, **or** modularity, view because its existence is seen by some linguists as removing the need for any subfield of linguistics called cognitive linguistics.

(Cuenca & Bach 2007:152)

Segment A, *the autonomy of language*, is reformulated in segment A' as *modularity*. These expressions are equivalent in that they can be used to define one another. A and A' are connected by the English RM *or*.

The reformulation model proposed by Polanco Martínez is slightly more nuanced than the model in (13).

(15) [A REFORMULATION A'(A $\cong$ A')] (Polanco Martínez 2016:20)

As in (13) above, segment A represents the original proposition, which undergoes reformulation to produce segment A'. The difference in this model is the use of the ' $\cong$ ' symbol. It denotes the gradient relationship between A and A', which can be seen in the different sub-functions of reformulation. An exact paraphrase is more semantically equivalent than some type correction, which is more equivalent than a distancing function. Polanco Martínez uses example (16) to illustrate his model.

(16) He hablado con él y dice que tiene que acabar un trabajo para mañana y que va por la mitad. **O sea/vamos** que no viene (Polanco Martínez 2016:22)

In (16) the speaker gives reasons why his friend cannot come to a party (A) but reduces his explanation to a simpler reason—that his friend is simply not coming (A'). Polanco Martínez (2016) notes that the use of different RMs is possible in this context, but that using different RMs results in a different interpretation of the utterance; thus, the gradient aspect of his model.

In the following sections I divide the discussion of reformulation into sections based on the paraphrastic/non-paraphrastic division supported by several researchers (e.g., Cuenca 2003; Cuenca & Bach 2007; Garcés Gómez 2003, 2010; Castillo Fadić & Insua 2017). The distinction between paraphrastic and non-paraphrastic reformulation is based on two factors, perspective shift and semantic/pragmatic equivalence or the lack thereof. In paraphrastic reformulations, the speaker's perspective generally remains the same in both S1 and S2, leading to semantic equivalence between the two segments. In non-paraphrastic reformulation, the speaker's

perspective shifts between S1 and S2, creating a difference in propositional content and thus eliminating complete semantic equivalence in both segments. I define these two notions more fully in the following two sections, each of which is further divided into subsections based on specific types of reformulation.

## 2.2 Paraphrastic Reformulation

Polanco Martínez (2016) defines paraphrase as process of reformulation in which S2 is reformulated “para extraer el elemento o los elementos esenciales de [S1] (recapitular o resumir) o para explicitar una conclusión derivada de [S1]” (2016:25). In other words, paraphrastic reformulation involves summary, recapitulation and/or explanation. The key aspect of paraphrase is that S2 contains the most important semantic pieces of S1, stated in a way the speaker believes to be more concise and communicatively effective. It “establishes a semantic or pragmatic relationship between two segments” (Garcés Gómez 2003:115, my translation). According to Bataller (2002), it is the simple restatement in some way of content expressed in S1 that defines paraphrastic reformulation. Polanco Martínez (2016) illustrates this with examples such as (17):

- (17) Hay que informar, pero de forma gratis, **vamos**, que no les cueste  
a usted dinerillo al tener que ver esta revista, como en otras ocasiones.  
(Polanco Martínez 2016:23)

The speaker in (17) uses the marker *vamos* to indicate that S2, *que no les cueste a usted dinerillo*, is a paraphrase of S1, *de forma gratis*. The propositional content—that is, the essential information—of both S1 and S2 is that the hearer will not have to pay any money for the service(s) being discussed.

In an earlier study on reformulation in French, Gülich and Kotschi (1995), like Polanco Martínez (2016), view paraphrase as conveying something already expressed in previous utterance, albeit with a certain degree of change (Gülich & Kotschi 1995:45, 47). They divide

paraphrastic reformulation into three subclasses: expansion, reduction, and variation. Their classification contrasts with that of Portolés (2001), who considers *explanation* (“explicación”) to be a separate reformulation subset, while Gülich and Kotschi (1995) consider it to fall under the subset of *expansion*. Classification of the function *recapitulation*, also proposed by Portolés (2001), has been contested as well. Garcés Gómez (2003), among others, finds *recapitulation* to be a function of non-paraphrastic reformulation. However, I consider it a function of paraphrastic reformulation. I discuss this contention more in section 2.2.2. Additionally in her classification, Garcés Gómez (2010, 2011) includes both correction and rectification as processes of paraphrastic reformulation. I, however, classify them as functions of non-paraphrastic reformulation. These two functions will be defined and discussed in section 2.3.

In the following sections I examine the categories of explanation and recapitulation individually, both of which I consider to be sub-functions of reformulation.

### 2.2.1 Explanation

Portolés defines *explicación* (‘explanation’) as a process of clarifying or explaining what the speaker said in the previous discourse segment (2001:142). The following examples illustrate Portolés’s notion of explanation:

(18) Tiene muy buena voluntad, **o sea**, es muy trabajadora (Portolés 2001:141)

(19) León es puro llano. **O sea** que puedes extenderte perfectamente (Schwenter 1996:863)

The speaker in (18), repeated from (5) in Chapter 1, clarifies the meaning of S1, *tiene muy buena voluntad*, with S2, *es muy trabajadora*, and connects the two segments with the RM *o sea*. He explains that the subject being a hard worker is related to her tenacity or persistence. In (19), the speaker explains that due to León’s flatness, one can travel very easily there.

Fernández (2000) examines the marker *quiero decir* and finds that it is used with the explanatory function, as well as other functions, to avoid misunderstanding and incomprehension and to repair them when they occur. She provides the following example to illustrate *quiero decir* with explanatory function.

- (20) La visita del Met, con el extraordinario talento de James Levine, Placido Domingo, Aprile Mollo, Juan Pons, Piero Faggioni y cantos otros, nos ha hecho a todos algo más ricos. **Quiero decir** espiritualmente. Aunque algunos, no se sabe por que, también se harán ricos de la otra manera (Fernández 2000:277)

In (20), the speaker describes the talent of great singers, noting how they have greatly “enriched” listeners. She then explains that she means spiritual enrichment, although others probably are enriched in the monetary sense as well.

According Casado (1991), explanation can have several functions, such as “evidenciación...de la expresión, del contenido, de la presuposición, de la inferencia..., o de la implicación” (1991:108). Basically, the explanation function clarifies explicit propositional content as well as implicit content, i.e., presupposition, inference, or implication, as seen in (21).

- (21) Un dos por ciento de la población filipina, **es decir**, un millón de personas, tiene como lengua materna el español (Casado 1991:108)

The speaker in (21) describes the percentage of Philippine people who have Spanish as their native language. He states that they account for 2% of the population, but then gives a further explanation, introduced by *es decir*, that 2% is about one million people.

While many researchers have studied markers that indicate that the following segment has an explanatory function, most focus on reformulation as an overall function rather than on explanation more specifically. However, I have chosen to include explanation as a subsection of reformulation in my classification because it is a clearly identifiable sub-function of reformulation. In the next section, I look at the function of recapitulation. 1

### 2.2.2 Recapitulation

I define the recapitulation function as a restatement of prior propositional content in a slightly different way. Garcés Gómez (2003) explains that recapitulation “consiste en volver sobre lo anterior para condensar en un último miembro lo que se considera más relevante” (2003:115). In other words, recapitulation involves... There is necessarily some degree of equivalence between the propositions of S1 and S2; that is, the semantic content in each segment must be equivalent to some degree.

Although Garcés Gómez (2003) and Gülich and Kotschi (1995) classify recapitulation as a type of non-paraphrastic reformulation, I have chosen to consider it as an operation of paraphrastic reformulation, as proposed by Polanco Martínez (2016). Portolés’s (2001) definition is somewhat ambiguous. He explains that the “miembro recapitulador puede mantener la misma orientación argumentativa de los miembros anteriores...pero también, además de la misma orientación, [puede] introducir miembros con orientación opuesta” (2001:143). In other words, Portolés asserts that a restated segment can either maintain the same discourse orientation (paraphrastic reformulation), or it can signal a change in said orientation, thus signaling a change in speaker perspective (non-paraphrastic reformulation). Thus, for Portolés, recapitulation can feasibly involve either a paraphrastic or a non-paraphrastic function. Based on Garcés Gómez’s (2003) definition of recapitulation above, this function actually appears to have a paraphrastic function. Nevertheless, Garcés Gómez maintains that recapitulation constitutes non-paraphrastic reformulation, which presents a certain grade of distance between both utterances, given that S2 would only constitute a partial paraphrase of S1. However, since I have defined paraphrastic reformulation as necessarily containing some degree of the same semantic content in both S1 and S2, for the purposes of the present study, I hypothesize that recapitulation is a type of

paraphrastic reformulation. The following examples of recapitulation from Garcés Gómez (2003:125) illustrate this function.

- (22) Abajo tenía dispuesto su almacén de telas moriscas y persas, objetos de culto, collares, diademas, brazaletes y otras joyas de princesas; **en fin**, toda suerte de tesoros traídos del sur y aun del oriente remoto para vender a los nobles leoneses
- (23) Es preciso que esa joven haga antesala. Es preciso que comprenda la autoridad, el orden, el poder...**en fin**, todo eso.

The speaker in (22) lists in S1 a number of expensive items to be sold to Leonese nobles. In S2, she summarizes said items into one group: all kinds of treasures brought from the south. In (23) the speaker again lists a number of things he feels that the subject should learn; he then refers to all of them with *todo eso*. In both examples the RM *en fin* signals to the speaker that what follows is a recapitulation or summing up of the contents expressed in S1.

The RMs most commonly classified as recapitulation markers (Portolés 2001; Garcés Gómez 2003) are *en suma*, *en resumen*, *en síntesis*, *en conclusión*, and *en fin*. It is interesting to note that these markers are more structurally complex than the grammaticalized markers of explanation. One might expect them to be more commonly found in written texts, rather than spoken discourse, due to writers' needing to summarize paragraphs, sections, chapters, articles and the like.

### 2.3 Non-paraphrastic Reformulation

Pons-Bordería (2001) defines non-paraphrastic reformulation as “a corrective move which displays a change in the speaker’s perspective, resulting in a reconsideration, distancing or even rejection of the perspective displayed in the preceding act” (Pons-Bordería 2001:224, cited in Bataller 2002). Similarly, Gülich and Kotschi (1995) propose that non-paraphrastic reformulation indicates a “change in utterance perspective, which at the same time indicates a certain degree of distance to the perspective contained in the reference expression” (1995:44).

The key aspect of non-paraphrastic reformulation is therefore the change of speaker perspective, signaled by the use of a RM. I would also argue that this type of reformulation entails much weaker semantic equivalence between discourse segments. This notion is supported by Gülich and Kotschi's (1995) idea of "invalidation," whereby "the validity of the reference expression is partly or completely 'annulled'" (1995:48).

Gülich and Kotschi (1995) divide non-paraphrastic reformulation into two categories: correction and dissociation. They then further divide dissociation into three subgroups: recapitulation, reconsideration, and detachment. Their classification is based on the strength of dissociation of S2 with S1—that is, the notion of invalidation. According to these authors, there is a gradient relationship between recapitulations (the weakest dissociation), reconsiderations (medium dissociations), and detachments (the strongest type of dissociation) (1995:47-8). They provide an example of dissociation in French indicated by the marker *après tous*:

- (24) C: ...et aujourd'hui comment considérez-vous c(e) moment-là: et comment pourriez-vous être marxiste ou avoir=encore en vous una PEtite graine de Marx, ...  
 L: mais . la GRAINE de Marx que j'ai gardée très certainement c'est euh c'est certaines pensées d(e) phiosophie de Marx= **après tous**: euh Marx a été . un . GRAND spécialiste de c'que nous app(e)lons les sciences sociales et humaines ... (Gülich and Kotschi 1995:48)

*C: ... and today, how do you consider that moment and how could you be a Marxist or still retain a small grain of Marxism ...?*

*L: but the small grain of Marx that I have most certainly kept is, uhh, is certain philosophical ideas of his. **After all**, uhh Marx was a renowned specialist of what we consider social and human sciences...*

(my translation)

According to their analysis, the expression, preceding the marker *après tous* (Eng: *after all*; Spa: *después de todo*), which is being reformulated after the marker is not a current part of the discourse, but a piece of mutual knowledge about Marxism and counterarguments against it, shared by both speaker and hearer. The reformulation, then, is the new, more appropriate viewpoint of Marx and his ideology (1995:48-9).

While I agree with Gülich and Kotschi (1995) on their classification of correction as a function of non-paraphrastic reformulation, I do not fully share their view on recapitulation, having already determined it to be a function of paraphrastic reformulation. Instead, I separate non-paraphrastic reformulation into two subcategories: rectification and distancing. Each of these two functions is treated in the following sections.

### 2.3.1 Rectification

Garcés Gómez (2010, 2011) considers rectification and correction to be two functions of paraphrastic reformulation and recognizes that they are distinct processes. For her, the main differences between them lie in the level of “invalidation” expressed by S2. However, Portolés (2001) and Gülich and Kotschi (1995), do not consider the two functions to be separate. Garcés Gómez (2010) defines rectification as the process in which

se anula parcial o totalmente lo expresado en el segmento de referencia porque no responde al estado de cosas que se pretende comunicar y se sustituye por lo expresado en el segmento reformulado; de este modo, la rectificación implica una anulación de lo dicho previamente y se realiza sobre la información conceptual transmitida o sobre los procesos inferenciales que se derivan de lo expresado. (2010:102)

The key aspect of rectification is partial or total annulment of the content expressed in S1. As one would expect, this annulment creates an extremely weak, if not non-existent, semantic relation between S1 and S2. Garcés Gómez specifies the following examples of rectification:

- (25) Después hablaré con más detenimiento de la sororidad como nueva forma de relación entre mujeres. Aunque tengamos experiencias dispares, y vivamos conflictos entre nosotras mismas, algunos de los cuales ya he analizado, a pesar de ellos, o **mejor dicho**, gracias a ellos, podemos sumar alternativas y procedimientos, elaborar formas de vida, construir nuevos deseos que integren la diversidad. (Garcés Gómez 2010:95-6)
  
- (26) Perdona, Ifigenia, pero tu padre ha muerto, lo quieras o no. Ya no se puede hacer nada. Es un asunto sin solución, o **mejor dicho**, solucionado por completo. La muerte es la solución drástica a los problemas de la vida. (Garcés Gómez 2010:96)

In (25), the speaker uses the RM *mejor dicho* to signal that what follows is a rectification, whereby the speaker changes the phrase *a pesar de ellos* into *gracias a ellos*. She begins to say that the difficulties and conflicts women face bring them together in spite of hardship; she corrects herself, however, letting the hearer know that these conflicts have actually been a good thing, having brought them together even more. In (26), the speaker changes his use of *un asunto sin solución* to *solucionado por completo*. This constitutes a full rectification of S1; that is, the speaker states that nothing can be done about the “problem”, referring to the death of the subject’s father, but then changes his view to express the fact that death is actually the most final solution to any problem.

The rectification function also includes corrections, which Garcés Gómez (2010) explains modify or correct the “mistakes” or poor word choices made by speakers in prior discourse. In other words, corrections act as repairs in discourse. Levinson (1983) describes repairs as being a broad category, including “word recovery problems, self-editings where no discernable ‘error’ occurred, corrections proper (i.e., error replacements) and much else besides” (1983:341).

Example (27) illustrates the sub-function of a correction.

- (27) cuando quería entrar en la oposición para...controlador, pues... tenía que ser... o sea, le pedían 3... **o sea**, <vacilación> 3 dioptrías se admitían como máximo, ¿no?  
 (Garcés Gomez 2011:98)

According to Garcés Gomez’s analysis, the speaker in (27) unwittingly uses the verb *pedir*, after which she uses *o sea* to indicate her erroneous use of *pedían* in S1, and that the segment that follows, with *admitían*, is more accurate.

Garcés Gómez (2011) and Polanco Martínez (2016) cite *mejor dicho*, *quiero decir*, *vamos*, *bueno*, *es decir*, and *o sea* as the most frequently occurring RMs which introduce the rectification function, while Portolés (2001) considers only *mejor dicho* and *más bien* to have this function. Interestingly, markers such as *bueno* and *vamos* are not always classified as RMs.

For instance, Portolés (2001) classifies *bueno* as a discourse operator with formulative function and does not consider *vamos* as pertaining to any of his RM sub-classifications. The polysemous nature of these RMs is discussed further in Chapter 4.

### 2.3.2 Distancing

The final subcategory of non-paraphrastic reformulation is distancing, or *detachment*. Garcés Gómez (2003), Gülich and Kotschi (1995), Martín Núñez (2013), Polanco Martínez (2016), and Portolés (2001) include distancing as a type of reformulation. Portolés (2001) defines distancing as presenting

expresamente como no relevante para la prosecución del discurso un miembro del discurso anterior a aquel que los acoge. Con ellos no se pretende formular de nuevo lo antes dicho, sino mostrar la nueva formulación como aquella que ha de condicionar la prosecución del discurso, al tiempo que se priva de pertinencia el miembro discursivo que lo procede. (2001:142-3)

The key aspect of distancing is that it signals the non-relevance of what is referred to in S1 with respect to S2; the speaker discursively ‘distances’ herself from what she has said in the previous discourse segment. The goal is not to restate or correct what was already said, but to show that following utterances are more relevant to the speaker’s communicative intent. For Garcés Gómez (2003), the process of distancing deals with changing the direction of the discourse (2003:115).

Martín Núñez (2013) offers what he considers to be an example of distancing in (28).

- (28) I: (...) últimamente como yo estoy media patuleca/ ee// y/ y ee/ no ando en/ no ando en/ en locomoción colectiva/ eeh / me gasto unos pesos y no me meto (...)  
E: ahora viene la última pregunta.  
I: de/ **de todas maneras**/ me sale más barato que andar en auto.  
E: en auto/ sí.  
I: mm.  
E: la bencina es muy cara
- (Martín Núñez 2013:186)

In (28) the speaker, speaking a variety of Chilean Spanish, explains the relative convenience of driving versus taking public transit. She states that she does not prefer using public transport, distancing herself from that idea by saying that it is simply much cheaper to drive.

Penec (2018) presents an example of distancing in written English discourse.

- (29) Mistress Amanda's crazed search for blame finally settled on her husband. She cursed him night and day and threw china dishes against the wall. When Master Ben sent for some medicine to calm her, she swallowed everything she could get her hands on.  
**Anyway**, that's what those who worked in the house said. (Penec 2018:184)

In her analysis, the English marker *anyway* introduces a distancing function. The narrator describes the crazy things *Mistress Amanda* does, but then distances herself from that description. She was not one of the people working in the house, so she does not want to say for sure.

While most researchers agree that distancing constitutes a sub-function of reformulation, Polanco Martínez (2016) is opposed to the idea and offers the following reasons as justification for this stance:

Los marcadores de distanciamiento actúan como mecanismos de formulación, de reorganización discursiva, a través de los cuales se presenta una desviación del punto de vista previo y se inicia una vía discursiva diferente de la anterior. Es cierto que estos marcadores también apuntan hacia lo dicho, en el sentido de que la nueva formulación anula la pertinencia de lo anterior y se presenta como la información que condicionará la continuación del discurso. Pero esta característica es, en cierto modo, común a muchos otros marcadores (si no a todos), en tanto que las operaciones semántico-pragmáticas que señalan los [marcadores discursivos] no afectan solo al segmento que introducen o sobre el que inciden, sino también a la interpretación del primer miembro y, de hecho, a la interpretación global del segmento resultante. Es decir, la interpretación del segmento introducido por un conector (pero, sin embargo, por tanto, etc.) o un operador (bueno, de hecho, en realidad, por supuesto, sin duda, por ejemplo, etc.) implica también la consideración del segmento previo para poder determinar la coherencia global de la enunciación. (2016:16)

His main argument is that the only reason distancing has been considered to be a type of reformulation is that it influences the previous discourse segment, nullifying its relevance and “validity”. He claims that this criterion is not sufficient to label distancing as a type of

reformulation, since many other markers and connectives also affect the interpretation of S1 (i.e., adversatives, additives, etc.). The question that remains, then, is whether or not we should consider distancing to be a function of reformulation. I hope to answer this question in Chapter 4.

A common function of markers typically classified as RMs is that of hedging. I classify this use of RMs under the umbrella sub-category of distancing. Schwenter (1996), working specifically with *o sea*, refers to these as *epistemic uses*. They reflect the speaker's degree of commitment to what she has said. Chodorowska-Pilch (1999) investigates the use of *vamos* with this function. She finds that *vamos* can be used by speakers to mitigate potentially face-threatening utterances, as seen in (30).

(30) Oye, una pregunta, **vamos**, a lo mejor no tiene mayor importancia  
(Chodorowska-Pilch 1999:349)

In (30), the speaker diminishes his intent to ask a question. This mitigation is signaled by the use of *vamos*. Hedging and mitigation are ways in which speakers can distance themselves discursively from a prior utterance, which is why they are included as sub-functions of distancing.

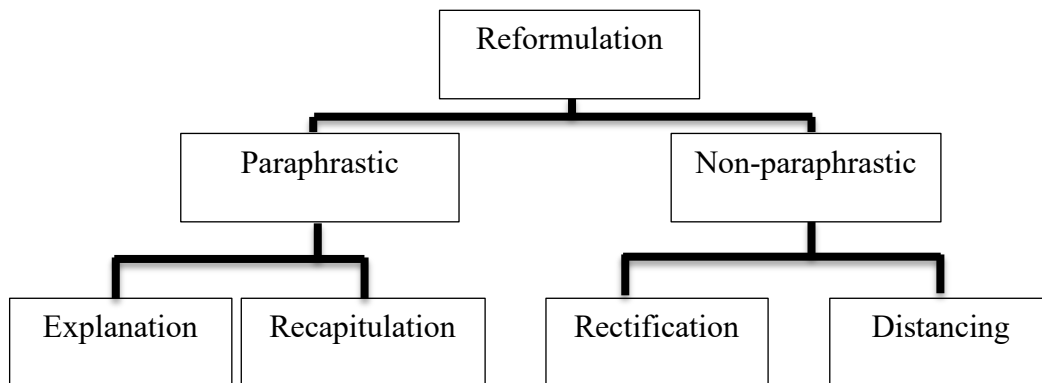
The most commonly used RMs to indicate distancing are *igual*, *en cualquier caso*, *en todo caso*, *de todos modos*, and *de todas formas*. Martín Núñez (2013) finds that the use of *igual* over other distancing markers is significantly higher. This is perhaps due to a higher level of grammaticalization of *igual* with respect to the other markers.

## 2.4 Framework for the present study

In this chapter, I defined several sub-functions of reformulation. Firstly, I distinguished between paraphrastic and non-paraphrastic reformulation. Paraphrastic reformulation deals with relative semantic equivalence of S1 and S2. By contrast, non-paraphrastic reformulation concerns a change of speaker perspective and typically entails less semantic equivalence between

propositions. Secondly, I defined and discussed four sub-functions within paraphrastic and non-paraphrastic reformulation: explanation, which defines, expands, or clarifies what the speaker said in prior discourse; recapitulation, which involves a restatement of the previous utterance in a different way; rectification, which adjusts, amends, or corrects the proposition of the previous utterance due to a change in speaker perspective or focus; and distancing, which results in a change of focus or topic and signals the non-relevance of the previous discourse.

I have chosen these four functions because I believe they each play a crucial role in defining processes of reformulation. Figure 1 illustrates the classification framework for my analysis in Chapter 4.



**Figure 1: The classification framework my analysis.**

## CHAPTER 3

### METHODOLOGY

#### 3.1 Corpus and Data Collection

In order to analyze the use and functioning of reformulation markers in Spanish discourse, I examined data from the *Proyecto para el estudio sociolingüístico del español de España y de América* Corpus (PRESEEA). This project consists of several corpus of transcriptions and audio recordings of 18 oral interviews elicited from nine women and nine men. The corpus divides the interviews by age group and level of education. Group 1 consists of participants between 20 and 34 years of age (6 participants); Group 2, between 35 and 54 (6 participants); and Group 3, 55 and over (6 participants). For education level, six participants approximately five years of education compose Group 1; six participants with approximately 10-12 years of education compose Group 2; and six participants with approximately 15 years of education compose Group 3. I limited the data analyzed to six interviews from the Madrid section of the corpus. The participants in these interviews included three female speakers and three male speakers, with a male and female speaker from each of the three educational levels. All six participants were between 20 and 34 years of age (Group 1). The interviews average 55.5 minutes each.

The authors of the PRESEEA corpus use a specific notation for their transcriptions. For each interview, the interviewer is labeled as “E” and the participant is labeled as “I” in the text of the transcription. Each interview is also given a specific coding, based on the sociolinguistic factors described above. For example, in “H12\_012,” the “H” label denotes a male participant (“M” denotes a female); the first number in the sequence indicates their age group (1, 2, or 3); the

second number, before the underscore, indicates their educational level (1, 2, or 3). The number sequence after the underscore indicates their number out of the overall number of interviews in the corpus. For the purposes of this study, this last number is irrelevant. All examples used in Chapters 4 and 5 are taken from PRESEEA and cited based on their given coding.

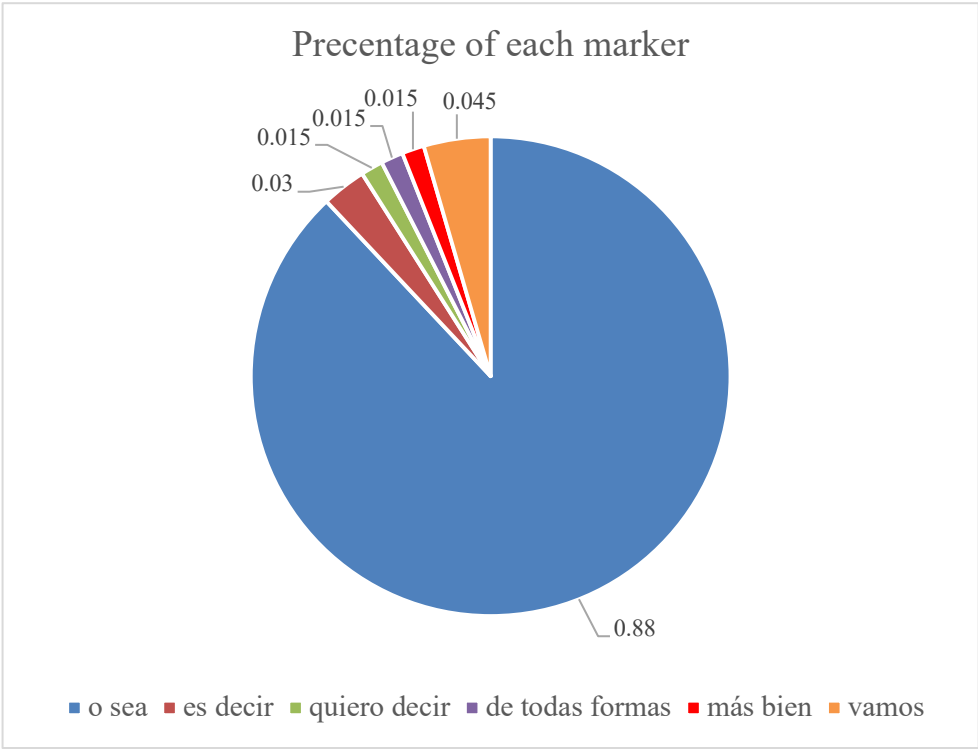
### 3.2 Initial PRESEEA Data Analysis

The data for this study show surprising initial results. My original objective was to analyze one representative RM from each class of reformulation, explanation, recapitulation, correction, rectification, and distancing, as defined in Chapter 2. However, upon searching the corpus for the most commonly cited RMs in each category, I found that many of those markers were not used by Madrilenian speakers at all. Many highly studied markers, such as *igual*, *mejor dicho*, *es más*, *a saber*, and *en fin*, were either not used by speakers, or were used, but not as markers of reformulation (i.e., they were used for other discourse functions).

I elected to analyze the RMs *o sea*, *es decir*, *quiero decir*, *de todas formas*, *más bien*, and *vamos*, because these were the markers found in the data. I exhaustively extracted 336 tokens from six interviews from the Madrid corpus, while identifying 200 tokens of these markers as RMs. I omitted markers used by the interviewers, those that were labeled as part of a dubious transcription, and those that appeared in quotations and thus indicated quoted speech. Table 1 and Figure 1 show the total number extracted for each marker.

**Table 1: The total number of tokens extracted from the data set.**

<b>Marker</b>	<b>Frequency</b>	<b>Percentage</b>
<i>o sea</i>	176	.880
<i>es decir</i>	6	.030
<i>quiero decir</i>	3	.015
<i>de todas formas</i>	3	.015
<i>más bien</i>	3	.015
<i>vamos</i>	9	.045
<b>Total:</b>	<b>200</b>	<b>1.000</b>



**Figure 1: The percentage of each marker in the data set.**

These initial results are quite surprising, given the incredibly high relative frequency of *o sea* tokens (88% of the total set) compared to substantially lower percentages of the other five markers. Frequent use of *o sea* has been attested to in oral discourse (Aydée Hernández 2016; Schwenter 1996); however, other researchers have also attested to high levels of *es decir* (Murillo 2016a) and *vamos* (Polanco Martínez 2016). These low levels of occurrence could be due to the fact that the data were produced by speakers under the age of 34; that is to say, speakers over 34 might be more likely to use the less frequently occurring markers.

Table 2 shows the total number of markers extracted based on the gender of the speaker.

**Table 2: The total number of RMs extracted based on gender.**

<b>Marker</b>	<b>Males</b>	<b>Females</b>	<b>Total</b>
<i>o sea</i>	89	87	176
<i>es decir</i>	5	1	6
<i>quiero decir</i>	2	1	3
<i>de todas formas</i>	1	2	3
<i>más bien</i>	2	1	3
<i>vamos</i>	6	3	9
<b>Total:</b>	<b>105</b>	<b>95</b>	<b>200</b>

We see that about 53% of tokens extracted were used by the men, with 47% used by the women.

The men used *vamos* somewhat more frequently than the women, while differences in the frequency of the other markers according to gender are not noteworthy.

Finally, Table 3 shows the number of markers extracted based on the speakers' level of education.

**Table 3: The total number of RMs extracted based on level of education.**

<b>Marker</b>	<b>Level of Education</b>		
	<u>1</u>	<u>2</u>	<u>3</u>
<i>o sea</i>	57	72	48
<i>es decir</i>	3	2	1
<i>quiero decir</i>	1	2	0
<i>de todas formas</i>	2	1	0
<i>más bien</i>	1	2	0
<i>vamos</i>	1	6	2
<b>Total:</b>	<b>64</b>	<b>85</b>	<b>50</b>

Table 3 shows that speakers from Group 2 (those with secondary education) used RMs more frequently, accounting for about 43% of the data.

### **3.3 Discourse Analysis Methods**

In the next chapter, I present my analysis of the data. For each token, I determine what subclass of reformulation occurs, note the syntactic position of each RM, and discuss the relationship of entailment present (or not) in each case. The criteria for determining S1 and S2 for each token was based on the discourse topic. In each instance, S1 was determined to be the topic being reformulated, while S2 is the reformulated content. The “boundaries” for each segment differ depending on the amount of propositional content determined to be in each segment (i.e., some segments are shorter (or longer) than others aleatorily).

In analyzing the entailment relationship present between segments, I examine entailment in terms of speaker assumptions, rather than based purely on lexical semantic content. In addition to S2 entailing strictly semantic content of S1, I propose that S2 can also entail the speaker *assumptions* of S1, by making the implicit content of S1 explicit. The explicitation of implicated material in reformulated discourse is corroborated by Murillo (2004), who states that RMs “[supply] further explicatures, which in their turn bring about the intended contextual effects, and by [make explicit] implicated premises and conclusions” (2004:2066).

CHAPTER 4  
ANALYSIS AND RESULTS

After extracting 200 tokens of RMs from the PRESEEA corpus, I analyzed them in their discourse contexts to determine what sub-function of reformulation they signaled, where they were located syntactically in each utterance, and whether or not they signaled a relationship of semantic entailment. Firstly, looking at syntactic position, I found that over 90% of the markers appear in medial utterance position, as in (32).

- (32) <cite> ¿tienes envidia? </cite> ¿envidia? pero si yo / mmm el Madrid que gane lo que quiera si yo no veo sus partidos // **o sea** que<alargamiento/> // no / no / no / no tengo ninguna envidia / (H13\_013)

Analysis of the data shows that the RMs examined occurred in either initial, medial, or final position. Table 4 presents the results of this analysis by position.

**Table 4: The total number of tokens based on syntactic position.**

Marker	Position			Total
	Initial	Medial	Final	
<i>o sea</i>	5	161	10	176
<i>es decir</i>	0	4	2	6
<i>quiero decir</i>	0	1	2	3
<i>de todas formas</i>	0	2	1	3
<i>más bien</i>	0	3	0	3
<i>vamos</i>	0	9	0	9
<b>Total</b>	<b>5</b>	<b>184</b>	<b>16</b>	<b>200</b>

Aijmer and Simon-Vandenberg (2011) and Crible (2017) observe that DMs can appear flexibly in any syntactic position. Furthermore, Schwenter (1996) includes an analysis of the

syntactic position of *o sea*, noting that *o sea* can appear in all three positions. However, the high percentage of markers in medial position in the PRESEEA data is not unexpected, because speakers tend to reformulate their preceding utterances, connecting two segments with RMs in discourse-medial position.

The 16 tokens of RMs in utterance-final markers is surprising only in that many of them do not seem to demonstrate any type of reformulation function. Rather, they appear to be used as *hedges* (Fraser 2010), or what Schwenter (1996) refers to as *epistemic uses*, which “[indicate] speakers' degree of commitment to what they say” (1996:865), as seen in (33).

(33) nada de agua / con el propio jugo de la cebolla y del tomate y de la carne // y del vino  
blanco / se hace e<alargamiento/>l jugo / **o sea** (M12\_010)

The speaker in (33) is describing a recipe to the interviewer. She tells part of the cooking process before using *o sea* as a hedge to finish her utterance, as if she is not quite sure of the recipe she has just described. This uncertainty is also suggested by her lengthening of the vowel in the article *el*. The category of hedges was not taken into consideration upon initial analysis, since it is not a function of reformulation. However, due to the highly polysemic nature of these markers (Polanco Martínez 2016), especially in the case of *o sea*, it is not surprising to find them used for other functions.

The 5 tokens found in turn-initial position all consist of cases of *o sea*. Turn-initial markers, specifically DMs and vocatives, have been analyzed in discourse (e.g., Heritage 2013; Heritage & Sorjonen 2018). Fiorentini and Sansò (2017) note that turn-initial RMs “capitalize on the fact that both markers, when used as RMs, may refocus on some specific parts of the first segment, especially when the reformulation introduces a new perspective on what has been said before” (2017:54). The following example illustrates the turn-initial use of *o sea*.

(34) I: yo creo que también la gente tampoco está concienciada de que realmente puede llegar  
a pasar algo <silencio/> no  
E: sí yo creo que ya está

I: **o sea** lo piensas si te lo dicen pero / que no es una cosa que digas (M11\_004)

The participant’s use of *o sea* is in response to the utterance produced by the interviewer; she uses it as a signal that she is going to reformulate what *he* has said.

Secondly, I looked at the classification of each token, based on the framework described in Chapter 2, which identified the following reformulation sub-functions: *explanation* (exp.), *recapitulation* (recap.), *rectification* (rect.), and *distancing* (dist.). Table 5 presents the number of tokens in the data found to have these sub-functions. The “Other” category contains tokens that I consider to be hedges.

**Table 5: The total number of tokens based on reformulative function.**

Marker	Classification					Total
	Exp.	Recap.	Rect.	Dist.	Other	
<i>o sea</i>	115	25	20	1	15	176
<i>es decir</i>	5	0	0	0	1	6
<i>quiero decir</i>	2	0	1	0	0	3
<i>de todas formas</i>	0	0	0	3	0	3
<i>más bien</i>	0	0	0	0	3	3
<i>vamos</i>	1	0	6	2	0	9
<b>Total</b>	<b>123</b>	<b>26</b>	<b>29</b>	<b>6</b>	<b>14</b>	<b>200</b>

By far the most common function used in this data set is that of explanation. This finding is not surprising, given the highly general nature of the explanatory function described in Chapter 2.

This function is illustrated by the following example.

(35) bueno yo estaba / medio saliendo con un chico // que había empezado  
hace<alargamiento/> nada // que era / el sobrino / de la novia / de mi tío / **es decir** //  
bueno / todo quedaba en familia (M14\_004)

The speaker in (35) is describing a situation in which she dated a man whom she discovered had distant ties to her family. She uses the marker *es decir* to signal a coming explanation, that being that she only dated men who had ties to her family.

The second most common function of the RMs found in the data set is that of recapitulation, as exemplified by (36). And, the third most frequently occurring function was rectification, as in (37).

(36) yo me he reído mucho / **o sea** yo me he reído muchísimo (H13\_013)

(37) lo tenía mal ¿no? <risas = "I"/> / **vamos** lo teníamos mal siempre todos ¿no? (H12\_007)

The speaker in (36) uses *o sea* to signal introduction of a restatement of his previous utterance, adding greater emphasis to the second segment with the use of *muchísimo*. In (37) the speaker corrects his first statement—that he ‘had it bad’—by using an the more inclusive first-person plural, which is signaled by the presence of the marker *vamos*.

Finally, I look at the relationship of semantic entailment (or lack thereof) between segments 1 and 2 in each utterance. Table 6 presents the number of tokens that include an entailment relationship.

**Table 6: The number of tokens with and without an entailment relationship.**

Marker	Entailment			Total
	Yes	No	N/A	
<i>o sea</i>	89	73	14	176
<i>es decir</i>	2	4	0	6
<i>quiero decir</i>	1	2	0	3
<i>de todas formas</i>	0	3	0	3
<i>más bien</i>	1	2	0	3
<i>vamos</i>	5	4	0	9
<b>Total</b>	<b>98</b>	<b>89</b>	<b>13</b>	<b>200</b>

In Table 6 reveals a higher number of tokens (49.5%) that consist of a semantic relationship of entailment, although this number is not significant compared to the number not consisting of an entailment (44%). Example (38) illustrates an entailment relationship.

- (38) según mis padres sí // **o sea** mis padres dicen que ahora es muy agresivo que da miedo /  
 ir por la noche por la calle // (M11\_004)  
 what her parents say ||- “according to her parents”

Here, the speaker describes what her parents think of their neighborhood. She says that, according to her parents, the neighborhood is more dangerous than it used to be. She then restates and expands this statement to talk about why they think it is dangerous. We see that there is an entailment relationship between “what her parents say” and what she says they say.

Examples (38) and (39) show tokens without this semantic relationship.

- (39) en mi casa<alargamiento/> nunca han preguntado **o sea** nunca te han dicho <cita> tienes  
 que ir a misa </cita> (H13\_013)

“nunca...han dicho” ¬||- “nunca han preguntado”

- (40) precioso / a mí es que me impactó un montón porque yo nunca había // yo / **o sea** había  
 salido de España pero (M13\_018)

“había salido” ¬||- “nunca había [salido]”

In (39) we understand that S2 *nunca ... han dicho* does not entail *nunca han preguntado*, because the fact that the speaker’s parents never told him explicitly to go to mass, does not entail that they have also never asked if he went to mass. Similarly, in (40) we see that there is no entailment, as the fact that the speaker had never left Spain, does not entail that she had never left the country.

I organize the rest of this chapter based on each marker analyzed in this study: *o sea*, *es decir*, *de todas formas*, *más bien*, *quiero decir*, and *vamos*. Within each subsection I further discuss the syntactic position of the marker in discourse, the sub-function of reformulation under

which they were classified, and whether there exists a relationship of semantic entailment between the segments they connect.

**4.1 *O sea (que)***

As I have already shown, the marker *o sea* accounts for 88% of the data collected. Due to the large number of tokens, this marker also demonstrates the most variability in terms of its function in the discourse. Polanco Martínez (2016) points out that “certain RMs can be stronger, more salient, and more easily accessible discourse options for expressing reformulation” (2016:24, my translation). The results of my analysis suggest that *o sea* appears to be the unmarked choice of speakers for expressing reformulation. Table 7 summarizes all of the characteristics of *o sea* examined.

**Table 7: Summary of the characteristics of *o sea*.**

<i>o sea</i>					
Position	Entailment	Classification			
		Exp.	Recap.	Rect.	Dist.
Initial	<i>Yes</i>	2	0	0	0
	<i>No</i>	3	0	0	0
Medial	<i>Yes</i>	57	25	6	0
	<i>No</i>	58	0	14	1
Final	<i>Yes</i>	0	0	0	0
	<i>No</i>	0	0	0	0

The polysemy of *o sea* is evident in the wide range of functions seen in Table 8. While explanation accounts for the largest number of tokens, recapitulation and rectification were also fairly frequent functions found in the data. The following examples illustrate each of these functions, signaled by *o sea*.

(40) mi vida cambió a positivo desde que conocí a mi novio // **o sea** yo me fui a Canarias a

trabajar cuando hubo el problema // porque<alargamiento/> ... (M11\_004)

The speaker in (40) uses *o sea* to signal an explanation of the first segment, in which she mentions that her life changed when she met her boyfriend. She then explains in segment 2 the circumstances under which they met and why her life changed.

In example (41) the reformulation is a recapitulation.

(41) es que en las empresas constructoras si no eres ingeniero de caminos no eres nada // <ruido = “respiración audible”/> pero no ya ingeniero / de caminos **o sea** tú eres ingeniero industrial y no eres nada en esa empresa (H13\_013)

The speaker comments that the only way to make it at his company is to be a civil engineer. He then restates the same information adding that even industrial engineers are not considered important at his workplace.

The following example involves the rectification of a preceding utterance, introduced by *o sea*.

(42) ¡hombre! de vez en cuando me gusta irme por ahí / de discotecas pero también es que<alargamiento/> // me da fo <palabra\_cortada/> <vacilación/> / **o sea** me <vacilación/> yo creo que tengo un poco de fobia a los sitios cerrados (M11\_004)

The speaker in (42) says that she likes to go out to dance, but then explains why she does not. She begins to say something but truncates her utterance and repairs herself. She uses *o sea* to signal to the hearer that she is making a correction. The reformulation in (43) is another example of rectification.

(43) bueno yo por lo menos suelo hablar de tú // ya con gente má<alargamiento/>s **o sea** de edad superior a la mía suelo hablar de usted / sí / yo sí (M11\_004)

Here, the speaker is describing with whom she does and does not use the pronominal address form *tú*, versus the formal *usted*. She begins her utterance describing the people with whom she uses *tú*, which is followed by *o sea* and a change in sub-topic to describing those with whom she uses *usted*. With this case of rectification, we see a change in her assessment (that she usually uses *tú*) to an admission that she uses *usted* with people older than she.

As we can also see in Table 8, there does not seem to be any pattern in terms of which reformulation functions *consistently* exhibit entailment and which do not. For the explanatory function, the number of tokens with entailment is almost exactly equal to the number of those without. It appears that for this function, entailment with *o sea* is just as likely to occur as not, as shown in (44) and (45).

(44) que ahora roban // <risas = "todos"/> **o sea** a mi hermano<alargamiento/> ya le han robado tres veces (M11\_004)

“ya...han robado tres veces” ||- “roban”

(45) realmente es ficticio **o sea** no estamos registrados en<alargamiento/> / como en la comunidad de Madrid ni en ningún lado / (H12\_007)

not being registered ¬||- the “fictitiousness” of the club to which he belongs

We see in (44) that “they have robbed three times” entails that they rob people, while the reverse does not hold true. In (45), the speaker is discussing his soccer club. He says that the club is not really legitimate, it is somewhat fake or unofficial—that is, they are not even registered in the community. We understand that not being registered for something does not entail that it is fictitious or fake.

The use of RMs for the corrective function can involve either segments related by entailment or segments where there is no semantic entailment, as seen in (46), where S2 entails S1, and (47), where no such entailments exists.

(46) todo el mundo se echó encima de mí / **o sea** se apoyó en mí // (M11\_004)

“se apoyó en mí” ||- “se echó encima de mí”

(47) E: ¿sí? / ¿por qué <vacilación/> / por qué<alargamiento/> decidiste ir al ejército? / I: pff bueno pues en parte por eso un trabajo fijo dos años // sabía que<alargamiento/> ahí a <vacilación/> a la hora de pagar al final de mes no pagan // y<alargamiento/> **o sea** / perdón / que<alargamiento/> no fallan en pagar quiero decir (H11\_002)

“no fallan en pagar” ¬||- “al final de mes no pagan”

In (46) the speaker is describing how her friends used her as support system. We see that, metaphorically, “they used me as support” entails “the threw themselves all over me”, while the reverse is not necessarily true. In (47) the speaker tells why he decided to join the army, stating that it was because they do not pay at the end of the month. He corrects himself, apologizing, saying that they do not *fail* to pay. We understand that “they do not fail to pay” does not entail “they do not pay.”

In the data analyzed for this study, the recapitulation function demonstrates a strong tendency towards the presence of entailment. All 25 cases of recapitulation with *o sea* exhibit entailment. This logically follows, given that recapitulation is a restatement of the same propositional content. Example (48) illustrates recapitulation function involving an entailment relationship.

(48) me estuve media hora riéndome de ellos / <risas = "E"/> pero porque es que eran dos niños **o sea** que me llegaban a la cintura<alargamiento/> un poco más arriba de ocho años tendrían ¿sabes? (M11\_004)

“me llegaban a la cintura...un poco más arriba de ocho años tendrían” ||- “eran...niños”

The speaker in (48) is describing a comical situation in which she was almost robbed by children. The description of the robbers as being “no more than 8 years old” entails that they were children.

By contrast, no relationship of semantic entailment occurred in the 9 sequences where RMs introduced utterances with the rectification function. Example (49) exemplifies the rectification function with no entailment relationship.

(49) en la planta de arriba no hay nada / <entre\_risas> es una </entre\_risas> **o sea** está el suelo puesto y las paredes pero hay un futbolí<alargamiento/>n (M12\_010)

“hay un futbolín” ¬||- “en la planta de arriba no hay nada”

The speaker is describing her friend’s apartment. She states that there is nothing upstairs, but then explains that there are just flooring and walls. The first half of S2 entails S1, however, the

adversative conjunction introduces information that cancels the entailment (there is foosball table).

## 4.2 *Vamos*

The marker *vamos* has the second-highest number of tokens in the data set, accounting for 4.5% of the RMs. Table 8 summarizes the characteristics of the instances of *vamos* as an RM examined.

**Table 8: A summary of the characteristics of *vamos*.**

Marker	Entailment	Classification			
		Exp.	Recap.	Rect.	Dist.
Medial	<i>Yes</i>	1	0	4	0
	<i>No</i>	0	0	2	2
Final	<i>Yes</i>	0	0	0	0
	<i>No</i>	0	0	0	0

In the data analyzed, *vamos* was found to signal the explanation, rectification, or distancing function. Polanco Martínez (2016:30) proposes that DMs such as *vamos* and *bueno* due to their underspecified and polysemic metadiscursive roles, take on the sense of the functional context in which they are used. Thus, *vamos* and *bueno*, can be used as RMs, as well as other types of DMs, based on their polysemic nature. This also explains why there are problems with the classification of markers. Specifically, unless they have a highly specified function and procedural meaning, many RMs, such as *bueno* and *vamos*, can act as indicators of other functions as well. Example (60) illustrates the explanation function introduced by *vamos*.

- (60) pues nada / primero cojo la carne picada / **vamos** hago espaguetis a<alargamiento/> a la boloñesa / entonces cojo la carne picada / (M13\_018)

Here, the speaker is describing a recipe she knows how to make. She starts her explanation of the process but forgets to tell what the recipe is for. She then uses *vamos* to signal inserting an explanation involving name the dish she is describing, and then continues her description of the recipe. In other words, *vamos* is used to introduce a clarification.

Example (61) shows the rectification function introduced by *vamos*.

- (61) pues<alargamiento/> </simultáneo> // pues sí me acuerdo por ejemplo<alargamiento/> en el año<alargamiento/> ¿qué sería? en el ochenta y<alargamiento/> <silencio/> / seis o por ahí / ochenta y cinco ochenta y seis que cayó aquí una<alargamiento/> // una nev <palabra\_cortada/> **vamos** / una nevada y<alargamiento/> una<alargamiento/> bajada de temperatura / tan bestial // que (H12\_007)

In (61), the speaker is describing a change in the weather. He starts telling his story but seems hesitant or unsure. He begins the word *nevada*, but truncates it, as if wanting to repair his word choice. *Vamos* is used to mark the repair, but the speaker uses the same word, *una nevada*.

Another example of the rectification function signaled by *vamos* is exhibited in (62).

- (62) eeh lo normal es que mucha gente<alargamiento/> // participa / **vamos** compite / esperando el fallo de los demás (H12\_007)

The speaker in (62) describes a boating competition in which he takes part. He says that normally many people participate, quickly rectifying his statement and substituting *participa* (participate) with the stronger word, *compite* (compete) partially cancelling the word choice of the previous segment.

Example (63) exemplifies the distancing function signaled by *vamos*.

- (63) E: cuéntame alguna receta que sepas / que sepas hacer  
I: pues<alargamiento/> no sé / son cosas tontas ¿no? pero **vamos** // te puedo dar una receta si te gusta el salmón ahumado / (H12\_007)

In this example, the speaker is again asked to describe a recipe to the interviewer. The speaker comments on the silliness of recipes—or perhaps the silliness of asking about them in the interview—but quickly distances himself from the preceding segment, in which underrates his

cooking abilities, by starting to explain his recipe. He uses *vamos* to signal the change of topic from his insecurity about cooking to describing a special dish.

In terms of semantic entailment, the tokens of *vamos* analyzed in this data set also follow the pattern of *o sea* and *es decir*; entailment is just as likely to occur as not. The following examples illustrate an entailment (64) and a non-entailment (65).

(64) E: ¿cómo es ella? Descríbemela  
I: es rubia // **vamos** / <entre\_risas> es teñida de rubio (M13\_018)  
“es teñida de rubio” ||- “es rubia”

(65) que yo creo que es la única paliza que me han dado / **vamos** / que yo <entre\_risas>  
recuerde </entre\_risas> (H12\_007)  
“que yo recuerde” ¬||- “creo que es la única paliza que me han dado”

In (64) the speaker jokes about her friend having blonde hair, admitting that she actually dyes it. We understand that having dyed blonde hair still entails being blonde. In (65) the speaker is recounting a time when he and his friends got beaten up. He says it is the only time he has ever received a beating, but quickly distances himself from that statement using *vamos* before admitting it is the only beating he remembers. It is understood that “what he remembers” is not entailed by “what he believes.”

### 4.3 *De todas formas*

With only 3 tokens of *de todas formas* in this data set, it is difficult to make claims about the nature of this marker and its functions. Table 9 summarizes the characteristics of *de todas formas* examined.

**Table 9: Summary of the characteristics of *de todas formas*.**

<i>de todas formas</i>					
Position	Entailment	Classification			
		Exp.	Recap.	Rect.	Dist.
Medial	<i>Yes</i>	0	0	0	0
	<i>No</i>	0	0	0	1
Final	<i>Yes</i>	0	0	0	0
	<i>No</i>	0	0	0	1

In this data set, two tokens of *de todas formas* signal only the distancing function, seen in (50).

- (50) se saltó el semáforo un<alargamiento/> / un señor mayor // y<alargamiento/> bueno tuve la<alargamiento/> / la suerte o<alargamiento/> los reflejos de poder pisar el freno pero vamos / **de todas formas** yo le di a él // (H12\_007)

The speaker in (50) describes a traffic accident caused by an older man running a traffic light. He conveys his displeasure by commenting that he was lucky enough to be able to slam on the brakes, but seemingly distances himself from that attitude, introducing a conclusion that he hit the man anyway. Murillo (2016b) cites conclusion as a function of segments introduced by RMs.

Since distancing is a function which signals the non-relevance of a previous segment, it follows that an entailment relationship would not be present. This is the case in (51).

- (51) hombre he escuchado pero de Madrid de lo típico de <cita> a ese le han pegado una paliza </cita> o<alargamiento/> / o los <extranjero> skings </extranjero> famosos que ahora se habla menos de ellos // parece que está la cosa más / controlada // <ruido = "voces femeninas"/> pero no aquí / **de todas formas** es que este es un barrio tranquilo / (M11\_004)

“este es un barrio tranquilo” ¬||- “lo típico de Madrid”

The speaker in (51) describes the neighborhood where she grew up, saying that she has heard all the bad things people say about Madrid (that people get beat up, etc.), but distances her neighborhood from that stereotype—it is much calmer and quieter. It follows that being a quiet neighborhood does not entail the idea of a typical Madrilenian neighborhood, if the typical neighborhood is dangerous.

#### 4.4 *Es decir*

Like *de todas formas*, there were few tokens of *es decir* in the data set. Table 10 summarizes the characteristics of the three tokens of *es decir* examined.

**Table 10: A summary of the characteristics of *es decir*.**

<i>es decir</i>					
Position	Entailment	Classification			
		Exp.	Recap.	Rect.	Dist.
Medial	<i>Yes</i>	2	0	0	0
	<i>No</i>	2	0	0	0
Final	<i>Yes</i>	0	0	0	0
	<i>No</i>	1	0	0	0

In this data set, *es decir* signals only the explanatory function, as seen in (52).

- (52) E: ¿a qué federación perteneces?  
 I: a la<alargamiento/> Real Federación Española de Motonáutica  
 E: uhum  
 I: *o sea es decir* nuestra<alargamiento/> modalidad está incluida<alargamiento/> junto con la de <vacilación/> las de motos de agua las de <alargamiento/> (H12\_007)

The speaker in (52) is describing his speedboating racing club. He gives the name of the group, the *Real Federación Española de Motonáutica*, and then explains that their category of competition is included in the same category as jet-skis. Interestingly in this example, the speaker uses a combination of two RMs, *o sea* and *es decir*. Fraser (2015) notes that when two DMs (in this case RMs) are combined in the same utterance, the first one used (in this case, *o sea*) has a “more general” function than the second one (in this case, *es decir*) (2015:49). Given the wide range of functions *o sea* signals in this data set, and the narrower range that *es decir* signals in the data, this corroborates Fraser’s findings. Further investigation of DM/RM combination in Spanish is beyond the scope of this thesis; however, this is an interesting topic for future research.

In terms of an entailment relationship, *es decir* follows the pattern of *o sea* in that entailment is as likely to occur as it is unlikely. The following two examples show an entailment relationship (53) and a non-entailment relationship (54). In (53), repeated from example (34), the speaker is describing the relationship a previous boyfriend had with her family.

- (53) bueno yo estaba / medio saliendo con un chico // que había empezado  
 hace<alargamiento/> nada // que era / el sobrino / de la novia / de mi tío / **es decir** //  
 bueno / todo quedaba en familia (M14\_004)

“todo quedaba en familia” ||- “era el sobrino de la novia de mi tío”

The perception of “everything staying in the family” entails the complex relations they have with each other’s families, described in S1. In other words, saying “todo quedaba en familia” entails a relationship such as uncle’s girlfriend’s nephew.

The speaker in (54) describes his relationship with his best friend.

- (54) E: ... ¿y por qué ese es tu mejor amigo? //  
 I: porque le conozco de toda la vida **es decir** hemos <vacilación/> / nos hemos  
 dejado<alargamiento/> de ir con otra gente / juntos // bueno con otros dos chavales  
 también pue<alargamiento/>s // será por algo / nos llevamos / muy bien (H12\_002)

not going out with other friends ¬||- knowing someone your whole life

He says that they have stopped going out with groups of other people because they have known each other their whole lives and they get along well, thus explaining the nature of their friendship. We understand that knowing someone your whole life does not entail deciding not to go out together with other friends.

#### 4.5 *Más bien*

The marker *más bien* occurred three times in the data set. Table 11 summarizes the characteristics of *más bien* examined.

**Table 11: A summary of the characteristics of *más bien*.**

<i>más bien</i>						
Position	Entailment	Classification				
		Exp.	Recap.	Rect.	Dist.	Other
Medial	<i>Yes</i>	0	0	0	0	1
	<i>No</i>	0	0	0	0	2
Final	<i>Yes</i>	0	0	0	0	0
	<i>No</i>	0	0	0	0	0

Each token of *más bien* signals a different function, none of which are part of the framework proposed in the study. The interviewer in (55) is asking about the participant’s neighborhood. He asks whether people think the neighborhood is dangerous; the participant corrects him, asserting that it is not.

- (55) E: no bueno ¿y <vacilación/> y la opinión de la gente? / ¿es que es un barrio<alargamiento/> peligroso o<alargamiento/>? //  
 I: no / **más bien** <silencio/> peligroso no // (H11\_002)

Not dangerous ||- not (dangerous)

Portolés (2001) classifies this marker as a rectification RM, but in this instance the speaker seems to hedge and then reaffirms his answer. It suggests that he is going to rectify his first answer (no), but then decides not to do so. In terms of entailment, something not being dangerous entails it being not (dangerous).

In (56), the speaker is describing a friend’s house at which she and her friends frequently stay/visit. She claims that the upstairs area is not furnished. However, she then admits that when they are there (*cuando vamos*), there is furniture:

- (56) en la planta de arriba no hay nada / <entre\_risas> es una </entre\_risas> o sea está el suelo puesto y las paredes pero hay un futbolí<alargamiento/>n / una camita y tal y<alargamiento/> juegos en plan una pelota / un billa<alargamiento/>r / para divertirte ¿no? más <vacilación/> **más bien** cuando vamos (M12\_010)

when we go there ¬||- “hay un futbolín, una camita, etc.”

She uses *más bien* to introduce a type of summing up of the situation: the items listed are what they encounter when they go there. There is no entailment relationship between “when they go there” and the list of items that comprise S1.

The third token of *más bien* analyzed in the data was determined to an intensifier, not an RM. This instance is categorized as “Other” in the analysis.

#### 4.6 *Quiero decir*

The data set yielded 3 tokens of the marker *quiero decir*. Table 12 summarizes the characteristics of these instances.

**Table 12: A summary of the characteristics of *quiero decir*.**

<i>quiero decir</i>					
Position	Entailment	Classification			
		Exp.	Recap.	Rect.	Dist.
Medial	<i>Yes</i>	1	0	0	0
	<i>No</i>	0	0	0	0
Final	<i>Yes</i>	0	0	0	0
	<i>No</i>	1	1	1	0

*Quiero decir* signals introduction of an explanation in two cases, and introduction of a corrective utterance in the third. However, none of the instances marks an entailment relationship. *Quiero decir* is also one of very few markers that maintains the reformulative function in turn-final position. In the following example, *quiero decir* introduces an explanation.

- (58) llevamos / tres años / de mm tres años / dos años que hemos estado<alargamiento/> los dos primeros pusimos un<alargamiento/> / un puesto en la playa // **quiero decir** yo he estado trabajando con ellos en verano / y<alargamiento/> ya lleva / eeh / este año hace un año la tienda / eh montaron una tienda allí en Ayamonte / y<alargamiento/> y ellos están ahí viviendo allí tienen ellos una casita muy chiquitita (M12\_010)

working with them in the summer ||- putting up two stands on the beach with them

In (58), the speaker is describing his work with a couple in Ayamonte. He comments that he works with the couple during the summer and that for the first two summers, they put up a stand at the beach. Here, there is an entailment relation between the segments: working with the couple during the summer entailed putting stands at the beach the first two summers.

Example (59) illustrates a rectification.

(59) pff bueno pues en parte por eso un trabajo fijo dos años // sabía que<alargamiento/> ahí a <vacilación/> a la hora de pagar al final de mes no pagan // y<alargamiento/> o sea / perdón / que<alargamiento/> no fallan en pagar **quiero decir** (H11\_002)

The speaker in (59), cited earlier as (46) is describing why he decided to join the army. He says that the army does not pay but corrects himself by stating that the army does not forget to pay. He first uses the marker *o sea* to signal his correction, but also uses *quiero decir* turn-finally to signal the same function, almost like a reminder to the hearer that the speaker has made a reformulation.

#### 4.7 Summary and Conclusions

Table 13 presents a summary of the characteristics of all the markers I investigated.

**Table 13: A summary of the characteristics analyzed.**

Marker	Position	Entailment?	Classification				
			Exp.	Recap.	Rect.	Dist.	Other
<i>o sea</i>	<b>Medial</b>	<i>Yes</i>	57	25	6		
		<i>No</i>	58		14	1	15
<i>de todas formas</i>	<b>Medial</b>	<i>Yes</i>				1	
		<i>No</i>				1	
	<b>Final</b>	<i>Yes</i>					
		<i>No</i>				1	
<i>es decir</i>	<b>Medial</b>	<i>Yes</i>	2				
		<i>No</i>	2				
	<b>Final</b>	<i>Yes</i>					
		<i>No</i>	1				1
<i>más bien</i>	<b>Medial</b>	<i>Yes</i>					1
		<i>No</i>					2
<i>quiero decir</i>	<b>Medial</b>	<i>Yes</i>					
		<i>No</i>	1				
	<b>Final</b>	<i>Yes</i>	1				
		<i>No</i>			1		
<i>vamos</i>	<b>Medial</b>	<i>Yes</i>	1		2		
		<i>No</i>			2	2	

## CHAPTER 5

### CONCLUSIONS

For this thesis I extracted and analyzed 200 tokens of discourse markers traditionally classified as RMs. My analysis had three objectives: to identify the sub-function of reformulation each marker signaled; to determine where they were located syntactically in relation to the discourse segments surrounding them; and to discern whether or not they signaled a relationship of semantic entailment. The markers examined were *o sea*, *de todas formas*, *es decir*, *más bien*, *quiero decir*, and *vamos*.

The first characteristic analyzed was syntactic position in the utterance, which could be turn-initial, medial, or final. I found that the majority of RMs occurred in utterance-medial position, which follows logically, since most speakers reformulate their own discourse, as shown in (66).

(66) pues unas chicas expusieron mi<alargamiento/> trabajo // **o sea** exactamente lo mismo  
(M12\_010)

In the data, those markers found in utterance-final position generally did not act as RMs, but instead as hedges, as seen in (67).

(67) es mucho más cómodo el autobús porque no tienes que hacer ningún traspaso ni nada / y<alargamiento/> nada ya llegar a clase / normalmente estoy<alargamiento/> seis horas / o sea salgo a la<alargamiento/>s dos y media / a las dos y media<alargamiento/> excepto los miércoles que estoy hasta las siete y media de la tarde // **o sea** (M12\_010)

When found in utterance-initial position, the RMs acted to signal a reformulation the speaker's previous turn, as in (68).

(68) I: y no sé haría un montón de cosas / tendría que pensarlo pero  
E: sí sí vete pensándolo porque llevas camino de<alargamiento/> ¿no? <risas = "E"/>  
I: **o sea** yo<alargamiento/> no descarto<alargamiento/> dedicarme a la política algún día  
(H13\_013)

Due to the paucity of literature on RM syntactic position, I believe these findings to be significant, although future studies with larger data sets should be used to corroborate them. Additionally, the frequency of occurrence of RMs in each position is largely dependent upon the genre of discourse analyzed. For instance, conversational interactions with frequent turn-taking might exhibit higher rates of turn-initial RMs. By contrast, lengthy monologues (story-telling, anecdotal discourse) would most likely contain more discourse-medial RMs. The use of turn-final RMs could be expected to occur in conversational discourse, particularly to mark the point at which the current speaker cedes the floor.

The second characteristic examined in this data set was the classification of each token. This was the most difficult part of the analysis due to the low degree of consensus among researchers as to an appropriate classification framework for reformulation and RMs. Gülich and Kotschi (1995), Portolés (2001), Pons-Bordería (2001, 2013, 2017), Garcés Gómez (2003, 2010), Cuenca (2003), and Murillo (2016 a, b) all propose different classifications for reformulation and RMs. The classification framework for this study divided reformulation into the four sub-functions proposed by Portolés (2001): explanation, which clarifies, defines, or expands what the speaker said in his previous utterance; recapitulation, which functions to restate in a different way what the speaker has just said; rectification, which partially or totally annuls the proposition of the previous utterance by amending, adjusting, correcting, or modifying it; and distancing, which signals the non-relevance of the previous discourse and can result in a change of topic or a hedge. This framework provided an adequate system for classifying the markers from this data set. After analyzing this data set, I also included an “other” category for those markers from the set of markers analyzed that did not exhibit a reformulation function.

The final characteristic analyzed in the study was the relationship of entailment present between the segments joined by each token. This relationship occurs between two propositions,

with the truth of one ensuring the truth of the other. In the case of reformulation, the propositions are the discourse segments connected by an RM. To the best of my knowledge, an entailment relationship between reformulated segments has not been investigated until now. I found that the explanation function allowed for an entailment relationship just as often as not. Only one function, recapitulation, allowed this relationship in all cases, while the distancing function allowed it in no cases. The rectification function disallowed it in most cases, but not all. Further study is needed to investigate this relationship between reformulated utterances.

While I believe these findings to be significant for the study of DMs and reformulation, I must acknowledge the shortcomings of the data and the analysis. While 200 tokens are not a small number, the majority (88%) of those tokens were of one marker (*o sea*); the other 12% is divided among the other five markers studied. The unusually small number of RMs found in the data set makes it difficult to make broad generalizations about their nature in discourse. In future studies, a more equal number of tokens of each RM, as well as a larger overall number of tokens would provide a more reliable and consistent analysis of RM function.

### **5.1 Suggestions for Future Study**

Several questions remain for future study. From a sociolinguistics standpoint, the variables of age, education level, and gender would be interesting to examine in to determine any correlation between RM use and these variables. In this study, I controlled the data set based on age, which could have influenced the markers used and their frequency. A brief secondary search of the PRESEEA corpus revealed higher numbers of the RMs *es decir*, *más bien*, and *quiero decir* used by speakers 34 and older. By contrast, these three RMs accounted for only 6% of RMs that appeared in the data analyzed in this study. I did not control for educational level. However, based on the data analyzed, Group 2 (participants with 10 or less years of schooling) tended to

use RMs more frequently. A more thorough study could corroborate this finding. The last factor controlled for was gender. The findings from the data analysis indicate no significant difference between male and female use of RMs; however, a larger number of participants would be needed to reach any clear conclusions regarding gender effects.

From a discourse pragmatic perspective, an investigation of combinations of RMs would be interesting to study. Work on DM combination has already been investigated (Fraser 2015). However, no work has been done, to my knowledge, on RM combination. Example (52), repeated here as (69), shows the combination of *o sea* and *es decir* in turn-initial position.

- (69) I: a la<alargamiento/> Real Federación Española de Motonáutica  
E: uhum  
I: **o sea es decir** nuestra<alargamiento/> modalidad está incluida<alargamiento/> junto con la de <vacilación/> las de motos de agua las de <alargamiento/> (H12\_007)

As analyzed in Chapter 4, this example illustrates the explanation function. As this is the only token of RM combining in this data set, further research is needed to determine whether or not the explanation function is the only one that permits combining. Research on which RMs permit combination with other RMs would be interesting to study as well. A future study could focus on which RMs combine with other RMs in spoken Spanish discourse.

Additionally, an examination of the translatability of RMs from Spanish discourse into English would be another extension of this thesis. RMs such as *o sea* and *bueno* are notoriously difficult to translate, due to their polysemic natures. A study of how television shows and/or movies dub their programs from English to Spanish and from Spanish to English, as well as the translation of subtitles would be quite interesting.

## REFERENCES

- Aijmer, K. & Simon-Vandenberg, A. 2011. Pragmatic markers. In *Discursive Pragmatics [Handbook of Pragmatics 8]*, Jan Zienkowski, Jan-Ola Ostman & Jef Verschueren (eds), 223–247. Amsterdam: John Benjamins.
- Aydée Hernández, M. 2016. “El marcador discursivo *o sea* en el español hablado de Medellín.” *Lingüística y Literatura*, 69, 295-314.
- Bataller, S. M. 2002. Reformulations and Relevance Theory Pragmatics: The Case of TV News Interviews. *Studies in English Language and Linguistics*, 12, Ed. F. Fernández, 41-98.
- Blakemore, D. 1987. Semantic constraints on relevance. Oxford: Blackwell.
- Blakemore, D. 1996. Are apposition markers discourse markers? *Journal of Linguistics*, 32, 325–347.
- Castillo Fadić, M. N. & Insua, E. S. 2017. El reformulador <es decir> en el español de Chile: Una propuesta de clasificación funcional. *Lenguas Modernas*, 49, 77-92.
- Chodorowska-Pilch, M. 1999. On the polite use of *vamos* in Peninsular Spanish. *Pragmatics*, 9(3), 343-355.
- Crible, L. 2017. Towards an operational category of discourse markers: A definition and its model. *Pragmatic markers, discourse markers, and modal particles: New perspectives*, (Eds. C. Fedriani & A. Sansó), 99-124.
- Cuenca, M. J.. 2003. Two ways to reformulate: A contrastive analysis of reformulation markers. *Journal of Pragmatics*, 35(7), 1069-1093.

- Cuenca, M. J., & Bach, C. 2007. Contrasting the form and use of reformulation markers. *Discourse Studies*, 9(2), 149–175.
- Del Saz Rubio, M. M. 2007. *English Discourse Markers of Reformulation*. Bern, Switzerland: Peter Lang.
- Fernández Bernárdez, C. 2000. “Quiero decir” como marcador de reformulación. *RILCE: Revista de filología hispánica*, 16(2), 263-288.
- Fiorentini, I. & Sansò, A. 2017. Reformulation markers and their functions: Two case studies from Italian. *Journal of Pragmatics*, 120, 54-72.
- Fischer, K. (ed.) 2006. *Approaches to discourse particles*. Elsevier.
- Fraser, B. 1990. An approach to discourse markers. *Journal of Pragmatics*, 14, 383-395.
- Fraser, B. 1996. Pragmatic markers. *Pragmatics*, 6(2), 167–190.
- Fraser, B. 1999. What are discourse markers? *Journal of Pragmatics*, 31, 931–952.
- Fraser, B. & Malamud-Makowski, M. 1996. English and Spanish Contrastive Discourse Markers. *Language Sciences*, 18(3-4), 863-881.
- Fraser, B. 2015. The combining of discourse markers—a beginning. *Journal of Pragmatics*, 86, 48-53.
- Garcés Gómez, M. P. 2003. Los marcadores de recapitulación y reconsideración en el discurso. *Revista de Investigación Lingüística*, 1(6), 111-141.
- Garcés Gómez, M. P. 2010. Marcadores de corrección y rectificación en los textos escritos. *Revista de Investigación Lingüística*, 16, 87-105.
- Garcés Gómez, M. P. 2011. Procesos de reformulación en el discurso oral: Corrección y rectificación. *Español actual: Revista de español vivo*, (96), 89-106.
- Gülich, E. & Kotschi, T. 1995. Discourse production in oral communication: A study based on French. *Aspects of Oral Communication*, 21, Ed. U. M. Quasthoff, 30-66.

- Heritage, J. 2013. Turn-initial position and some of its occupants. *Journal of Pragmatics*, 57, 331-337.
- Heritage, J. & Sorjonen, M.-L. 2018. Introduction: Analyzing turn-initial particles. In J. Heritage & M.-L. Sorjonen (Eds.), *Between turn and sequence: Turn-initial particles across languages. Studies in Language and Social Interaction*, 31, 1-22.
- Labov, W. & Fanshel, D. 1977. *Therapeutic discourse*. New York: Academic Press.
- Levinson, S.C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Murillo, S. 2004. "A relevance reassessment of reformulation markers." *Journal of Pragmatics*, 36, 2059–2068.
- Murillo, S. 2016a. "Sobre la reformulación y sus marcadores." *Cuadernos AISPI*, 8, 237-258.
- Murillo, S. 2016b. "Reformulation markers and polyphony: A contrastive English-Spanish analysis." *Languages in Contrast*, 16(1), 1-30.
- Polanco Martínez, F. 2016. Marcadores del discurso y contextos funcionales: La reformulación como construcción discursiva. *Onomázein*, 34, 14-35.
- Pons Bordería, S. 2013. "Un solo tipo de reformulación." *Cuadernos AISPI*, 2, 151-170.
- Pons Bordería, S. 2017. "Volviendo sobre *un solo tipo de reformulación*: una respuesta a Silvia Murillo." *Cuadernos AISPI*, 10, 153-172.
- Portolés, J. 2001. *Marcadores del discurso* (2nd ed.). Barcelona: Ariel.
- PRESEEA. 2014. *Corpus del Proyecto para el estudio sociolingüístico del español de España y de América*. Alcalá de Henares: Universidad de Alcalá.
- San Martín Núñez, A. & Guerrero González, S. 2016. Los marcadores de reformulación en el corpus PRESEEA de Santiago de Chile. *Forma y Función*, 29(2), 15-38.
- Schiffrin, D. 1987. *Discourse Markers*. Cambridge: CUP.
- Schourup, L. 1999. Discourse markers. *Lingua*, 107, 227-265.

Schwenter, S. A. 1996. Some reflections on *o sea*: A discourse marker in Spanish. *Journal of Pragmatics*, 25, 855-874.

Sperber, D. & Wilson, D. 1995. Relevance. *Communication and Cognition*. Oxford: Blackwell.

Zwicky, A. M. 1985. Clitics and particles. *Language*, 61, 283-305.