

The Middle Voice and Connected Constructions in Ibero-Romance

*A variationist and
dialectal account*

Carlota de Benito Moreno

John Benjamins Publishing Company

Studies in Language Variation

29

The Middle Voice and Connected Constructions in Ibero-Romance

Studies in Language Variation (SILV)

ISSN 1872-9592

The series aims to include empirical studies of linguistic variation as well as its description, explanation and interpretation in structural, social and cognitive terms. The series covers any relevant subdiscipline: sociolinguistics, contact linguistics, dialectology, historical linguistics, anthropology/anthropological linguistics. The emphasis will be on linguistic aspects and on the interaction between linguistic and extralinguistic aspects – not on extralinguistic aspects (including language ideology, policy etc.) as such.

For an overview of all books published in this series, please see benjamins.com/catalog/silv

Editors

Peter Auer
Universität Freiburg

Frans Hinskens
Meertens Instituut &
Radboud Universiteit
Nijmegen

Paul Kerswill
University of York

Editorial Board

Suzanne Aalberse
Universiteit van Amsterdam

Arto Anttila
Stanford University

Gaetano Berruto
Università di Torino

Jenny Cheshire
University of London

Katie Drager
University of Hawai'i at Mānoa

Katarzyna Dziubalska-
Kołaczyk
Adam Mickiewicz University in
Poznań

Jürg Fleischer
Philipps-Universität Marburg

Peter Gilles
University of Luxembourg

Brian D. Joseph
The Ohio State University

Johannes Kabatek
Universität Zürich

Pia Quist
University of Copenhagen

Anne-Catherine Simon
Université catholique de
Louvain

Sali A. Tagliamonte
University of Toronto

Øystein Alexander Vangsnes
UiT The Arctic University of
Norway

Juan A. Villena Ponsoda
Universidad de Málaga

Volume 29

The Middle Voice and Connected Constructions in Ibero-Romance
A variationist and dialectal account
by Carlota de Benito Moreno

The Middle Voice and Connected Constructions in Ibero-Romance

A variationist and dialectal account

Carlota de Benito Moreno

University of Zurich

John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

DOI 10.1075/silv.29

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2022020593 (PRINT) / 2022020594 (E-BOOK)

ISBN 978 90 272 1142 2 (HB)
ISBN 978 90 272 5758 1 (E-BOOK)

© 2022 – John Benjamins B.V.

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

John Benjamins Publishing Company · <https://benjamins.com>

Table of contents

Acknowledgements	IX
CHAPTER 1	
Reflexive constructions: An introduction	1
1.1 Introduction	1
1.2 Reflexive constructions in the varieties under study: State of the art	4
1.2.1 Purely reflexive and reciprocal verbs	7
1.2.2 The relationship of the reflexive verb and its unmarked counterpart	10
1.2.3 Semantic criteria in typologies of paradigmatic reflexive constructions	15
1.2.4 “Difficult cases”	17
1.2.5 Summary	18
1.3 The diachrony of reflexive constructions	19
1.3.1 Overall development	21
1.3.2 Indirect reflexive constructions	22
1.3.3 From reflexive to reciprocal	23
1.3.4 From reflexive to middle	23
1.3.5 Reflexive intransitive and transitive verbs	24
1.3.6 Summary	26
1.4 Typological approaches to reflexive and middle constructions	26
1.4.1 Voice and diathesis	27
1.4.2 The middle voice	28
1.4.3 The middle voice as a derived voice	31
1.4.4 The middle voice as a basic voice	33
1.5 Theoretical framework	38
1.5.1 Language change and variation	38
1.5.2 Mechanisms of language change	40
1.5.3 Spatial diffusion of language change	41

CHAPTER 2

Methodology

43

- 2.1 Introduction 43
- 2.2 The COSER data 43
 - 2.2.1 Selected interviews 44
 - 2.2.2 Collection of examples 47
 - 2.2.3 Analysis of the data 50
 - 2.2.4 Note on spelling and linguistic representation 51
- 2.3 Other sources 52
 - 2.3.1 Video questionnaire 52
 - 2.3.2 Written examples 55
- 2.4 Technical notes 56
 - 2.4.1 Statistical methods 56
 - 2.4.2 Geographical representation 58

CHAPTER 3

The middle voice in Spanish

59

- 3.1 Introduction 59
- 3.2 The middle voice as a derived voice 60
 - 3.2.1 Subject deletion diathesis: Anticausative verbs 60
 - 3.2.2 Object-deletion diatheses: Absolute constructions 66
 - 3.2.3 Object-deletion diatheses: De-objective verbs 70
 - 3.2.4 Subject-demoting diathesis: Conversive verbs 75
 - 3.2.5 Object-demoting diathesis: Antipassive verbs 78
 - 3.2.6 No change of diatheses: Reflexive intransitive verbs 82
 - 3.2.7 No change of diatheses: Non-reversible verbs 86
- 3.3 The middle as a basic voice: Semantics of the middle domain 91
 - 3.3.1 Body action middles 92
 - 3.3.2 The cognition middle 98
 - 3.3.3 Spontaneous events 103
 - 3.3.4 Naturally reciprocal events 105
 - 3.3.5 Verbs with no semantic ascription in Kemmer's account 106
 - 3.3.6 Pseudocopulative or attributive verbs 108
 - 3.3.7 The RM in auxiliary verbs 111
- 3.4 The middle voice in Spanish and Galician: A basic or a derived voice? 113

CHAPTER 4

Anticausative verbs in Spanish

119

- 4.1 Introduction 119
- 4.2 Different kinds of anticausative verbs 120
 - 4.2.1 Rest of the territory 124
 - 4.2.2 North-western varieties 127
- 4.3 Externally and internally caused events 129
- 4.4 The aspectual properties of the predicate 137
- 4.5 The event structure of the predicate: The acceptability of datives 142
- 4.6 Animacy of the subject 149
- 4.7 Generalised linear mixed model 153
- 4.8 Conclusions 154

CHAPTER 5

Reflexive intransitive verbs

157

- 5.1 Introduction 157
- 5.2 Analogical effects of the semantic class of the verb 158
- 5.3 Analogical effects of the animacy of the subject 162
- 5.4 Intransitive reflexive verbs of translational motion 165
- 5.5 *Morir(se)* and *caer(se)* 177
- 5.6 Corporal (physical and psychological) processes 185
- 5.7 *Quedar(se)*, *estar(se)* and *esperar(se)* 192
- 5.8 Spontaneous *pasar(se)* 204
- 5.9 More sporadic analogies 207
- 5.10 Conclusion 209

CHAPTER 6

Reflexive transitive verbs

215

- 6.1 Introduction 215
- 6.2 Indirect reflexive constructions 216
 - 6.2.1 Recipient datives 219
 - 6.2.2 Autobenefactives 220
 - 6.2.3 Possessive datives 223
 - 6.2.4 Indirect detransitivised constructions 225
- 6.3 Transitive verbs: General considerations 228
- 6.4 Inherently autobenefactive transitive verbs 234
- 6.5 Verbs of consumption 237
 - 6.5.1 Emphasis in autobenefactives and reflexive transitive verbs 242
- 6.6 Verbs of spending time 246

6.7	Verbs of cognition	251
6.7.1	<i>Saber(se)</i> ‘to know’, <i>aprender(se)</i> ‘to learn’, <i>estudiar(se)</i> ‘to study’	253
6.7.2	<i>Creer(se)</i> ‘to believe’, <i>imaginar(se)</i> ‘to imagine’	257
6.7.3	<i>Dar(se) cuenta</i> ‘to realise’: Cognition light verb structures	259
6.8	Corporal processes	260
6.9	Verbs of finding and forgetting: Lack of intentionality	261
6.10	Transitive reflexive verbs of translational motion	266
6.10.1	Transitive reflexive verbs of translational motion with a direct object that conveys a path	274
6.11	Allobenefactives	276
6.12	Conclusion	278
CHAPTER 7		
	Conclusions	281
7.1	A variationist study of the middle voice	281
7.2	The nature of the middle voice in the varieties under study	282
7.3	The RM in verbs with no change of diathesis	284
	References	289
APPENDIX 1		
	Localities interviewed with the corpus COSER	301
APPENDIX 2		
	Questionnaire videos	305
APPENDIX 3		
	Localities interviewed with the questionnaire	311
APPENDIX 4		
	List of analysed verbs and their classification	313
APPENDIX 5		
	List of verbs analysed in corpus eseuTenTen11	369
	Index	373

Acknowledgements

For this book and the research it describes, which was part of my doctoral dissertation, I owe a debt of gratitude to a great many people. First, to my supervisors, Inés Fernández-Ordóñez and Johannes Kabatek, as well as all those involved in guiding or evaluating the dissertation in some way: Leonid Kulikov, Rosario Álvarez, Cristina Sánchez López, Amaya Mendikoetxea, Andrés Enrique-Arias, Miriam Bouzoita, Enrique Pato, Bert Cornillie and Álvaro Octavio de Toledo y Huerta. To the anonymous reviewers who read the manuscript for this book and provided useful comments and ideas, as well as to Frans Hinskens, who very patiently helped me through the editorial process. Clarifying my thoughts, looking at my data in a new light, and obtaining information on languages of which I am not a native speaker was made far easier thanks to the many fruitful discussions (through all possible channels) with so many colleagues and friends: Albert Wall, Álvaro Octavio de Toledo y Huerta, Ana Estrada Arráez, Ander Egurtzegi, Araceli López Serena, Bea Martín, Cristina Bleorțu, David Gerards, Eduardo Louredo, Enrique Pato, Javier Rodríguez Molina, Ángela Di Tullio, Curdin Derungs, Etagu Alemu, Flurina Wartmann, José Luís Forneiro, Olivier Iglesias, Péter Jeszenszky, Sascha Völlmin, Tanja Samardžić and Xurxo Diz Pico. Sabine Güsewell helped me to understand what is bootstrapping and how to perform it. John Barlow thoroughly reviewed my writing. Many colleagues, friends and family members helped me to record the videos for the questionnaire, to test it, and to contact speakers. These speakers also generously shared their time with me, and provided very valuable information, although perhaps they could not see the point of it. To everyone mentioned above, I would like to express my sincere thanks. I received institutional and financial support in many phases of the process, from the Spanish Ministry of Education (FPU Grant), the Spanish Ministry of Science and Innovation (projects FFIL2009-10817 and FFI2011-12419-E), the Autonomous University of Madrid, the Romanisches Seminar and the University Research Priority Program *Language and Space* from the University of Zurich. Without these grants and support, none of the work would have been possible.

Finally, I am lucky to have great friends and the best family, who supported me (in both the English and the Spanish sense of the word) through the two long processes of writing the dissertation and then the book. *Muchísimas gracias a todos.*

Reflexive constructions

An introduction

1.1 Introduction

Why do reflexive markers (henceforth RMs, following Kemmer's 1993 terminology) exhibit a great many different functions across languages? How did these functions come about? Is there a single function that can account for all the used of RMs used in a given language? I am not the first person to ask these questions. They are, however, the starting point of this book, which will address them in a very specific way, one based on a large corpus of reflexive constructions from spoken varieties of Spanish and Galician, and taking a variationist and dialectal approach. Three preliminary questions arise here: (1) why not consider all reflexive constructions? (2) why focus on Spanish and Galician?, and (3) why a corpus-based dialectal and variationist perspective?

1. The reflexive constructions that we will focus on in this book are (a) those that are paradigmatic, i.e. they can appear in all persons, and (b) those in which the RM cannot be understood as a full referential pronoun. That is, excluded from this study are constructions that are both semantically and syntactically reflexive or reciprocal (where the RM is a referential pronoun), plus passive and impersonal constructions, where the RM can only appear in the 3rd person. The constructions that remain are often described with the term "middle" (not always accompanied by *voice*, but by *constructions* or *verbs*) (Fernández Ramírez 1986; Bobes Naves 1974a; Alcina & Blecua 1975; Babcock 1970; Benavides 2010; Monge 1955, Álvarez, Regueira & Monteagudo 1986, Álvarez & Xove 2002),¹ and are a rather heterogeneous class. Moreover, there have been several accounts of these middle voice or middle constructions from a typological perspective, with somewhat contradictory findings (Geniušienė 1987; Kemmer 1993; Kulikov 2011, Zúñiga & Kittilä 2019, Inglese forthcoming). One initial goal of this study is to provide a corpus-based typology of middle and related constructions that allow us to compare the behaviour of the RM in these constructions with previous typological accounts.

1. Other authors reserve the term for smaller subclasses: see Sánchez López (2002); NGLE (2009).

2. The focus on Spanish and Galician is justified by the fact that these closely-related neighbour languages show rather different productivity in terms of the reflexive constructions under study. Hence they are very good candidates for a contrastive analysis that aims to shed light on the evolution of the different functions of the RM. Moreover, a finer-grained analysis can be achieved by using dialectal data instead of looking solely at comparisons of the Standard languages.
3. The presence of the RM in the constructions analysed is not systematic in any of the varieties under study. That is, the same verb in the same syntactic construction might show the RM or not, being optional to some extent. A quantitative variationist account is well fitted both for the study of the linguistic factors that affect the behaviour of the RM and for the inclusion of a fine-grained dialectal perspective. Moreover, a corpus-based analysis is the only approach that fits the functional perspective on language adopted here.

The book has four main goals. First, it seeks to contribute to the typological discussion on the middle voice, in which two main competing conceptions of this voice can be found. While there is broad agreement that most of the constructions addressed in this study do indeed make up the middle voice of these languages (this holds for all Romance languages), one question must be addressed here: is the middle voice in these varieties a derived or a basic voice? (see Section 1.4). It is not difficult to find examples where the attachment of the RM is associated with a change of diathesis, but there are also cases where it is not, thus providing evidence for both possibilities. In order to answer this question, a quantitative variationist study of the alternation between verbs marked with the RM and unmarked verbs will be undertaken, so as to understand the productivity of the RM depending on the change of diathesis, on the one hand, and on the situation type, on the other. That is, it will be crucial in what follows to understand not only where the RM might appear, but also where it does not appear. In order to evaluate the two different typological proposals outlined above, it is necessary to separate form and function and to define them independently, for “[e]ven though such a distinction runs counter to the Saussurian concept of the sign it is essential in an area where markers often overlap” (Moyse-Faurie 2008: 107). Therefore, we will study the behaviour of the form (the presence or absence of the RM) depending on the function, considering two groups of notional domains: the existence of a change of diathesis, and the situation types as described by Kemmer (1993). This is done in Chapter 3, where the data suggest that the middle voice in the varieties under study is better described as a derived voice, that is, is associated with valency-reducing functions.

However, because the association with a change of diathesis is not enough to explain all the variability in the distribution of the RM, a second goal of this book is to achieve a deeper understanding of the factors that affect such variation.

Chapter 4 is devoted to a finer-grained analysis of anticausative verbs, which constitute one of the diatheses in which the RM has the highest type-frequency, despite not being completely systematic. Chapters 5 and 6 deal with situations that do not fit the description of the middle voice as a derived voice, that is, when there is no change of diathesis, be it with intransitive verbs (Chapter 5) or with transitive verbs (Chapter 6). These cases (especially the latter) are further from the middle prototype, but I will argue that they are semantically connected to the core functions of the middle voice or the reflexive diathesis. In all these cases the presence of the RM is optional to some extent, and this study aims to identify which factors (mostly intralinguistic, but spatial distribution is also taken into account) determine this variability.

Third, this study seeks to provide a detailed description and typology of the various formally reflexive constructions that cannot be considered reflexive, reciprocal, passive or impersonal in the varieties under study. While there are many descriptions and classifications of these constructions in the literature, almost none of them are based on the exhaustive analysis of corpus data (with the exception of Cartagena's (1972) excellent study, based on written Spanish). This means that several reflexive constructions have either gone unnoticed in the literature or have not received much attention. The typology of reflexive constructions provided in this book is primarily based on their syntactic behaviour (i.e. the changes of diatheses that they undergo, see Chapter 3), but also on the semantics of the verb, especially for those that undergo no change of diathesis (discussed in Chapters 5 and 6). As mentioned, such typology aims at being systematic and exhaustive (as it can possible be). Being systematic, it can also inform cross-linguistic studies about the middle voice, despite being based on only two languages. Being exhaustive, it allows to rethink previous accounts of these constructions in these languages, since it includes large amounts of data that have been previously disregarded. The current study also provides the first exhaustive description of the geographical distribution of these constructions, and is, as far as I know, the only quantitative study that investigates such a large number of reflexive constructions.

Finally, the book explores the diachronic connections between different subtypes of reflexive constructions and proposes hypotheses on the diachronic development of these subtypes, especially regarding those with no change of diathesis (discussed in Chapters 5 and 6). Specifically, I explore the hypothesis that these constructions developed through paradigmatic and syntagmatic analogy with both constructions where the RM indicates a change of diathesis and semantically reflexive constructions. While no historical data are analysed here, I will take both the statistical distribution of the same linguistic factors and the geographical overlap as hints of such diachronic connections, this being a standard method in dialectological studies (see Section 1.5).

This first chapter focusses on reviewing and explaining the main concepts of the study. A brief introduction to previous descriptions of middle constructions in the languages under study is provided in Section 1.2. The diachrony of the RM in these languages is discussed in Section 1.3. The various typological approaches to the middle voice are examined in Section 1.4. Finally, a discussion of the theoretical underpinnings of the research is given in Section 1.5.

1.2 Reflexive constructions in the varieties under study: State of the art

In this section I briefly outline the main features of reflexive constructions in Spanish, Galician and Asturian based on the previous literature. A large array of paradigmatic reflexive constructions have been described for all these languages (especially for Spanish). The RM is used in all of them to mark reflexive and reciprocal constructions, which can be further divided into direct and indirect constructions. The RM is also used in several contexts of derived intransitivity – that is, it marks intransitive verbs in which the unmarked counterpart is transitive. But the RM is also possible in contexts where no change of valency takes place: it attaches to originally intransitive and transitive verbs. Moreover, all these languages have at least some reflexive verbs for which no unmarked counterpart is found.

Given this complex situation, central to the literature here have been both typologies of the different kinds of reflexive constructions and the syntactic status of the reflexive pronoun in these constructions.² In this section, I will mainly focus on discussing these typologies, and will provide only brief notes on the morpho-syntactic status of the reflexive pronoun.

Reflexive constructions in these languages have traditionally been defined on formal grounds: they show an unstressed pronoun that agrees in person and number with the subject of the verb, i.e. a reflexive unstressed pronoun (Cartagena 1972: 21).

The reflexive unstressed pronouns in the languages under study are the following:

1. 1st sg.: *me* (Sp., Gal., Ast.,)
2. 2nd sg.: *te* (Sp., Ast., accusative in Gal.) ~ *che* (dative, Gal.)
3. 1st pl.: *nos* (Sp., Gal., Ast.) ~ *mos* (Ast.)
4. 2nd pl.: *os* (Sp.) ~ *vos* (Gal., Ast.)
5. 3rd: *se* (Sp., Gal., Ast.)

2. While the literature on reflexive constructions in Spanish is very extensive, work on reflexive constructions in Galician and Asturian is significantly less common.

From this list, *se* is the only item that is unequivocally reflexive, since *me*, *te*, *nos* and *os* are used in non-reflexive contexts (see (1)). It is for this the reason that reflexive constructions are often referred to as “constructions with *se*”.

(1) Standard Spanish

- a. Yo me vi. (reflexive *me*)
I REFL.1SG see.PST.1SG
‘I saw myself.’
- b. Tú me viste. (non-reflexive *me*)
You ACC.1SG see.PST.2SG
‘You saw me.’
- c. Él se vio. (reflexive *se*)
He REFL.3 see.PST.3SG
‘He saw himself.’
- d. *Tú se viste. (non-reflexive *se*)
You REFL.3 see.PST.2SG
‘You saw yourself.’
- e. Tú la viste. (non-reflexive 3rd person unstressed clitic)
You ACC.3SG see.PST.2SG
‘You saw her.’

Designing a typology of reflexive constructions in these languages is a complicated matter, and scholars have not yet done so satisfactorily (see a recent review in Fábregas 2021). The difference between paradigmatic (sometimes “pronominal”) and non-paradigmatic (or “syntactic”) uses of *se* is one of the most pervasive in the literature (Bobes Naves 1974a, b; Gómez Torrego 1992; Otero 2002 for Spanish; Álvarez, Regueira & Monteagudo 1986; Álvarez & Xove 2002; Freixeiro Mato 2002 for Galician, but see Fábregas 2021 for a recent critical perspective). This is a formal criterion based on the fact that some reflexive constructions admit verbs in every grammatical person (and hence every reflexive pronoun), while some others only admit 3rd person verbs (i.e. only *se*). Such a distinction separates the reflexive passive and impersonal constructions (non-paradigmatic), which have a non-explicit indefinite human agent (see (2)), from all other reflexive constructions (paradigmatic). The division is crucial to the current study, since I will not address non-paradigmatic uses here. Non-paradigmatic uses have received a great deal of attention from a variety of perspectives, featuring not only in diachronic studies (Monge 1955; Elvira 2002; Melis & Peña-Alfaro 2007) but also in quite a number of synchronic syntactic analyses, from various theoretical frameworks (see Martín Zorraquino 1979; Turley 1998; Mendikoetxea 1999b; Pedersen 2005, among many others).

- (2) a. Se ven gaviotas desde aquí.
REFL.3 see.PRS.3PL seagulls from here
'One can see seagulls from here.'
- b. Se disfruta de mucha tranquilidad desde aquí.
REFL.3 enjoy.PRS.3SG of many peace from here
'It is very peaceful here.'

Once this formal criterion has been established, a variety of syntactic and semantic criteria have been used to organise the various typologies. Syntactic criteria seem to be the most prominent in the literature, and I will address these first.

Most scholars agree on a distinction between “purely” reflexive and reciprocal verbs (see (3)) on the one hand, and all other reflexive constructions on the other (see (4)). This is justified syntactically, since only the former admit emphatic phrases, as in Sp. *a sí mismo* // Gal. *a si mesmo* // Ast. *a sí mesmu* (lit. ‘to REFL self’) in reflexive verbs, or Sp. *el uno al otro ~ entre sí* // Gal. *un ó outro ~ entre si* // Ast. *ún a otru* (lit. ‘the one to the other’ ~ ‘between REFL’) in reciprocal verbs (Alcina & Blecuá 1975; Martín Zorraquino 1979; Klein 1987; Sánchez López 2002; Otero 2002; Cano 1981 for Spanish; Cidrás Escáneo’s 1991; Álvarez & Xove’s 2002 and Freixeiro Mato’s 2002).

- (3) a. Me veo a mí misma en ese espejo.
REFL.1SG see.PRS.1SG to REFL.1SG self in that mirror
'I see myself in that mirror.'
- b. Ti vícheste reflectida no espello.
You see.PST.2SG-REFL.2SG reflect.PTCP in.the mirror
'You saw yourself reflected in the mirror.'
- (Galician, Álvarez & Xove 2002: 555)
- c. Se vieron el uno a otro desde lejos.
REFL.3 see.PST.3PL the one to other from far
'They saw each other from far away.'
- (4) a. Me levanto (*a mí misma) muy temprano.
REFL.1SG wake.up.PRS to REFL.1SG self very early
'I wake up very early.'
- b. Me comería una tortilla (*a mí misma)
REFL.1SG eat.COND.1SG a omelette to REFL.1SG self
ahora mismo.
now self
'I would eat an omelette right now.'
- c. ¡No te enfades (*a ti mismo) por
no REFL.2SG get.mad.PRS.2SG to REFL.2SG self for
esa tontería!
that silly.thing
'Don't get mad for such a silly thing!'

Accepting emphatic phrases with a stressed pronoun is a property that these constructions share with other unstressed pronouns (see (5)), which justifies proposing a different status of the RM in each type. Accordingly, in “purely” reflexive and reciprocal verbs the RM is a pronoun, comparable to other unstressed pronouns such as *lo* or *la*, that is, it has referential and argumental properties. In the remaining constructions, however, it is deemed to be an intransitivising device or a verbal morpheme (Martín Zorraquino 1979; Klein 1987; Gómez Torrego 1992; Sánchez López 2002; Otero 2002; Arús 2006; NGLE 2009; Benavides 2010). Some authors, however, insist on attributing it pronominal status at least in some subtypes of the latter category (Alarcos 1970; Bello 1981 [1847]). Less frequently, some authors have proposed that the RM is no longer a pronoun at all, as Cartagena (1972) argues for *se* (but not *me*, *te*, *nos*, *os*).

- (5) Lo vi a él
 ACC.3SG see.PST.1SG to OBL.3SG
 ‘I saw him.’

The distinction between purely reflexive and reciprocal verbs on the one hand, and the remaining constructions with *se* on the other, is also crucial in the present study, since I will only focus on the latter. Hence, in the following description of the state of the art of purely reflexive and reciprocal constructions, I will limit myself to those aspects relevant for the analysis of other reflexive constructions, such as the distinction between direct and indirect constructions.

1.2.1 Purely reflexive and reciprocal verbs

Subdivisions within purely reflexive and reciprocal constructions are more controversial. A distinction between direct constructions (where the RM is the direct object, (6)) and indirect ones (where the RM is the indirect object, (7)) in the purely reflexive subclass is acknowledged in several works on Spanish (Fernández Ramírez 1986; RAE 1973; Alarcos 1970, 1994; Bello 1981 [1847]; Martín Cid 2004; Martín Zorraquino 1979; Gómez Torrego 1992; Cartagena 1972; Bobes Naves 1974a; Cartagena 1972; Cano 1981), but largely ignored in many others, especially in more recent work (Klein 1987, Ávila López Pedraza 2009, Benavides 2010; Otero 2002; NGLE 2009). This is probably due to the lack of interest in purely reflexive verbs as a class in the literature. Where the distinction is noted, the more basic status of direct reflexives as compared to indirect reflexives is often recognised. Fernández Ramírez (1986), for example, terms direct reflexives “proper reflexives” (*reflexivos propios*), as opposed to reflexive verbs with a direct object (*verbos reflexivos con complemento directo*). The distinction between direct and indirect constructions in reciprocal constructions is virtually never mentioned in the literature on Spanish,

with a few exceptions, such as Alcina & Bleca (1975); Martín Zorraquino (1979) and Gómez Torrego (1992).

- (6) Se miraba detenidamente en el espejo.
REFL.3 look.PST.3SG slowly in the mirror
'S/he looked at himself/herself slowly in the mirror.'
- (7) Se enviaba flores fingiendo ser una admiradora.
REFL.3 send.PST.3SG flowers pretending be.INF a admirer
'She sent herself flowers pretending to be an admirer.'

However, the distinction is more prominent in the literature on Galician and Asturian because these languages differ greatly from Spanish in this respect, showing a lower productivity of the indirect reflexive pronoun. For instance, Galician lacks indirect reflexive pronouns almost completely (Lantes 1980, Álvarez, Regueira & Monteagudo 1986; Cidrás Escáneo 1991; Álvarez & Xove 2002; Freixeiro Mato 2002), as illustrated by the contrast in (8).

- (8) (Galician, Álvarez, Regueira & Monteagudo 1986: 179)
- a. Lavei(*me) as mans.
wash.PST.1SG-(REFL.1SG) the hands
'I washed my hands.'
- b. Queimei(*me) un dedo.
burn.PST.1SG-(REFL.1SG) a finger
'I burnt my finger.'

A reflexive indirect pronoun may only occur in Galician when there is an unequivocal need to specify that the subject and the recipient are coreferential, as in (9) (Cidrás Escáneo 1991; Freixeiro Mato 2002).

- (9) (Galician, Freixeiro Mato 2002)
- a. O médico diagnosticouse unha hepatitis.
the doctor diagnose.PST.3SG-REFL.3 a hepatitis
'The doctor diagnosed himself/herself as having hepatitis.'
- b. A costureira fíxose unha saia para cumpreanos
a seamstress make.PST.3SG-REFL.3 a skirt for birthday
da filla.
of.the daughter
'The seamstress made herself a skirt for her daughter's birthday.'

It has also been noted that indirect reflexive pronouns are frequent with speech verbs in Galician (see (10)) (Álvarez, Regueira & Monteagudo 1986, Álvarez & Xove 2002).

- (10) Pregúntome a miúdo que será dela.
 ask.PRS.1SG-REFL.1SG to often what be.FUT.3SG of.OBL.3SG
 'I often ask myself what she is up to.'
 (Galician, Álvarez, Regueira & Monteagudo 1986: 179)

Interestingly, the tendency away from reflexive datives in Galician is not found in reciprocal constructions, where the RM shows no restrictions (Álvarez, Regueira & Monteagudo 1986, Álvarez & Xove 2002; Freixeiro Mato 2002):

- (11) Galician
 a. Déronse moitas labazadas.
 give.PST.3PL-REFL.3 many slaps
 'The slap each other several times.' (Freixeiro Mato 2002: 155)
 b. Ana e María déronse bicos.
 Ana and María give.PST.3PL-REFL.3 kisses
 'Ana and María kissed each other.' (Álvarez & Xove 2002: 556)

Asturian shows intermediate behaviour between Spanish and Galician, since both the indirect reflexive and the indirect reciprocal are possible, but while in the latter the presence of the RM is a requirement to convey such a meaning, in the former it is only a possibility (D'Andrés 1994; García García 1989). According to D'Andrés (1994), the RM only appears in the indirect reflexive when coreferentiality with the dative is not interpreted by default (see (12a)), or to mark contrastive emphasis (see (12b)), and otherwise is absent (see (12c)). Suárez Fernández (1990), however, claims that the RM cannot appear in indirect constructions in Asturian.

- (12) (Asturian, D'Andrés 1994)
 a. Tú regalastite'l coche.
 you give.as.a.gift.PST.2SG-REFL.2SG-the car
 'You gave yourself the car as a gift.'
 b. –¿Llavóte la cabeza'l peluqueru? –Non, llavéme
 wash.PST.3SG the head-the hairdresser no wash.PST.1SG-REFL.1SG
 la cabeza yo mesmu.
 the head NOM.1SG self
 '– Did the hairdresser wash your head? – No, I washed my head myself.'
 c. Tú llaves la cabeza.
 you wash.PRS.2SG the head
 'You wash your head.'

The lack of attention that has been paid to the distinction between direct and indirect constructions in Spanish is rather surprising, since it has been said that it lies at the heart of the diachronic connections between the different types of *se* constructions (Bobes Naves 1974a). Such scant interest may be due to the dominating

synchronic perspective of most analyses of reflexive constructions. The difference between direct and indirect constructions will play a prominent role in the current study, since the data show clear differences between intransitive and transitive reflexive verbs.

Much more complex are the subdivisions of the remaining paradigmatic reflexive constructions, which I will now turn to. As noted above, in these constructions the RM no longer behaves like a pronoun, since they do not admit emphatic pronominal phrases (see Otero 2002; Martín Zorraquino 1979; Gómez Torrego 1992; Maldonado 1999; Cano 1981 among others). The RM is thus analysed as a component of the verb itself (Gómez Torrego 1992; Otero 2002; Arús 2006). This is a rather heterogenous class which not all authors group together (see, for example, Sánchez López 2002) and is very frequently defined negatively, including any reflexive paradigmatic construction that does not fall under the purely reflexive and reciprocal subgroup (Arús 2006; RAE 1973), or indeed is sometimes not defined at all (Otero 2002). As we have said, many authors refer to these constructions using the term “middle” (Fernández Ramírez 1986; Bobes Naves 1974a; Alcina & Blecua 1975; Babcock 1970; Benavides 2010; Monge 1955; Álvarez, Regueira & Monteagudo 1986, Álvarez & Xove 2002). In what follows, I review the most relevant criteria used in different typologies, towards providing a broad picture of the various parameters that have been taken into account in their description, and beginning with syntactic criteria.

1.2.2 The relationship of the reflexive verb and its unmarked counterpart

Most typologies take into account the different syntactic relationships that can be found between the reflexive verb and its unmarked (non-reflexive) counterpart. There are at least three possibilities: (a) the transitivity properties of the reflexive verb change as compared to its unmarked counterpart; (b) the transitivity properties of the reflexive verb remain the same as in its unmarked counterpart, and (c) there is no unmarked counterpart (Cartagena 1972; Fernández Ramírez 1986; Klein 1987; Gómez Torrego 1992; Lantes 1980, Álvarez, Regueira & Monteagudo 1986, and, partially, Alarcos 1970; Martín Zorraquino 1979). I discuss these in the remainder of this section.

The RM is often involved in transitive/intransitive alternations, according to which the reflexive form is intransitive and has an unmarked transitive correlate (Hernández 1966; Babcock 1970; Cartagena 1972; Alcina & Blecua 1975; Martín Zorraquino 1979; Fernández Ramírez 1986; Klein 1987; Gómez Torrego 1992; Sánchez López 2002; NGLE 2009 for Spanish; Álvarez, Regueira & Monteagudo 1986, Álvarez & Xove 2002; Freixeiro Mato 2002, for Galician; D’Andrés 1994 for Asturian). For this reason some authors have proposed that the general function of

the RM is to intransitivise the verb to which it attaches (Cartagena 1972; Babcock 1970; Bogard 2006).³

Some accounts dissect this intransitivising function in more specific subclasses. One of the most prominent changes in intransitivity associated with the RM is the so-called *causative alternation* (Mendikoetxea 1999a, Sánchez López 2002; Bogard 2006; NGLE 2009: 34.6, Vivanco 2016): the unmarked verb is a transitive verb that typically takes an agent (as the subject) and a patient (as the direct object), while the reflexive verb is an unaccusative intransitive verb with a patient subject:

- (13) a. María rompió el vaso.
 María break.PST.3SG the glass
 ‘María broke the glass.’
 b. El vaso se rompió.
 the glass REFL.3 break.PST.3SG
 ‘The glass broke.’

For Galician, Álvarez, Regueira & Monteagudo (1986) and Álvarez & Xove (2002) describe (a subset of) these verbs on semantic terms, noting that they imply a transformation or change of the subject and that they take experiencer subjects (even when the subject is inanimate). They note that this construction does not need to be marked by the RM in Galician.

- (14) (Álvarez & Xove 2002: 557).
 a. A roupa secou (co sol)
 the clothes.SG dry.PST.3SG with.the sun
 ‘The clothes dried (with the sun).’
 b. O prato rompeu(se)
 the plate break.PST.3SG-REFL.3
 ‘The plate broke.’

Similarly, D’Andrés (1994) notes that some of these verbs in Asturian need the RM in the intransitive pattern (*afogar* ‘to drown’, *llevar* ‘to stand up’, *echar* ‘to lie down’, *volver* ‘to turn, to go back’, *mancar* ‘to hurt’, *poner* ‘to put’, *sentar* ‘to sit down’, *derrodiyar* ‘to kneel down’), whereas several others do not. The latter group

3. Cartagena (1972) actually observes that the RM is often used in transitive sentences and hence attributes a broader function to the RM, which he calls a marker of “non-obliquity” (*no oblicuidad*). By this, he means that the RM indicates that the action is not transferred to a participant different from the subject, which most often results in the intransitivisation of the verb, but not exclusively. Bogard (2006) considers it to be an argument-reducing device (*clausurador argumental*), a broader syntactic function that enables him to classify together not only intransitive reflexive verbs with an unmarked transitive counterpart but also the reflexive and passive function. He considers that the reduction of the transitivity is a constant in reflexive constructions.

can be further subdivided in those verbs that admit the RM, although being more common without it (*cansar* ‘to rest’, *nublar* ‘to get cloudy’, *fartucar* ‘to get sick, tired’, *torcer* ‘to bend’), and those that never take the RM (*callar* ‘to be quiet, to shut up’, *plasmear* ‘to stun’, *avieyar* ‘to get old’).

Another possible intransitivising alternation consists of an unmarked transitive verb that also takes two participants: a cause (in the subject position) and an experiencer (in the direct object position). In the reflexive counterpart, the experiencer occupies the subject position and the cause argument may either be coded as a prepositional object or remain unexpressed (see (15)) (Cartagena 1972; Alcina & Blecua 1975).⁴

- (15) a. Verte alegró a María.
 see.INF-ACC.2SG make.glad.PST.3SG to María
 ‘Seeing you made María happy.’
 b. María se alegró (de verte)
 María REFL.3 make.glad.PST.3SG of see.INF-ACC.2SG
 ‘María was happy (to see you).’

Another intransitivising alternation that has been noted, especially in the most recent literature, is the so-called antipassive alternation (Alcina & Blecua 1975; Bogard 1999a, 1999b, 2006; Sánchez López 2002; NGLE 2009 for Spanish; Álvarez, Regueira & Monteagudo 1986, Álvarez & Xove 2002; Freixeiro Mato 2002 for Galician; D’Andrés 1994 for Asturian). Here the reflexive verb can preserve the two participants that it has in the unmarked transitive counterpart, but the second one (typically a patient) is no longer a direct object (as in the unmarked counterpart) but a prepositional object (see (16)). This prepositional object can be omitted.

- (16) a. María decidió ir.
 María decide.PST.3SG go.INF
 ‘María decided to go.’

4. The fact that these verbs can preserve a prepositional object has gone unnoticed in several works, which have nevertheless made a distinction between anticausatives and these verbs, mostly based on semantic reasons. Hernández (1966) distinguishes between “physically internal reflex intrinsic *se*” (*se intrínseco reflejo interior físico*, i.e. anticausatives) and “psychically internal reflex intrinsic *se*” (*se intrínseco reflejo interior psíquico*, such as *alegrarse*). That is, the difference is mainly based on the semantic class of the verbs that appear in both types, since the latter mostly consists of emotion verbs. Fernández Ramírez (1986), for instance, talks about the “*alegrarse*-type” (*‘be-happy-type’*), as opposed to other anticausatives with human subjects with no “active reading” – this difference is mainly based on the different thematic roles of the subject in each reflexive type. Sánchez López (2002) also follows this criterion, distinguishing between anticausative verbs with a subject that is a notional object, and anticausative verbs with an experiencer subject (see also Martín Zorraquino 1979).

- b. María se decidió a ir.
 María REFL.3 decide.PST.3SG to go.INF
 ‘María decided to go.’
- c. María se decidió.
 María REFL.3 decide.PST.3SG
 ‘María made a decision.’

Lastly, very few studies have analysed cases like (17) separately, where the intransitive reflexive preserves the agent/experiencer in the subject position, while the patient object is demoted. Cartagena (1972: 73) considers them a subtype within the intransitivized class, attributing them a *totum pro parte* reading.

- (17) a. Juan arriesgó su vida.
 Juan risk.PST.3SG his life
 ‘Juan risked his life.’
- b. Juan se arriesgó.
 Juan REFL.3 risk.PST.3SG
 ‘Juan took a risk.’

As regards reflexive verbs whose transitivity properties remain the same as in the unmarked counterpart (see (18)), it is common to make a further distinction between intransitive and transitive verbs (Martín Zorraquino 1979; NGLE 2009). Transitive verbs like *comerse* ‘to eat’, *tomarse* ‘to have, to take’ and *creerse* ‘to believe’ were typically grouped with the indirect reflexive in earlier work (Hernández 1966; Cartagena 1972; Bobes Naves 1974a; b; RAE 1973; Fernández Ramírez 1986; Bello 1981 [1847]; Cano 1981; Gómez Torrego 1992) on the basis that they all have a direct object. The reflexive pronoun in those verbs has usually been considered an ethical or interest dative.

- (18) a. Juan se bebió un vaso de vino.
 Juan REFL.3 drink.PST.3SG a glass of wine
 ‘Juan drank a glass of wine.’ (Otero 1999)
- b. Te comías las rosas.
 REFL.2SG eat.PST.2SG the roses
 ‘You ate the roses.’
 (Juan Ramón Jiménez, *II Antol. poética*, apud Fernández Ramírez 1986)

These cases, where the RM appears with transitive verbs, are claimed not to exist in Galician (Freixeiro Mato 2002; Instituto de la Lengua Gallega 1975; Lantes 1980) or in Asturian (D’Andrés 1994). According to these authors, examples such as those in (19) and (20) are not grammatical in these languages.

(19) Galician

- a. *Bebinme unha cervexa
 drink.PST.1SG-REFL.1SG a beer
 'I drank a beer.' (Freixeiro Mato 2002)
- b. *Merecese uns azoutes
 deserve.PRS.3SG-REFL.3 some spankings
 'S/he deserve a spanking.' (Álvarez, Regueira & Monteagudo 1986)

(20) (Asturian, D'Andrés 1994)

- a. *Buscáivos un trabayu que dea dineru.
 look.for.IMP.2PL-REFL.2PL a job that give.SBJV.3SG money
 'Look for a job that pays well.'
- b. *L'escolín tráxose los llibros nuna cartera.
 The-pupil bring.PST.3SG-REFL.3 the books in.a rucksack
 'The pupil brought the books in a rucksack.'

The existence of intransitive verbs that may or may not take the RM in Spanish (see (21)) has been observed since early work in the field (Hernández 1966; Cartagena 1972; RAE 1973; Alcina & Blecua 1975; Fernández Ramírez 1986). Many authors have noted that, despite the fact that the reflexive counterpart has the same argument structure as the unmarked counterpart, they may differ in meaning (Alarcos 1994; Cartagena 1972; Martín Zorraquino 1979 among others).

(21) (Alarcos 1994)

- a. Voy a casa. ~ Me voy a casa.
 go.PRS.1SG to home ~ REFL.1SG go.PRS.1SG to home
 'I'm going home' ~ 'I'm leaving to go home.'
- b. Está quieto. ~ Se está quieto.
 be.PRS.3SG still ~ REFL.3 be.PRS.3SG still
 'S/he is still' ~ 'S/he stays still.'

Reflexive intransitive verbs are not very common in Galician and they are restricted to a few movement and stative verbs (*ir(se)* 'to go', *marchar(se)* 'to leave', *estar(se)* 'to be', *quedar(se)* 'to stay') (Álvarez, Regueira & Monteagudo 1986; Freixeiro Mato 2002). Galician does not admit the RM with certain intransitive verbs that can be reflexive in Spanish, such as *adormecer* 'to fall asleep' or *morir* 'to die' (Freixeiro Mato 2002).

A similar observation applies to Asturian. The RM does not attach to originally intransitive verbs, except for a very few cases (*dir* 'to go', *escapar* 'to escape', *quedar* 'to stay', *valir* 'to be worth' and *llegar* 'to arrive'), and even here the RM is rare (Suárez Fernández 1990; D'Andrés 1994). García García (1989) goes further, claiming that the RM can never appear with intransitive verbs in Asturian.

Some studies on Spanish group all these verbs together (whether intransitive or transitive). For instance, Klein (1987) considers them all under the “stylistic *se*” subtype, on the grounds that they are not shared by all speakers, show dialectal variation, and have a stylistic rather than a grammatical effect. Grouping intransitive and transitive pronominal verbs in one subclass seems to be more frequent in recent work, as a consequence of the recent tendency to analyse the reflexive pronoun in these constructions as an aspectual or perfective marker (De Miguel & Fernández Lagunilla 2000; Otero 2002; Bogard 2006). This interpretation will be discussed in more detail below in Chapters 5 and 6 (see also De Benito Moreno 2015, 2021).

Finally, we turn to reflexive verbs with no unmarked counterpart (*quejarse* ‘to complain’, *arrepentirse* ‘to regret’, *enterarse* ‘to learn’, etc.), which have received several names: inherently pronominal verbs; reflexes (*reflejos*) (Bello 1981 [1847]); pronominal (Bello 1981 [1847]; NGLE 2009; Gómez Torrego 1992 for Spanish; Álvarez & Xove 2002 for Galician) or purely pronominal verbs (Sánchez López 2002). The most prevalent opinion is that the RM forms part of the verb and cannot be separated from it (Alarcos 1970; Cartagena 1972; Klein 1987; Gómez Torrego 1992; Otero 2002, Ávila López-Pedraza 2009). Lantes (1980) points out that such verbs often take an oblique complement in Galician, and D’Andrés (1994) makes a similar observation for Asturian

Reflexive verbs whose unmarked counterpart has a very different meaning from the one showed by the reflexive form (*acordar* ‘agree’ vs. *acordarse* ‘to remember’, *portar* ‘to take’ vs. *portarse* ‘to behave’) have sometimes been grouped in this category (Alcina & Blecua 1975; Alarcos 1994; Sánchez López 2002; NGLE 2009). That is, when the differences of meaning between both verbs are too large, it is often assumed that they are not related to each other (at least from a synchronic point of view). It is commonly acknowledged that there may be some difference of meaning between the unmarked verb and its reflexive counterpart, but it is also commonly agreed that these differences are not only small, but systematic. Authors like Klein (1987) and Cartagena (1972) also acknowledge the similarities between these two classes (that is, *acordarse*, *portarse* on the one hand and *arrepentirse*, *quejarse* on the other), since they both consider the RM to be a lexical component in these cases. However, they prefer to categorise them in two different subtypes, making clear the relevance of having an unmarked counterpart.

1.2.3 Semantic criteria in typologies of paradigmatic reflexive constructions

Thus far, I have focussed on the syntactic criteria that have been used to establish different subtypes of reflexive verbs, be it the grammatical status of the RM or the relationship between the reflexive verb and its unmarked counterpart. But semantic

criteria have also been used in the classification of reflexive constructions, often in combination with syntactic ones. We have already seen the delicate interweaving of syntactic and semantic criteria at work in the differentiation of different “intransitivised” reflexive constructions (such as the type of verb or the thematic roles its participants take, as in *levantarse* ‘to wake up, to stand up’ vs. *alegrarse* ‘to be glad’). Two main semantic criteria are found in global analyses of reflexive constructions: the semantics of the verb, and the animacy of the subject.

The semantic class of the verb plays a role in the classification and description of reflexive constructions in some accounts. Fernández Ramírez (1986) notes that anticausative verbs with human subjects are very often verbs of movement. Hernández’s (1966) account distinguishes four types of what he calls “intrinsic se”, depending on the semantic class of the verb: physically internal reflex (*reflejo interior físico*), psychically internal reflex (*reflejo interior psíquico*), dynamic reflex (*reflejo dinámico*, i.e. verbs of movement) and reflexes with inceptive aspect (*de aspecto inceptivo*, i.e. *dormirse* ‘to fall asleep’, *morirse* ‘to die’). Cano (1981) establishes a difference between verbs of emotion, of movement, and those with reciprocal meaning (*juntarse* ‘to get together’, *reunirse* ‘to meet’, *acumularse* ‘to accumulate’).

Alcina & Blecua (1975) actually talk about “movement and stative verbs in the intransitive construction” when discussing intransitive verbs that can take the RM with no change of valency. Similarly, Martín Cid (2004) seems only to include verbs of movement in this subclass. Martín Zorraquino (1979) does not use the semantics of the verb as a classificatory criterion but notes that the possibility of the absence or presence of the RM with no (significant) change of meaning, nor a change of diathesis, is very common among verbs of movement.

Finally, Cartagena (1972: 66–71) provides a very detailed classification of what he calls “non-oblique pronominal constructions” (i.e. intransitivised reflexive verbs) in semantic terms, with five main subgroups: psychological verbs, verbs of movement, verbs of appearance, verbs largely of disappearance, and verbs of change of state. Because of the quantitative importance of verbs of movement and verbs of change of state, he makes further subdivisions within these two types, taking into account the meaning of the unmarked variant.⁵

5. Within movement verbs, Cartagena (1972) distinguishes verbs that have the meaning of providing someone with a specific position or posture (*acostar* ‘to lie down’); verbs that mean to move something or someone (*mover* ‘to move’); verbs that mean to interrupt the movement of something or someone movement (*detener* ‘to stop’), and verbs that mean to change the posture or body position of something or someone (*agachar* ‘to bend down’). Within verbs of change of state, he groups separately those verbs that indicate general transformations (*modificarse* ‘to change’); verbs that have the meaning of acquiring a specific property (*enfriarse* ‘to get cold’); verbs that indicate the altered composition or consistency of an object, which typically require an inanimate subject (*agujerearse* ‘to pierce’); and verbs that refer to the alteration of the degree of resistance of someone or something, typically referring to animate subjects (*cansarse* ‘to rest’).

Besides the semantic class of the verb, another very common semantic consideration made in the literature on reflexive verbs is the animacy of the subject. Fernández Ramírez (1986) distinguishes two classes of anticausative verbs – note that he does not use this term – according to whether their subject is human (the subclass “transitive verbs neutralization”) or inanimate (“reflexive verbs with inanimate subject”). The very same distinction is fundamental to Cano’s (1981) typology of reflexive constructions. Similarly, Benavides (2010) classifies anticausative with human subjects in the reflexive *se* class, while claiming that anticausatives with inanimate subjects belong in the middle *se* class. Cartagena (1972), although without using subject’s animacy as a classificatory criterion, notes that some reflexive subtypes tend to have either human or animate subjects, while others are more inclined to take inanimate subjects.

1.2.4 “Difficult cases”

Two small subgroups of reflexive constructions, mentioned only by a few authors, should also be mentioned here. One type that has proved difficult to classify within most typologies is the so-called *mediopassive* or *quasipassive* (Babcock 1970; NGLE 2009: 41.110–q, see (22)). These have sometimes been grouped with anticausatives (Benavides 2010), but have also been considered a subtype of the passive reflexive (Mendikoetxea 1999b; NGLE 2009). Following this latter interpretation, I omit them from the present study.

- (22) a. Las luces reflectantes se ven fácilmente.
 the lights reflector REFL.3 see.PRS.3PL easily
 ‘Reflector lights are easily visible.’ (Mendikoetxea 1999b: 1662)
- b. Juan se espanta fácilmente.
 Juan REFL.3 scare.PRS.3SG easily
 ‘Juan scares easily.’ (Babcock 1970: 40)

On the other hand, examples like (23), which show a non-reflexive dative encoding an affected participant, have also been considered a special subtype of reflexive constructions. Martín Zorraquino (1979) considers them a special subgroup of anticausatives with inanimate subjects. Sánchez López (2002) considers them to be a special subgroup of *middle se* and points out that in some cases the dative is obligatory (*olvidársele algo* ‘to forget something’), as does Cano (1981). Gómez Torrego (1992) places them in the subgroup of “doubly pronominal verbs” (*verbos doblemente pronominales*) within the non-paradigmatic class, although he notes that the reflexive is a verbal morpheme (as in anticausatives). This subgroup includes verbs like *antojarse* ‘to feel like doing something’, *ocurrirse* ‘to have an idea’ and *olvidarse* ‘to forget’ (Gómez Torrego 1992). The NGLE (2009: 41.13b) also identifies them

as a specific subtype, although it considers them analogous to unaccusative verbs with an argumental experiencer (NGLE 2009: 35.2ñ), as I do in this study (thus including them in the analysis).

- (23) a. Se le rompieron los platos
REFL.3 DAT.3SG break.PST.3PL the plates
'He broke the plates accidentally.' (Martín Zorraquino 1979: 115)
- b. A Pepe se le ha estropeado el coche por
to Pepe REFL.3 DAT.3SG have.PRS.3SG break.PTCP the car for
segunda vez
second time
'Pepe's car broke for the second time.' (Sánchez López 2002: 103)

1.2.5 Summary

In this section, I have offered a comprehensive overview of the different classificatory criteria that have been used in the general literature on reflexive constructions in these languages. The main criteria that are most commonly found are based on either formal grounds (paradigmatic and non-paradigmatic uses), syntactic ones (the grammatical status of the RM, the relationship of the reflexive verb with its unmarked counterpart), or on semantic grounds (the animacy of the subject and the verb's semantic class), all of which are relevant for the present study.

From the literature review it is clear that reflexive constructions in Galician and Asturian show a lower productivity as compared to Spanish. Lantes (1980, 1986) notes a lower productivity of these constructions in the varieties spoken in the north-western area of the Iberian Peninsula (that is, Asturian, Portuguese and Galician). He claims that the main differences between Spanish and Galician lie in (1) the impossibility of indirect reflexive constructions, (2) the rare presence of the RM with "occasional pronominal verbs" (*pronominais ocasionais*) like *ir*, and (3) the resistance of Galician to clitic clusters when one of them is reflexive.

It should be noted, however, that these claims are not free of language purism, which may impact on their accuracy.⁶ This becomes clear if we consider quotes such as the following, where the agrammaticality of indirect reflexives is claimed despite the fact that they are actually documented in the language: "Despite of being ungrammatical, indirect reflexives can be found in literary texts: *O de Tamames lavouse os pés no obrador* (lit. 'The one from Tamames washed himself the feet at the workshop' [Cunqueiro])" (Lantes 1980: 37–8), my translation. In any case, it is very

6. Consider one of the aims of his study: "The influence of Castilian, especially in the literary language, is of outstanding importance, which is why it is urgent to clarify and separate what is Galician from what is not" (Lantes 1980: 1, my translation).

relevant to the issue to observe that speakers of Galician themselves acknowledge the use of the RM as a potential source of interference with Spanish. Kabatek (2000) observes that, although both (24a) and (24b) are well-documented forms in traditional spoken Galician, speakers who feel unsure about their use of Galician often state that they tend to consciously avoid the use of the reflexive form, considering it to be Spanish interference.

- (24) a. vou sentar
 go.PRS.1SG sit
 'I am going to sit down.'
- b. voume sentar
 go.PRS.1SG-REFL.1SG sit
 'I am going to sit down.'

The complex picture depicted in this section can be explained on historical grounds. That is, the broad range of constructions where the RM appears can be seen as the product of a historical evolution of this marker, which has been constantly expanding its functional range since Latin. Thus, a satisfying understanding of the synchronic situation here requires a solid understanding of this historical evolution and vice versa; that is, a detailed analysis of the synchronic situation, such as the one presented in this book, can inform our interpretation of the historical developments, which is indeed one of the aims of this book. In the following section I will briefly review the literature on the diachronic development of reflexive constructions in the languages under study.

1.3 The diachrony of reflexive constructions

The historical development of reflexive constructions in Spanish, Galician and Asturian has not received as much attention as their synchronic behaviour. Notwithstanding, some studies have been devoted to the diachronic origin of these constructions and some others, despite taking a fundamentally synchronic perspective, have hinted at possible diachronic connections.

Although this study is mainly concerned with the synchronic situation of the varieties under examination, the (plausible) historical evolution of formally reflexive constructions is crucial for our goal, for at least two reasons. First, the general theory of diffusion of language change enables us to correlate different dialectal stages with different diachronic states, which is why the data used in this study can be used as indirect evidence of the direction of diachronic developments (see Section 1.5.1).⁷

7. This idea also underlies García García's (1989) claim that the lower frequency of *se* in Asturian indicates that this language is at an earlier stage of development than Spanish.

Second, the description and analysis of a situation of variation (or a diasystem) necessarily rests on specific hypotheses as to how this situation came about, that is, on the different stages of development (see Section 1.5.1). In this section I will offer a brief description of the main historical developments concerning reflexive constructions as described in the literature.

The general picture, on which most scholars agree, divides the extension of the reflexive pronoun into four stages. From its originally reflexive meaning, the pronoun would have developed middle or intransitivising functions and would later have taken on the passive function. The development of the impersonal function is the most recent (Lapesa 2000; Melis & Peña-Alfaro 2007). Only the passive and the impersonal functions are Romance developments, while the other functions of the RM are already found in Latin (Monge 1955). The development of new functions of the RM in Romance is typically depicted as a competition of the reflexive pronouns with the middle/passive Latin morphology, in which the RM progressively adopts more functions (Monge 1955: 11–12).

Enrique-Arias & Bouzouita (2013), in a study of the frequency of the clitics in Spanish from a diachronic perspective, note that the reflexive clitic is the one whose frequency increases the most (as opposed to non-reflexive dative and accusative clitics). They interpret this situation as a consequence of its increasing level of grammaticalization. On the one hand, they show that the frequency of each function of the reflexive has increased from the Middle Ages to our times. On the other hand, they also show that the total number of functions of the RM increases, since in Medieval Spanish the reflexive passive was quite rare and the reflexive impersonal is not documented, while these two functions are relatively frequent in more modern stages.

This general picture coincides with the implicational hierarchies found in typological descriptions. Geniušienė (1987) proposes the following hierarchy for the development of the “objective recessive diatheses” from reflexive systems:

- (25) Decausative > passive > impersonal (Geniušienė 1987)

Similarly, Heine (2000) proposes the grammaticalization chain in (26) in his study of reflexive and reciprocal markers in African languages. The order of the chain is based on their nominal characteristics (associated with the display of nominal morphology), which decrease from left to right, and their affixation characteristics – associated with less phonetic substance –, which increase from left to right. The point of departure for the Romance reflexive would be the third step of this hierarchy, since *se* already had reflexive meaning in Latin.

- (26) Nominal meaning > emphatic > reflexive > reciprocal > middle > passive
(Heine 2000)

While this picture is largely agreed upon, we have to admit that it is not very informative for our purposes, since we are not concerned with reflexive, reciprocal, passive or impersonal constructions here, but with middle constructions, including the decausative diathesis mentioned by Geniušienė (1987). Thus, many questions remain unanswered. Are there different phases in the development of this function? Is there an expected diachronic order in the development of the different constructions subsumed under the middle function? Is there a difference between the evolution of direct and indirect reflexive constructions? Is the reciprocal a necessary intermediate step between the reflexive and the middle/intransitivising functions? What is the origin of the RM that attaches to intransitive and transitive verbs with no change of valency (and which is therefore not decausative)?

Some of these questions will concern us throughout this study. In the following section, I will consider what has been said about the overall development of reflexive constructions in Spanish (in Section 1.3.1), the specific developments of the indirect reflexive (Section 1.3.2), the reciprocal function (Section 1.3.3), the development of middle functions (Section 1.3.4), and the spread of the RM to contexts where there is no change of valency (Section 1.3.5).

1.3.1 Overall development

Most of the reflexive constructions documented in Modern Spanish already existed in Old Spanish (Lapesa 2000). The texts collected in *Orígenes del español* (i.e. texts until the eleventh century) show examples of the direct reflexive construction (already quite lexicalised, close to middle contexts), the indirect reflexive, anticausatives with inanimate subjects, intransitive and transitive verbs, and even a couple of cases of possible passive constructions (Martín Zorraquino 2002).

All modern reflexive constructions were found as early as the thirteenth century, with the sole exception of the impersonal reflexive (Enrique Arias & Bouzouita (2013). Crucially, the productivity of all the constructions considered increases gradually over time, especially those functions where the RM indicates a change of diathesis. This increase in productivity does not only refer to a higher number of examples, but also to a greater number of verbs found in some categories – especially in the anticausative, the antipassive, and what they call aspectual categories, where no change of valency is found Enrique-Arias & Bouzouita (2013).⁸

8. Bogard's (2006) claim that there is no substantial increase in the productivity of the RM arises from a flawed quantitative method. He compares the occurrences of each reflexive construction to the total of reflexive constructions in a given period. Such a calculation is inadequate, since the different reflexive constructions apply to different contexts. Bogard even claims that the purely reflexive and reciprocal function have decreased, but this observation is motivated only by steeply

1.3.2 Indirect reflexive constructions

Work on the development of the indirect reflexive constructions from Latin to Spanish is almost non-existent. Most authors seem to assume that these functions are inherited directly from Latin. However, Luque Castro (2007) does notice that Latin indirect reflexive constructions are notably less frequent than in Spanish. He proposes that indirect reflexive constructions with verbs of dressing in fact derive from the reanalysis of Latin constructions with an accusative and an ablative object.

He observes that the most common structure for expressing this notion in Latin is the one in (27). In such a construction, the reflexive pronoun *se* is an accusative pronoun, referring to the affected participant of the verb action. On the other hand, the garment (*tunica*) is in the ablative case and shows an instrumental reading. According to Luque Castro (2007), the loss of the case distinction in Latin would have led to a structure with two participants in the same case. This would have caused the reanalysis of the reflexive pronoun as the indirect object and of the instrument noun phrase as the direct object.

- (27) Expoliavit se Ionathas tunica
 remove.PST.3SG REFL.ACC.3 Ionathas.NOM robe.ABL
 ‘Jonathan took off his robe.’ (Samuelis, I, 18: 4, apud Luque Castro 2007)
- (28) Juan se quitó la tunica
 Juan REFL.3 remove.PST.3SG the robe
 ‘Juan took off his robe.’

While this reanalysis could account for sentences like (28), Luque Castro considers it unlikely for verbs with a body part as the direct object is (see (29)), since constructions such as (30a), where the accusative encodes the body part and the ablative encodes the instrument placed on the body, are rather rare. He hypothesises that the origin of these constructions is a double accusative construction facilitated by the fact that Latin also had the possibility of encoding the item of attire put on or removed from the body in the accusative (30b). To date, we lack a quantitative study of textual data that confirms these hypotheses.

- (29) Me cubro la cabeza con un turbante
 REFL.1SG cover.PRS.1SG the head with a turban
 ‘I cover my head with a turban.’ (Luque Castro 2007)

increasing number of passive examples (Bogard 2006: 855). It is unlikely that the productivity of purely reflexive and reciprocal functions of the RM has decreased, since there is no other way of forming reflexive and reciprocal direct constructions in Spanish than the attachment of the RM.

(30) (Luque Castro 2007)

- a. Haec locutus gladio latus cingor.
 this say.PTCP sword.ABL wide.NOM circle.pass.1SG
 ‘This said, I surrounded my waist with the sword.’
 (Petronio: Satyricon, 82)
- b. Clipeumque auroque trilicem lorica
 shield.ACC-and gold.DAT-and of.three.threads.ACC breastplate.ACC
 induitur
 put.on.PRS.3SG
 ‘And he puts on his shield and his three-golds-breastplate.’
 (Virgilio: Aeneidos, 7: 639–640)

1.3.3 From reflexive to reciprocal

The development of reciprocal meanings from originally reflexive sources is extremely common cross-linguistically (König & Gast 2008, Maslova 2008). Maslova (2008) claims that “the roots of linguistic reciprocity are reflexive” (Maslova 2008: 233) and suggests that multiple-participant reflexive constructions are the most likely diachronic source of reciprocal constructions. However, not all reflexive-based middle systems develop reciprocal functions, and the reflexive may acquire middle functions without developing the reciprocal one (Nedjalkov 2007b). This is especially prominent in North-Germanic and East-Germanic languages. Accordingly, “semantic expansion of reflexive-middle markers does not necessarily include the reciprocal meaning” (Nedjalkov 2007b: 186).

Unfortunately, as far as I know, there are no specific studies on the actual development of the reflexive pronoun as a reciprocal marker in Latin and the Romance languages. This function of the RM is already documented in Latin, where it was more frequent in colloquial speech than in the writing of classical authors (Bassols 1956: § 291). However, the use of *se* alone as a reciprocal marker was not common in Latin, since the pronoun typically appeared with other unambiguous reciprocal phrases, such as *inter se* or *inuicem*, which could also mark reciprocity on their own, without the co-appearance of *se* as a full pronoun, like Romance languages require today (Baldi 1979; Ernout & Thomas 1984).

1.3.4 From reflexive to middle

A recurrent observation when seeking connections between Latin reflexive pronouns and Romance reflexive constructions is the fact that some Latin constructions are marked by the accusative pronoun *se* while others are marked with the dative pronoun *sibi*. Unfortunately, the lack of case-morphology in most modern Romance reflexive pronouns means that this distinction is not obvious (see Larochette 1939).

An interesting proposal here is made by Cennamo (1999), who studies the behaviour of both accusative *se* and dative *sibi* in Latin, proposing that they develop into different kinds of “pleonastic reflexives”. She examines reflexive intransitive and anticausative verbs of motion, speech and emotion, and concludes that, in Late Latin, *se* and *sibi* mark different kinds of intransitive verbs. On the one hand, *se*, originally being a marker for anticausative forms, marks atelic intransitive verbs with a controlling participant (an actor), that is, unergative verbs. *Sibi*, on the other hand, marks unaccusative verbs (i.e. typically telic verbs with an undergoer subject). Interestingly enough, Cennamo (1999) claims that the origin of these values of *sibi* does not lie in constructions where the reflexive pronoun is required by the verb, but in an emphatic construction where *sibi* attaches to adjectives. This construction, she argues, later spread to clauses conveying a change of state and location, where *sibi* equated to *per se* and produced spontaneous readings:

In particular, *sibi* takes up the inactive domain of the R-form, gradually weakening its original meaning of *Dativus Commodi* (where it was mainly confined to High participants, i.e., to animate, human, willful [sic] entities) till it ends up marking the spontaneous manifestation of a situation (whether action, event or state) regardless of the animacy of the surface subject. No External Causer is either overtly expressed or implied (recall the difference between *iumentum se refrigerat* and *iumentum sibi refrigerat*), and *sibi* is also used with inanimate participants which have no Control over the process (*tubergula sibi nascuntur*). *Se* on the other hand ends up marking External Causation, as in the so-called ‘passive’ use of the pattern *se* + active verb. (Cennamo 1999: 134)

By the eighth and ninth centuries, the path followed by *sibi* (from high to low unaccusativity) and the path followed by *se* (from low unaccusativity to unergativity) converged, while the loss of case distinction in the reflexive pronoun neutralised the two forms (Cennamo 1999). This proposal is extremely interesting, since it suggests that the RM that appears with verbs of change of state and location verbs on the one hand, and the RM used with verbs that convey mental processes, speech acts and other actions on the other hand, have different origins. The former would have developed from the dative reflexive pronoun *sibi*, while the latter would have developed from the accusative reflexive pronoun *se*.

1.3.5 Reflexive intransitive and transitive verbs

As regards the extension of the RM to intransitive and transitive verbs, the distinction between the Latin accusative and dative reflexive mentioned above is also crucial. Accordingly, many of the functions of the Spanish RM can be explained because of its double Latin origin: the accusative form *se* and the dative form *sibi* (Bobes Naves 1974a). It has been claimed that reflexive intransitive verbs like *irse*

‘to go-REFL’ and *morirse* ‘to die-REFL’ and reflexive transitive verbs like *beberse* ‘to drink-REFL’ and *comerse* ‘to eat-REFL’ come from Latin constructions with an ethical dative (Bobes Naves 1974a; Cano 1981; Fernández Ramírez 1986). However, it has also been claimed that the RM in intransitive verbs of motion has a double origin, namely, an analogy with anticausative verbs of motion (Cuervo 1954 [1907]: § 102) and as a natural product of the ethical dative, for which no analogical explanation is required (Martín Zorraquino 2002).⁹

Maldonado’s (1989) proposal is based purely on semantic connections. Rejecting the extended idea that the RM does not add any meaning when attached to intransitive verbs, he proposes that “these values of *se* correspond to the semantic value [of] Latin middle verbs in which the activity of the subject undergoing a corporeal or emotional change was emphasized in several ways” (Maldonado 1989: 345, emphasis in original). These are the uses he calls “energetic *se*”, where the RM highlights unexpected energy. On this view, the RM develops such a function because it limits the energy transfer of the event to the single participant of the event. Allegedly, the most prevalent function of the RM with intransitive verbs in the thirteenth century is a focussing one. From this function, energetic *se* developed in several stages, and its behaviour was already systematic by the sixteenth century:

Looking at the evolution of **Energetic *se*** it is rather clear that it starts applying to verbs of motion specifically to elements fixed in a supporting surface, then it applies to natural events such as rain and death where animates are included, then the number of verbs taking *se* increases considerably to the extent that there are many metaphorical extensions such as *irse* ‘fart’. At this stage it also applies to verbs that carry inherently an accidental or an inchoative meaning. And in the latest stages of the grammaticalization process it applies to human beings in verbs of motion even when the locative goal is present. It is in this stage that the construction not only covers a wide range of semantic space but also it covers an extensive variety of discourse environments to the extent that written formal discourse evading the use of *se* is interpreted as rather artificial in current Spanish.

(Maldonado 1989: 356, emphasis in original)

Besides these different diachronic interpretations of the facts, it seems clear that reflexive intransitive verbs of movement were especially frequent in Medieval Spanish and indeed already in Vulgar Latin (Martín Zorraquino 2002). In Old Spanish the RM was only common with *ir* and it gradually extended to other verbs of motion (Larochette 1939). In a corpus of texts from the twelfth to the sixteenth century,

9. Luque Castro (2006) proposes a rather different scenario, again based on syntactic reanalysis of several Latin constructions (accusative subjects in infinitive clauses, double object constructions, and impersonal psychological verbs with an accusative experiencer). This proposal, however, mostly leaves aside the different semantic values that have been attributed to these verbs.

irse ‘to leave’ accounts for more than half of the cases of this construction, followed by three other verbs of movement (*tornar* ‘to return’, *partir* ‘to leave’ and *venir* ‘to come’) (Bogard’s 2006).

1.3.6 Summary

Some main issues concerning the diachronic extension of constructions with the RM arise from the review of the literature. First, indirect reflexive constructions were a rather uncommon construction in Latin and are probably a Romance development. Second, the reciprocal construction, although connected to the purely reflexive construction, is not a necessary step for the middle functions of the RM to come about. Finally, the loss of case morphology in the RM of Romance languages obscures the fact that it conflates the accusative and the dative Latin reflexive pronoun – this difference might also be responsible for the very different possibilities that the RM show in modern Romance, especially concerning the paradigmatic reflexive constructions that can be called middle in a very broad sense. In the next subsection I focus on the concept of *middle* through a consideration of the typological literature.

1.4 Typological approaches to reflexive and middle constructions

In this section I briefly describe the state of the art of the literature on reflexive and middle constructions from a typological and cross-linguistic perspective. The large number of functions that reflexive constructions can have cross-linguistically explains the significant attention that has been paid to this linguistic category, closely related to the middle voice, currently a hot topic in the field. Moreover, because reflexive constructions recurrently show similar functions across languages, they have often been approached through cross-linguistic comparison. Of course, linguistic categories identified typologically need to be considered in prototypical terms, since the situation where two languages behave in exactly the same way regarding a given phenomenon is very rare (Bickel 2007). The constructions studied in this book are no exception:

More often than not, different ways of expressing middle, reflexive or reciprocal situations in one language will not match the strategies available in another language, even if the languages in question are genetically and geographically very close. However, there are also instances of similarity as a result of language contact or universal cognitive schemas. (Moyse-Faurie 2008: 107)

In Section 1.4.1 I will distinguish the concepts of voice and diathesis. In Section 1.4.2 I review the concept of the middle voice as discussed in the literature on the languages under study here and in Sections 1.4.3 and 1.4.4 I describe the two main concepts of middle voice available in the typological literature.

1.4.1 Voice and diathesis

The concepts of voice and diathesis have received a relatively heterogeneous treatment in the literature (for a recent historiographic study of the concept, see Signes Codoñer 2016). As Klaiman (1991) notes, it is not rare that voice is used as an intuitive concept, with no previous definition. Klaiman identifies three type of phenomena that can be described as voice phenomena in the literature:

1. **Derived voice:** in this conception, voice is related to the notion of argument alignment. Verbs have a default configuration (active voice) which may be altered in some constructions. A typical example of a derived voice is the passive voice, where the syntactic alignment of the two nuclear arguments of a transitive verb (agent and patient) is altered.
2. **Basic voice:** in this conception, voice is not related to verbal valency, but to the existence of a change in the status of the subject. The middle voice as used in the literature for classical languages like Latin and Classical Greek is a clear example of this view: “[t]raditionally [...], the middle in opposition to the active is said to express a distinction of meaning relating to alternate views of the logical subject’s participation in the denoted action. The middle, in contrast to the active, signals the subject’s coincidence with the locus of the action’s principal effects” (Klaiman 1991: 27).
3. **Pragmatic voice:** in this conception, voice is related to information structure and might refer to a change in default relationship between participants and thematic roles or in the ontological prominence of participants. An example of this view is the inverse voice, in which the verb is morphologically marked if the object is higher in the animacy hierarchy than the subject.
4. In the varieties under study, the middle voice has either been associated with changes in the verbal valency or linked to a more affected subject; that is, interpreted either as a derived voice or as basic voice, respectively. The same can be said for the typological literature, where accounts of the middle both as derived and basic can be found. A review of these account is given in the next section.

Despite the Romance tradition of not differentiating between voice and diathesis, I believe that distinguishing these terms in the spirit of the Leningrad Typological Group (see Kulikov 2011, Zúñiga & Kittilä 2019) is useful when speaking about

the derived voice (as García-Miguel 2001, 2010 does for Spanish). The distinction appropriately differentiates between syntactic behaviour and morphological markedness, for which Martín Zorraquino (1979) already used two different terms: diathesis and “diathety” (*diateidad*). According to Kulikov, diathesis refers to the syntactic configuration of the verb – it is the “pattern of mapping of semantic arguments onto syntactic functions (grammatical relations)” (Kulikov 2011: 370). As such, this notion is closely related to verbal valency, since the number of verbal arguments is crucial information to establish its diathesis (Kulikov 2011).

Voice, however, is a morphological concept and refers to “[the] regular encoding of diathesis through verbal morphology” (Kulikov 2011: 371). This distinction between the two terms is very useful, since diathesis and voice do not always match perfectly. As Kulikov (2011) notes, it is not rare to find diathesis alternations that are morphologically unmarked cross-linguistically. This phenomenon is called “lability”, and labile verbs are those that might alter their syntactic alignment without morphological marking (such as *break* in English: *I broke the vase / The vase broke*).

Moreover, voice morphemes are often polysemic and encode several diathesis alternations (García-Miguel 2001, Kulikov 2011). In these cases, “one morphological voice corresponds to a number of diatheses, a ‘diathesis cluster’ or ‘family’” (Kulikov 2011: 393). This morphological pattern is meaningful, since “diatheses belonging to the same cluster normally share some feature(s)” (Kulikov 2011: 393). As discussed in Section 1.4.3 below, this is the case for the middle voice when interpreted as a derived voice (Kulikov 2011, 2013, Zúñiga & Kittilä 2019, Inglese forthcoming).

1.4.2 The middle voice

The term “middle voice” originally referred to a morphological category of Greek verbs, whose meaning largely matches the meaning associated with many of the reflexive constructions in the varieties under study here. However, the Ibero-Romance grammatical tradition has consciously avoided the term, largely due to the fact that the Greek and Latin middle voice is an inflectional category of the verb, while the RM is (originally) a pronoun (Martín Zorraquino 1979; Pena 1982; Maldonado 1999). Other authors, such as Cartagena (1972), have rejected the term because there is no perfect match between “middle meanings” and the contexts where the RM appears. Many authors, however, have defended the opposite view, that is, that the RM is a middle voice marker in Spanish (Larochette 1939; Babcock 1970; Cano 1981; Pena 1982; García-Miguel 1985; Maldonado 1999). Even among those who have preferred to avoid the term, it is not rare to find the use of the adjective “middle” (*medio*, *medial*) to refer to some reflexive constructions of these varieties (see for instance NGLE 2009; Bobes Naves 1974a; Alcina & Blecua 1975).

Most often, the middle voice in Spanish is associated with the meaning of this voice in the classical languages, that is, with the status of the subject regarding the verb. Pena (1982) applies the classical distinction between active ([–affected subject]) and middle voice ([+affected subject]) to Spanish, claiming that the RM marks the reflexive middle and the internal middle voices, while the BE periphrastic passive marks the passive middle. A similar account can be found in García-Miguel (1985), who nevertheless believes that all three subtypes of middle described by Pena (1982) can be marked by the RM in Spanish. Larochette (1939) and NGLE (2009) propose that the middle voice is used when the subject of the verb undergoes a change of state (see also García-Miguel 2001). For Alcina & Blecua (1975), the subject of a verb in the middle voice is involved completely in the verbal action.

From a Cognitive Grammar perspective, Maldonado (1999) picks up on another of the recurrent notions that have been associated with the middle voice: the fact that it is a mixed voice, where the subject shows both active and passive properties (Klaiman 1991).¹⁰ On this view, however, the intermediate position of the middle voice corresponds to the flow of energy transferred from the agent to the patient, which is maximal in transitive constructions and amounts to zero in absolute constructions (with no direct object or RM) (Maldonado 1999).

From a different perspective, the relationship between the middle voice and the change of valency of the verb has also been highlighted, claiming that the main or only function of the RM is to intransitivise the verb (Babcock 1970; Heredia 1999; Cano 1981). At this point it we might mention the relationship between middle voice and transitivity, since

[v]oice systems exist in order to express divergences from canonical event types that fall at opposite extremes along a scale of semantic transitivity, a scale independently motivated by its effects on linguistic marking patterns other than voice. Thus, transitivity is the broader phenomenon within the framework of which voice phenomena should be understood. (Kemmer 1994: 221–2)

As we know, two competing concepts of transitivity are found in the literature. In the more traditional of these, transitivity is a syntactic concept referring to the fact that transitive verbs take a direct object, while intransitive verbs do not. The more modern view, however, was initiated by Hopper & Thompson (1980), who proposed a semantic and gradual conception of transitivity. On this view, transitivity can be

10. In Klaiman's words: "Originally, the middle seems to have been conceived as a compromise category displaying characteristics of both the active and the passive. In a middle construction, the viewpoint is active in that the action notionally devolves from the standpoint of the most dynamic (or Agent-like) participant in the depicted situation. But the same participant has Patient-like characteristics as well, in that it sustains the action's principal effects" (Klaiman 1991: 3).

broken down into several binary parameters. Transitivity becomes a property of the sentence (and not only of the verb) and the degree of transitivity of a sentence can be ranked on a continuum from low to high transitivity. While this is not the place to explain the ten transitivity parameters, it is worth noting that the strongest prediction of Hopper & Thompson's (1980) model is that these features co-vary: "whenever an obligatory pairing of two Transitivity features occurs in the morpho-syntax or semantics of a clause, THE PAIRED FEATURES ARE ALWAYS ON THE SAME SIDE OF THE HIGH-LOW TRANSITIVITY SCALE" (Hopper & Thompson 1980: 254, emphasis in original).

The middle voice, however, seems to be a special case regarding this generalisation, which fits with its mixed character. Both Kemmer (1993) and Maldonado (1999) place the middle voice in an intermediate position on the continuum between the transitive and intransitive prototypes, because middle markers are sometimes associated with low transitivity features and sometimes to high-transitivity parameters.

Let us illustrate this with an example. Hopper & Thompson (1980) use the Spanish examples in (31) to suggest that the presence of the RM *se* is associated in Spanish with the high-transitivity pole, since it co-varies with punctuality: with some verbs, the RM is used with punctual events, which is a high-transitivity feature, as opposed to its low-transitivity counterparts, i.e. states.

- (31) (Hopper & Thompson 1980: 266)
- a. Juan durmió (toda la noche).
 Juan sleep.PST.3SG all the night
 'Juan slept through the night.'
 - b. Juan se durmió (*toda la noche).
 Juan REFL.3 sleep.PST.3SG all the night
 'Juan fell asleep.'

However, the examples in (32) offer the exact opposite interpretation, since the presence of the RM co-varies with the low-transitivity pole of the parameter of number of participants. When *dormir* 'to sleep' shows higher transitivity by taking two participants (in (32a)), the RM is ruled out, while it can appear when the verb has only one participant (32b) – a low-transitivity feature.

- (32) a. Juan (*se) durmió al niño.
 Juan REFL.3 sleep.PST.3SG to.the child
 'Juan put the child to bed.'
- b. El niño (se) durmió.
 the child REFL.3 sleep.PST.3SG
 'The child slept/fell asleep.'

The conception of the middle voice as an intermediate category might account for its mixed behaviour, but poses a problem of potential circularity, since every combination of Hopper & Thompson's (1980) parameters would be in principle possible. The analysis of the limits of such a combination would require a different study and will not be addressed in the current book. Since the association between the middle voice and intransitivisation in the varieties under study has been made under the traditional concept of (in)transitivity, I will also refer to such an interpretation in what follows. Whenever the effect of the semantic parameters associated with transitivity in Hopper & Thompson's (1988) account is considered, I will not propose particular hypotheses of how to deal with the behaviour of the RM according to those parameters and the global transitivity of the clause.

Finally, we should note a particular use of the term "middle", mostly within the generative framework. In the generative literature it is frequent to find the name "middle construction" to refer to constructions which can only appear in the 3rd person (and hence belong in the non-paradigmatic subclass of reflexive constructions, Section 1.2.4). These constructions do not convey an event, but rather a property of the subject. They always have a generic interpretation, do not admit perfective tenses and typically have a modality adverb (see (33)). As noted above (Section 1.2.4), these constructions will be omitted from this study, since I consider them to be a subtype of the passive reflexive.

- (33) Las camisetas de algodón se lavan fácilmente
 the shirts of cotton REFL.3 wash.PRS.3PL easily
 'Cotton shirts wash easily.' (Mendikoetxea 1999b: 1654)

In what follows, I will review two typological approaches to the middle voice. In Section 1.4.3 I describe Kulikov's (2011) approach to the middle as a derived voice and in Section 1.4.4 I discuss Kemmer's (1993) typological approach to the middle as a basic voice.

1.4.3 The middle voice as a derived voice

The relationship between reflexive constructions, the middle voice, and intransitive diatheses has often been noted in the literature. Take, for instance, this statement from König & Gast (2008): "Reflexive pronouns in French, by contrast, are used [...] for the so-called 'middle-domain', i.e. as markers of derived intransitivity [...]" (König & Gast 2008: 6). Even if derived intransitivity (or detransitivisation) is often taken as the main or core function of the middle voice, it has also been noted that the match between middle voice and intransitivity is not perfect (Kulikov 2013; Mous 2007, Inglese forthcoming). In Konso (Cushitic), for example, "[v]erbs

with a middle derivation can be transitive [...] and even ditransitive [...]. Middle derivation changes the meaning of the frame of the verb but does not necessarily reduce the number of arguments to one; middle verbs need not become intransitive” (Mous 2007: 214).

It is important to highlight that, in conceptions of the middle as a derived voice, the different diatheses marked by the middle voice are generally compared to a “more basic” structure, which is typically the transitive active construction (Klaiman 1991, García-Miguel 2010). This characterisation as “more basic” might refer to several aspects, such as being morphologically unmarked, being more frequent in the discourse (and hence less marked), being the earliest construction from a historical point of view, or being the starting point of a syntactic derivation. The three first interpretations are empirical matters, while the latter is purely theoretical. Whenever I talk about demotion, removal or promotion of verbal arguments in this book, such notions will be intended as helpful descriptive terms in the comparison of two diatheses of a verb, and not as theoretical proposals of transformational processes.

From a cross-linguistic point of view, middle voice marking is typically associated with at least seven diatheses: reflexive, reciprocal, anticausative, conversive, passive, antipassive and autobenefactive (Kulikov 2011, see also Zúñiga & Kittilä 2019 and Inglese forthcoming). All these diatheses can show reflexive marking in the varieties under study, plus the de-objective diathesis, which Kulikov (2011) considers together with the antipassive. However, I will consider it as an independent type of diathesis, following Geniušienė (1987).

The effect of these diatheses on the verbal participants is not the same (Kulikov 2011, Zúñiga & Kittilä 2019). Accordingly, three types of diatheses can be distinguished: diatheses *sensu stricto* (conversive and antipassive diatheses), diatheses *sensu latiore* (anticausative and de-objective diatheses) and operational diatheses (reflexive, reciprocal and autobenefactive diatheses). While in diatheses *sensu stricto* all the semantic roles selected by the verb are preserved, in diatheses *sensu latiore* some of them are lost. That is, while the agent is removed in the anticausative diathesis, it is only demoted from the subject position to an oblique position in the passive and conversive diatheses. Similarly, the patient or theme is removed in the de-objective diathesis, but demoted to an oblique position in the antipassive. Finally, all semantic roles are preserved in operational diatheses, but undergo some kind of operation. The reflexive diathesis shows coreferentiality between the subject and the object (be it direct or indirect), while in the reciprocal diathesis this coreferentiality arises between two different events – and in this sense it can be said that a reciprocal sentence is related to at least two basic clauses (with the same verb). The autobenefactive diathesis is affected by two derivations, namely, the addition of a

One of the main properties on which the relative elaboration of events is based is the relative distinguishability of participants, i.e. “the degree to which a single physico-mental entity is conceptually distinguished into separate participants, whether body vs. Mind, or Agent vs. Unexpectedly contrasting Patient” (Kemmer 1993: 66). Thus, while in transitive constructions the two verbal participants are presented as maximally different, the degree of distinguishability between participants in intransitive sentences is minimal (for there is only one). Since the two verbal participants are coreferential in reflexive sentences, the degree of relative distinguishability is lower than in transitive sentences. Middle sentences show an even lower degree of distinguishability, because in these sentences there is an expectation of coreferentiality (as opposed to the reflexive domain, which affects events that more often than not have distinct participants).

Kemmer’s (1993) differentiation between the reflexive domain and the middle domain is based on the existence of two-form languages (as opposed to one-form languages), which have two distinct markers, namely, a reflexive marker and a middle marker. A relevant typological characterisation can be made regarding the morphological complexity or phonological weight of these two markers: when a difference between them can be found, the reflexive marker tends to be heavier than the middle marker. A prototypical example comes from Russian, where reflexive sentences are marked by the independent pronoun *sebjja*, while middle sentences are marked by the suffix *-sja* (Kemmer 1993).¹² Although the varieties under study here are one-form languages, showing the same marker for the reflexive and the middle domain, this difference between a heavy and a light marker can be projected to some extent in these varieties, due to the existence of prepositional emphatic reinforcements (Sp. *a sí mismo*, Gal. *a si mesmo* ‘to one self’, see Section 1.2) that can only appear in the reflexive domain (see also De Benito Moreno 2015).

Kemmer classifies the types of events that show a low degree of relative distinguishability of participants or elaboration of events and which tend to show middle marking cross-linguistically. In her description of the different types of events that show middle marking in the languages of the world, Kemmer uses

12. The motivation for this typological tendency has been explained in terms of iconicity. For Kemmer (1993), “[g]reater phonological/morphological substance is associated with a greater degree of distinguishability of participants, and a greater degree of distinguishability of events” (Kemmer 1993: 121). For Haiman (1983), however, this iconicity is related to the expectation of coreferentiality and hence “marked form corresponds to marked meaning” (Haiman 1994: 1632). Haspelmath (2008: 45) agrees with Haiman (1994) in that this tendency is related to the “principle of economical coding of predictable information”, and goes further by suggesting that this principle is ultimately motivated by token frequency – verbs which show coreferentiality more often are those marked by the light marker.

situation types, that is to say, “sets of situational or semantic/pragmatic contexts that are systematically associated with a particular form of expression” (Kemmer 1993: 7). She identifies a number of situation types that are typically associated with middle-marking across the languages of the world and proposes a semantic map that links the semantic relation between such types.

In what she calls “the body action middle”, Kemmer differentiates between grooming verbs (‘to shower’), change in body posture events (‘to stand up’, ‘to sit down’), non-translational events (‘to turn’, ‘to bow’) and translational motion events (‘to go’, ‘to come’, ‘to run’). These four categories are ordered according to their proximity to reflexive constructions (see Figure 2). In the so-called “cognition middle”, which includes mental events, Kemmer identifies cognition, emotion and perception verbs. Another situation type related to the middle domain involves spontaneous events, “which designate changes of state of an entity” (Kemmer 1993: 142) and in which no agent is encoded. Finally, she considers two domains often coded by the RM (the indirect reflexive and the reciprocal) and claims that the same notion of relative distinguishability of participants and events can be used to differentiate between situation types more prone to be marked by the RM and those more prone to be marked by the middle marker.

The different situation types identified by Kemmer will be described in more detail in Chapter 3. For the moment, it will suffice to mention them and to describe briefly the semantic relation they have with other situation types. In order to do so, it is useful to adopt Kemmer’s visual representation, provided in Figure 2. Here, spatial distance is intended to represent semantic distance and the lines connecting different types represent direct semantic connections between situation types. The situation types in bold are the direct constructions that will be addressed in this volume.

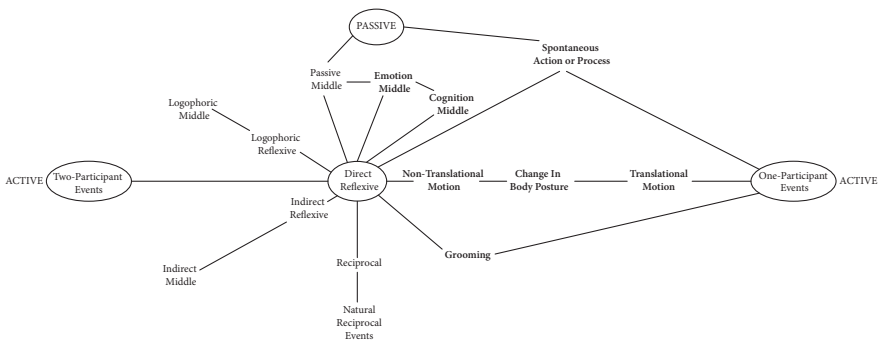


Figure 2. Semantic relations among the middle and other situation types (adapted from Kemmer 1993)

Maldonado's (1999) applies Kemmer's (1993) idea of the gradual distinguishability of events and participants to Spanish reflexive constructions. He claims that, in prototypical reflexive constructions, the agent and the patient are conceptualised in different mental spaces, while in middle constructions there is only one mental space. Thus, all middle constructions share two features: (1) the focus on the change of state of the thematic patient, and (2) the merely schematic representation of the inductive force. More specific meanings depend on the characteristics of the verb and the specific conceptualisation of the scene. Some key aspects of his analysis are the attribution of a general semantic function to the RM in middle constructions, that is, treating the middle voice in Spanish as a basic voice, and the rejection of a syntactic function of the RM (such as cancelling a verbal argument). For Maldonado (1999), the main function of the RM is to place the focus on the change of state and to depict the involved participant with intermediate – i.e. middle – characteristics between agents and patients. Similarly to Kemmer (1993), he relies on the characteristics of the verb to derive more specific features of the different middle constructions and proposes a net of semantic relationships that interconnects all these different functions. Two examples suffice to illustrate the claims of his analysis:

1. In bivalent constructions with an experiencer, the RM either decreases the control level of agentive experiencers (see (34a, b)) or increases the emotional participation of “thematic experiencers” (see (34c, d)). Therefore, the presence of the RM brings the experiencer to a middle domain where the participant is neither completely active nor fully thematic.

(34) (Maldonado 1999)

- a. Recordé su cara.
remember.PST.1SG his/her face
'I remembered his/her face.'
- b. Me acordé de su cara de pronto.
REFL.1SG remember.PST.1SG of his/her face of early
'I suddenly remembered his/her face.'
- c. No sabes cuánto me alegre
not know.PRS.2SG how.much ACC.1SG make.glad.PRS.3SG
verte.
see.INF-ACC.2SG
'Seeing you makes me incredibly glad.'
- d. No sabes cuánto me alegre de
not know.PRS.2SG how.much REFL.1SG make.glad.PRS.1SG of
verte.
see.INF-ACC.2SG
'I am incredibly glad of seeing you.'

2. In passive, impersonal and anticausative constructions, instead of linking the presence of the RM with the deletion of the agentive argument, Maldonado proposes that the RM allows the inductive force to be present, but in a merely schematic way:

(35) (Maldonado 1999)

- a. En el temblor, los platos se rompieron.
in the earthquake the plates REFL.3 break.PST.3PL
'The plates broke in the earthquake.'
- b. La maleta se perdió en el aeropuerto.
the suitcase REFL.3 lose.PST.3SG in the airport
'The suitcase was lost in the airport.'

In this study I will mainly stick to Kemmer's (1993) categories, in that Maldonado's (1999) work relies heavily on Kemmer's study and also because the approach in that study allows for a clearer definition of the different categories. This better suits a variationist account based on empirical data, since it allows for a more objective operationalisation of the data. To conclude this section, it is worth noting the four characteristics of middle voice systems that have cross-linguistic validity according to Kemmer (1993) and hence to contextualise them in relation to the present study:

1. Not all verbs in a given situation type take the middle marker. In Chapter 3, I compare the effect of syntactic and semantic parameters on the distribution of the RM in order to investigate which of the two conceptions of the middle voice better fits the behaviour of the RM in the varieties under study.
2. The middle marker is often optional, its presence or absence not leading to any significant difference of meaning. The variationist perspective adopted in this study aims to arrive at a better understanding of the factors underlying this optionality.
3. Middle-marked verbs with an unmarked transitive counterpart are very frequent. As noted above, this is related to the conception of the middle voice as a derived voice and the middle marker as an intransitivising device. The suitability of the two interpretations of the middle voice in the varieties under study will be empirically addressed in this study.
4. Middle-marked verbs with no unmarked counterpart at all are also common (see also Inglese forthcoming, who considers them a necessary element of middle voice systems). I will call them non-reversible verbs, following Geniušienė (1987). These verbs are problematic for purely syntactic accounts of the middle, since there is no diathesis alternation (or no alternation at all).

In the next section, I briefly review some of the theoretical notions related to the study of language that are crucial for an understanding of the perspective adopted in this study.

1.5 Theoretical framework

In this section I present the main theoretical assumptions on which this study is based. Generally speaking, I feel comfortable describing my approach as what Newmeyer (2005) calls holistic functionalism, and I subscribe to the statement that “[t]here is no direct linkage between external functions and grammatical properties. The influence of the former on the latter is played out in language use and acquisition and (therefore) language change and is manifested only typologically” (Newmeyer 2005: 175). This point underlines why most theoretical observations made in this section derive from the theory of language change, even though my study itself focusses on a synchronic (i.e. static) stage of the varieties explored. In Section 1.5.1 I discuss the relationship between language change and language variation, in Section 1.5.2 I review the mechanisms of language change pertinent to this study, and in Section 1.5.3 I consider the spatial diffusion of language change.

1.5.1 Language change and variation

From a general perspective, I subscribe to Weinreich, Labov & Herzog’s (1968: 162) view of language as “an orderly heterogeneous system in which the choice between linguistic alternants carries out social and stylistic functions, a system which changes with accompanying changes in social structure”. This well-known conception is the starting point for most modern work on sociolinguistics and historical linguistics. Within this perspective, linguistic variation can be described by means of the concept of the linguistic variable (the dependent variable), which consists of at least two linguistic variants that show different probabilistic distributions according to different linguistic and extra-linguistic factors (the independent variables). In this study the linguistic variable has two possible variants: the presence of the RM and its absence.

One of the crucial characteristics of this conception of language is that it provides a straight-forward link between synchrony and diachrony, since language change is explained by the fact that linguistic systems are intrinsically variable. Conversely, synchronic stages can only be explained through recourse to historical explanations (Coseriu 1973). Thus, it often happens that the behaviour of two (or more) competing synchronic strategies cannot be satisfactorily explained in synchronic terms (Bybee 2001). Moreover, this generalisation can also be applied to other dimensions, such as the spatial (diatopy) or social (diastraty) axes of variation and change. This in turn explains why in the current study we will resort to a theory of language change – although it is a synchronic study, not a “syntopic” one – since we will be comparing different (but closely related) varieties.

Accordingly, I concur with the claim that linguistic explanations must be historical explanations, understood in Coseriu's (1973) sense: "the general problem of [language] change lies in establishing the ways and conditions of that change" (Coseriu 1973: 112, my translation). Since these ways and conditions do not necessarily always trigger the same consequences, linguistic explanations cannot be causal explanations: "causal explanations do not, and cannot, fully explain the intentional, and fundamentally historical, activity human beings are engaged in when they create language" (Willems 2013–4: 110, see also Coseriu 1973; Itkonen 2013–2014). Linguistic explanations can also not be teleological, since languages are not intentional subjects and individual speakers cannot foresee the ultimate consequences of the changes in which they are a part (Coseriu 1958; Itkonen 2013–2014). Therefore, one of the goals of this study is to establish the ways and conditions that can explain the different developments of the RM.

Another consequence of this conception of language is that the object of inquiry is not linguistic competence, but performance, since, as noted by Saussure, language changes in speech (Saussure 1945 [1916]). This has a straight-forward effect on the type of data used, which must be real instances of speech; that is, this will be a strictly usage-based account. I will discuss the methodology and data used in this study in Chapter 2.

A full understanding of language change requires that at least two phases be distinguished: innovation and diffusion or adoption. This distinction was popularised initially by Weinreich, Labov & Herzog (1968), but in fact had already been fully developed by Coseriu (1958, 1973). As Coseriu (1973: 79–80) puts it, language change only occurs in the diffusion phase, since while innovation is an individual act, diffusion is a social act. According to this view, only social reasons can explain why an innovation is adopted by and diffused through a given community. As noted above, I will focus on the conditions for the innovations in the evolution of the RM, and towards understanding its diffusion, a diachronic study must be undertaken. However, this study is also concerned with the geographical distribution of the variation to be analysed, which has a direct link to one of the aspects of social diffusion: diffusion across space. In what follows I address (very briefly) three crucial concepts that have to do with the cognitive processes underlying linguistic change: reanalysis, analogy, and frequency (Section 1.5.2). In Section 1.5.3 I describe the two most popular models of spatial diffusion of language change.

1.5.2 Mechanisms of language change

One of the main mechanisms of innovation described in the literature is reanalysis, which refers to the novel interpretation of a sequence. Reanalysis can be understood in syntactic terms:

Reanalysis [...] is a mechanism which changes the underlying structure of a syntactic pattern and which does not involve any immediate or intrinsic modification of its surface manifestation. [...] While the surface structure is not directly affected by reanalysis, underlying structure is. (Harris & Campbell 1995: 61)

On this view, reanalysis might affect several aspects of the structure, such as the grammatical relationships or the category labels. According to Harris & Campbell (1995), the only requirement for a reanalysis to occur is that a sequence can be interpreted as having two different underlying structures, that is, no semantic ambiguity is needed according to these authors. However, for other authors, reanalysis is a semantic change based on the more general process of metonymy. On this view, “semantic change is the really important phenomenon in reanalysis, and [...] rebracketing and relabeling of the constituent structure are merely dispensable side-effects” (Detges & Waltereit 2002: 168–9).

Reanalysis, which is considered to constitute the beginning of many language changes, is an abrupt process, one that might originally affect a single sequence. “[T]he gradual mapping out of the consequences of the reanalysis” (Timberlake 1977: 141) is known as the actualisation of reanalysis. As Harris & Campbell (1995) have noted, this concept can be applied to the spread of changes not originated by reanalysis, for which they propose the term extension. A clear result of this extension is that the new construction might affect more lexical items, suffering a process of schematization, as Bybee (2015) calls it, “because it results in the construction covering more distinct lexical items, thereby creating a schematic category” (Bybee 2015: 172).

An interesting proposal about how this extension or schematization takes place is given by Fischer (2008), who places great value on the notion of analogy. While analogy has typically been applied to morphological changes that extend through paradigmatic similarity (Coseriu 1973; Elvira 1998, see Bybee 2015), Fischer uses a broader interpretation of the concept, which includes syntagmatic similarity.¹³ According to Fischer (2008), language is organised into two types of connections, iconic and indexical. Iconic connections, being based on semantic or formal similarity, are paradigmatic, while indexical connections arise through association

13. Bybee (2015) also acknowledges the similarities between the extension of a syntactic construction and analogical levelling.

within syntagmatic relationships. That is to say, a new construction might not only expand to new lexical items, but might affect these new lexical items gradually, starting in those contexts that are also syntagmatically similar to the starting point of the construction. In this conception, analogy is a very powerful tool and is the main mechanism behind the extension of language change, which spreads through similarity among contexts (De Smet 2012). A consequence of great importance for the current study is that “an item’s behavior subsequent to reanalysis is codetermined by its behavior on its original analysis” (De Smet 2012: 603). Therefore, we expect similarities between the behaviour of a sequence in its old and new usages, which should be traceable through quantification. Thus, we expect analogical extension to follow an ordered path through new contexts depending on the similarity of these to the original context.

Finally, we need to highlight the role of frequency in language change. This notion has received considerable attention since Bybee’s (1985) observations. The main idea is that “conventionalization through repetition creates grammar” (Bybee 2007: 6). This is justified in cognitive terms: the human brain is sensitive to the effects of repetition. On this view, the frequency of linguistic items, sequences or constructions has a direct effect in how they are stored in the speaker’s memory, in how they are connected to other stored items and in how accessible they are (Bybee & Hopper 2001).

Type frequency (as opposed to token frequency) refers to the number of different elements that can appear in a given “pattern of language” and has a determinant role on the productivity of a construction (Bybee 2001, 2007), because the higher the type frequency of a construction, the greater the analysability of the construction (Bybee 2001). Accordingly, the more productive a construction, the more it will spread to new elements (Bybee 2001, 2007; Bybee & Thompson 2007). As a consequence, categories with a large number of members show a stronger tendency to attract new members than categories with a small number of members (Bybee 2001; Bybee & Thompson 2007).

1.5.3 Spatial diffusion of language change

After some observations on linguistic mechanisms of language change, a brief note on its spatial diffusion is in order. Two of the most common models of spatial diffusion discussed in the literature are the wave model and the gravity model (Chambers & Trudgill 1980; Britain 2004). In the former, the extension of a pattern across space is crucially determined by geographic proximity – the strength of an innovation is an inverse function of the distance to the focus of the innovation. In the latter, it is not only euclidian distance that affects spatial diffusion, but also demographic size – larger towns receive linguistic innovations before smaller towns,

even if the later are geographically closer to the focus of the innovation, because they are better connected.

Since the data for this study come from the last generation of rural speakers before massive urban migration, and hence urban data are not dealt with here (see Chapter 2), we expect the wave model to be more appropriate for explaining the geographic patterns that arise from the analysis. The wave model is typically illustrated using a diagram similar to the one provided in Figure 3. A given change will be more deeply rooted in the focus or centre of the innovation, that is, the new linguistic pattern will appear in more contexts (symbolised as letters in Figure 3) in the location of the focus of the innovation. The new pattern will spread gradually, broadly as a function of geographical distance. A geographical pattern like this allows us to establish the implicational hierarchy that is followed in the diffusion of the new pattern, and hence a diachronic reconstruction – in Figure 3 we can assume that the new pattern reached context C only after having reached context B, which in turn was reached after context A.

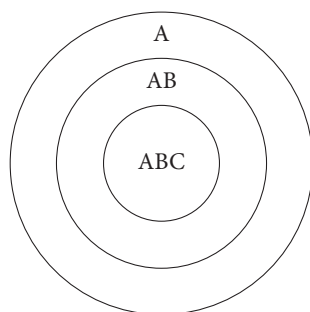


Figure 3. Idealised map of wave diffusion (adapted from Haas 2010)

Although this model does not capture the effect of social factors or the fact that diffusion might not follow a constant rate through all contexts (Haas 2010), the asymmetrical distribution of different contexts in space is an important indication of different diachronic stages. In Chamber & Trudgill's words, "spatial diffusion of linguistic innovations often constitutes a kind of reflection of the other types of diffusion" (Chambers & Trudgill 1980: 182). Or, as Fernández-Ordóñez puts it, "the linguistic history of a territory is written in its dialect areas" (Fernández-Ordóñez 2011: 24, my translation). Accordingly, the dialect data for this study will be used both as a synchronically descriptive tool and as a method to obtain diachronic hypotheses, which it is hoped will be followed up in future historical studies.

Methodology

2.1 Introduction

The study of reflexive constructions from a dialectal and variationist perspective requires the analysis of real speech. In this chapter I describe in detail both the corpora and the techniques used in the collection and analysis of the data. A sound description of the research method is essential for both assessing the validity of the research and allowing for its reproducibility. As Gordon (2005: 958) puts it: “Regardless of the procedures used, what is crucial is that researchers be careful in making generalizations on the basis of their samples and that they provide details about their sampling procedures so that others can judge the validity of the claims”.

This study is based largely on spoken language, mostly from the Spoken and Audible Corpus of Rural Spanish (henceforth COSER, after its Spanish acronym) (Subsection 2.2). Other sources have been used sporadically, these described in Subsection 2.3. In Subsection 2.4 I provide a brief note on the statistics and visualisation techniques used.

2.2 The COSER data

The Spoken and Audible Corpus of Rural Spanish (COSER) is a project initiated in 1990 by Inés Fernández–Ordóñez. Its goal is to collect large spoken samples of rural varieties of European Spanish for the study of dialectal morphosyntax. In order to do so, COSER combines the goal of traditional linguistic atlases, i.e. to study rural varieties, with the main sociolinguistic tool, i.e. the semidirected interview.

This combination is especially appropriate for the study of dialectal morphosyntax for two main reasons. On the one hand, rural and elderly speakers who have always lived in their place of birth – which is also where the interview takes place – are representative of a linguistic layer which is difficult to find in other sources.¹ This linguistic layer is especially interesting for the study of dialectal variation, since it

1. That is, the informants of COSER mostly conform to the label NORM label (‘non-mobile older rural males’), coined by Chambers & Trudgill (1980) to refer to the prototypical informant in traditional dialectology. However, COSER includes women, who represent a little more than half of the informants of the corpus (52% in May 2022).

often shows a linguistic continuum, offering a much direct access to the genetic relationships between different varieties, which are related by spatial proximity. Such a situation is not expected in the study of urban varieties, whose mutual relationships are affected by other parameters, such as the size of the urban centre (Britain 2004).

On the other hand, the use of semidirected interviews substantially improves the traditional methodology of linguistic atlases, which typically obtain their data through closed questionnaires. The semidirected interview reduces the attention paid by the speaker to their own speech, since they are not aware that the ultimate goal of the task is to study their speech. This minimises the effect of the observer's paradox (Labov 1972). The collection of large samples of free speech also allows for the possibility of documenting multiple cases of a given linguistic phenomenon, which in turn enables the quantitative analysis of the data. Moreover, the semidirected interview allows for the study of morphosyntax, which was traditionally ignored in most linguistic atlases and for which the use of questionnaires is often problematic.

The procedure followed for collecting the COSER interviews is simple. A speaker who fulfils the COSER requirements (that is, an elderly native, originally from the place of interview or closely nearby, and who has not spent long periods of time away) is found and asked whether they accept being interviewed for around one hour. The mean duration of the COSER interviews is 64 minutes (in October 2019). Interviews are conducted by teams of two to five people who have normally never met the informant before. This might lead to the objection that the distance between interviewers and interviewee could affect the speech of the latter. Hence, two points that must be taken into account. First, the informant is a first-hand expert on the topic or topics of the interview, which reduces the potential inhibition of informants as a result of them being interviewed by a team of university students. Second, the effect caused by the asymmetry of the situation is expected to affect all interviews similarly, suggesting that the comparison of the interviews should offer valid results, especially in relative terms.

2.2.1 Selected interviews

A total of 151 COSER interviews have been analysed for the current study. These amount to approximately 199 hours of material, and were conducted in all parts of the peninsular Spain, except Catalonia.²

2. COSER is an ongoing project, and interviews from Catalonia had not been conducted at the time of the present study. However, most of this region had indeed been included by the end of 2021

Since Spain is a multilingual country, some considerations have to be made regarding the language of the interviews and the informants: (1) in the Region of Valencia, all interviews were conducted in villages which are traditionally monolingual Spanish; (2) in Biscay and Gipuzkoa (the Basque Country) all interviews were conducted in Spanish, with both monolingual (Spanish) and bilingual (Basque/Spanish) speakers; (3) in Asturias, most speakers were bilingual (Spanish/Asturian) but preferred to use Spanish during the interview; (4) in Galicia, all the speakers were speakers of Galician (most of them bilingual with Spanish), but some preferred to use Spanish during the interview. Since not enough interviews were conducted in Galicia to confirm the effect of the chosen language on the presence of the RM,³ all interviews are considered together.

Map 1 shows the geographic distribution of the interviews. The list of localities, their duration, and which language was used, can be consulted in Appendix 1.



Map 1. Distribution of the COSER interviews analysed

The distribution of interviews by province is shown in Table 1. As can be seen, for most provinces, four to five interviews were analysed. However, some areas (Murcia, Region of Valencia, Galicia and most of Andalusia) show a lower density of interviews, due to the fact that interviews in these areas were conducted more recently, so fewer transcriptions are available. As partial compensation for this, some of the interviews selected for these areas were longer than usual (see Appendix 1).

3. Galicia has also not been exhaustively interviewed for COSER yet, but an extraordinary interview campaign was carried out in 2012 to collect data for the present study.

Table 1. Number of interviewed localities per province

Province	Number of localities	Province	Number of localities
Álava	4	La Coruña	3
Albacete	4	La Rioja	4
Alicante	2	León	5
Almería	2	Lugo	2
Asturias	4	Madrid	4
Ávila	4	Málaga	2
Badajoz	4	Murcia	2
Burgos	5	Navarra	5
Cáceres	5	Orense	3
Cádiz	3	Palencia	5
Cantabria	5	Salamanca	4
Castellón	2	Segovia	4
Ciudad Real	4	Sevilla	2
Córdoba	2	Soria	4
Cuenca	4	Teruel	4
Granada	2	Toledo	5
Guadalajara	4	Valencia	2
Guipúzcoa	3	Valladolid	6
Huelva	1	Vizcaya	5
Huesca	4	Zamora	4
Jaén	4	Zaragoza	4

A final observation must be made regarding those cases where the speech of more than one person was recorded. While the COSER interviews preferably record one single speaker, it is not rare that either other speakers participate sporadically (passers-by) or that the informants insist on being interviewed together with their spouse or friends. All informants that fulfil the COSER requirements as described above were considered for the analysis, meaning that the whole interview is taken as representative of the speech of the interviewed locality. This is justified by the fact that the corpus does not normally allow for a comparison of speakers from the same village, and hence a differentiation between speakers makes little sense.

In the next Subsection I describe the methodology used to extract the relevant examples from the COSER interviews.

2.2.2 Collection of examples

All relevant examples were extracted by hand, through the careful reading of the revised transcriptions.⁴ However, this method is not free from problems, since the identification of all cases where the RM *can* appear is a challenging matter. Most relevant contexts are easily recognisable through the combination of native intuition and a reading of the literature, but this leaves out those examples that do not belong to my native variety and that have not been described in the literature. Such cases where the presence of the RM is not linked to a change of valency are especially complicated, since there is no way of reducing the potential envelope of variation; potentially every verb could show the RM.

In order to reduce the effect of these complications, I collected exhaustively all cases of those verbs which are known to take the RM in some contexts (even if only rarely, like *estar* ‘to be’) and of some verbs which are known to take the RM (or the middle marker) in languages with a similar middle voice system (such as *oler* ‘to smell’, cf. Lat. *odōror* ‘to smell’, apud Kemmer 1993: 137; *nacer* ‘to be born’, cf. Rom. *a se naște*). Retrospectively, this turns out to be a somewhat exaggerated prevision, since the “unexpected” reflexive forms that were documented were so scarce that they could not be compared with their unmarked counterparts. However, for verbs which are frequent overall (such as *estar* ‘to be’, *saber* ‘to know’, *hacer* ‘to make’, *coger* ‘to get/catch’, *comprar* ‘to buy’ and *buscar* ‘to search’) this was a fruitful prevision that enabled the assessment of the global frequency of the reflexive form, or even the relevance of some linguistic and extra-linguistic parameters in the distribution of the RM, as will be discussed later.

At any rate, the fact that the RM was “unexpectedly” documented with some verbs means that the envelope of variation was not always collected exhaustively (i.e. it does not contain all the cases with and without the RM) for all verbs. During the analysis I indicate this whenever it is the case. Most of the time, however, all contexts are taken into account, and this is to be understood as the default situation. The normalization techniques used for the representation of frequencies in those cases where unmarked contexts were not collected systematically are also indicated in the analysis.

Due to the extremely high frequency of some verbs (such as *ir* ‘to go’, *llevar* ‘to carry’, *estar* ‘to be’, *saber* ‘to know’), these were only extracted in a subset of the interviews. I refer to the subset of interviews where all verbs were considered as “subcorpus E” (after *exhaustive subcorpus*) and to the set of all interviews as “subcorpus NE” (after *non-exhaustive subcorpus*). The default situation is that the data

4. In order to reduce the number of transcription mistakes, I personally revised each transcription by listening to the recording. Any remaining errors are therefore my responsibility.

belong to the whole corpus (subcorpus NE), so I only explicitly mention subcorpora when subcorpus E is used.

Subcorpus E contains 66 localities (see Appendix 1), whose distribution is shown in Map 2. The distribution of considered places by province is given in Table 2. Subcorpus E contains at least one locality per province and includes all interviews from the north-western area, which is very relevant for our study. Since this is a small area in comparison to the rest of the territory, all interviews were analysed exhaustively, in order to yield the greatest number of items for comparison from such a small area. In order to maintain an approximately even density distribution of localities in the rest of the territory, some provinces are represented by more than one locality.



Map 2. Distribution of the localities in subcorpus E

Finally, when analysing data from spoken corpora, problems of classification often arise. The following is a list of the decisions made regarding these problems:

1. If the informant repeated a sequence many times, using it as discursive tag or filler, I keep only one occurrence, discarding all others. I removed occurrences of *yo qué (me) sé; qué sé yo* ‘what do I know’; *¿sabes?; ¿saben?; ¿sabe?; ¿sabéis?; ya sabes* ‘you know?’; *¿ves?; ¿ve?; ¿ven?; ¿veis?* ‘see?’; *ya verás; verá* ‘you will see’; *que no veas* ‘a lot’; *mira; mirad/mirar; miren; mire* ‘see’; *yo (me) recuerdo* ‘I remember’; *(me) cago en* ‘damn’.
2. I also discarded frequent instances of repeated discourse (unless they preserved their lexical meaning and were not used as mere discursive markers), such as *y ya está* ‘and that’s all’; *lo que pasa (es) que* ‘what happens is that’; *y se terminó / y se acabó* ‘and that’s all’; *(no) ves que* ‘can’t you see that?’; *no sé qué/cuántos* ‘I don’t know what/how many’.

3. Cases of *venir* ‘to come’ in constructions such as *el domingo que viene* ‘next Sunday’ were also discarded.
4. Instances of “collective repeated discourse” such as idioms, sayings, songs or recited texts, names of games or nicknames were also discarded. Recited texts composed by the speaker were retained.
5. Examples which were ambiguous between two or more reflexive constructions were discarded.
6. All instances of verbs in the passive/impersonal reflexive construction or with the so-called spurious *se* (where *se* is an allomorph of the non-coreferential dative pronoun *le* in a clitic cluster) were discarded, since the RM is not compatible with these constructions.
7. Verbs contiguous to an unintelligible fragment or followed by an interruption in the discourse which prevented the appropriate interpretation of the example were not considered. Similarly, contexts labelled “not completely intelligible” in the transcription were not taken into consideration.

Table 2. Number of interviewed localities per province in subcorpus E

Province	Number of localities	Province	Number of localities
Álava	2	La Coruña	3
Albacete	1	La Rioja	1
Alicante	1	León	5
Almería	1	Lugo	2
Asturias	4	Madrid	1
Ávila	1	Málaga	1
Badajoz	1	Murcia	1
Burgos	1	Navarra	1
Cáceres	2	Orense	3
Cádiz	2	Palencia	3
Cantabria	5	Salamanca	1
Castellón	1	Segovia	1
Ciudad Real	2	Sevilla	1
Córdoba	1	Soria	2
Cuenca	1	Teruel	1
Granada	1	Toledo	1
Guadalajara	1	Valencia	1
Guipúzcoa	1	Valladolid	1
Huelva	1	Vizcaya	1
Huesca	1	Zamora	1
Jaén	1	Zaragoza	2

2.2.3 Analysis of the data

The data extracted from each interview according to the method presented above amounted to a total of 39,323 examples. These were analysed according to a number of parameters; the locality of the interview, the presence of the RM, the attested verb, the type of diathesis alternation, the semantic category of the verb according to Kemmer's classification (see Section 1.4.4), the animacy of the subject, the verb person, the verb tense, and whether the verb appears in a verbal periphrasis.

Once all the interviews were analysed, the results were collated in a single table, which was reviewed several times in order to reduce errors. In what follows, I list the possible values for each parameter.

1. Presence of the RM: Yes ~ No.
2. Attested verb: verb lexeme. As a rule, I did not group as the same lexeme morphological (prefixed) variants such as *llenar* – *enllenar*, *bajar* – *abajar*, since it is not clear that they share the same syntactic behaviour.
3. Type of diathesis alternation: Anticausative ~ Conversive ~ De-objective ~ Absolute ~ Antipassive without explicit object ~ Antipassive with explicit oblique object ~ Non-reversible ~ Intransitive / ~ Transitive ~ Auxiliary. Whenever the classification of a verb was unclear, I followed DRAE's classification of its transitivity properties (transitive, intransitive or both) and the *Diccionario de construcción y régimen de la lengua castellana* (Cuervo 1994 [1886]) for information on the existence of prepositional objects.
4. Semantic category of the verb: Naturally reciprocal ~ Autobenefactive ~ Grooming ~ Grooming (metaphorical) ~ Body process ~ Non-translational motion ~ Non-translational motion (metaphorical) ~ Change in body posture ~ Change in body posture (metaphorical) ~ Translational motion ~ Translational motion (metaphorical) ~ Spontaneous event ~ Spontaneous event (metaphorical) ~ Cognition event ~ Perception event ~ Emotion event ~ Pseudocopulative ~ Other.
5. Animacy of the subject: Animate ~ Inanimate. Animate subjects include humans, animals that are alive and collective nouns with a human interpretation. Inanimate subjects include all remaining possibilities (including vehicles, institutions, and dead animals).

In the analysis of specific categories, a number of extra parameters were analysed:

1. For autobenefactive constructions and transitive verbs (see Chapter 6), the delimitation of the object was considered. The types of objects that are considered either delimited or non-delimited are listed in Section 6.5.

2. For intransitive and transitive verbs of motion (Section 5.4, 6.10), the presence of an explicit direction adjunct was considered, differentiating between adverbs, prepositional adjuncts (and the preposition used) and the absence of direction adjuncts.
3. For anticausative and some intransitive and transitive verbs (Chapters 4, 5, and 6), the presence or absence of a pronominal dative was considered, as well as its thematic role.
4. For some specific subtypes of constructions or verbs, other parameters were considered. These are explained in the appropriate sections.

2.2.4 Note on spelling and linguistic representation

The spelling of Spanish and Galician used in this book requires some explanation. For the examples in Spanish, the COSER criteria are followed, which represent all morphosyntactic variants of the informant, and also some phonetic variants (mostly those that involve the addition or removal of a phonetic segment).⁵ For Galician examples, the official spelling is used, this sometimes adapted to represent the morphosyntactic and phonetic variants of the speakers.

In the lists of verbs attested in the different constructions, Spanish is mainly used, sometimes also representing the use of cognate verbs in Galician, in order to offer an easier way of reading the data, and also to unify the data to allow for the use of statistical software. Obviously, Galician verbs with no Spanish cognate or with a Spanish cognate that was not attested in the interviews are offered in Galician.⁶

5. Most conversational marks were, nevertheless, omitted. I have only preserved those that refer to a non- or almost non-intelligible fragment ([A-Inn] and [A-PIn:] respectively) and the ones identifying the interviewer or informant turn (E: and I: respectively). I have also preserved the label [NP], which replaces proper names in order to preserve anonymity. Last, because speech often shows repetitions and digressions, the examples provided in this book have sometimes been shortened for the sake of clarity and brevity, by removing linguistic material that is not crucial for the interpretation of the example.

6. Dialectal derivative morphological variants of verbs are offered, since the presence of a given prefix might affect the verb's syntactic behaviour. These variants are not frequent enough to explore their effect in the presence of the RM, but I offer this information anyway, in the hope that it might be useful for future work on morphological variation.

2.3 Other sources

In this section I will briefly describe the remaining data sources I have used, namely, a visual questionnaire (Subsection 2.3.1) and written texts (Subsection 2.3.2).

2.3.1 Video questionnaire

This study was originally part of a larger project which included syntactically reflexive and reciprocal constructions, which are under-represented in the COSER interviews. Because of this, it was necessary to design an elicitation task that would provide enough examples of these constructions to conduct a quantitative analysis. For this purpose a video questionnaire was designed.

The advantages of collecting data through a questionnaire are obvious: questionnaires produce highly controlled and comparable examples. Traditional dialectology relied mainly on translation questionnaires (Chambers & Trudgill 1980). However, these imply a strong effect of the question on the answer, casting doubts on the validity of the data. Indirect questions or visual aids, such as the drawings or herbarium used in the data collection for the Linguistic Atlas of the Iberian Peninsula, are not well suited for the production of complete sentences, that is, for the study of morphosyntax (Sanchis Guarner 1953). Written questionnaires designed to elicit introspective judgements have been used in the study of constructions with *se* (Nishida 1994; D'Introno & González & Rivas 2007; Garita & Rojas 1993), but there are some convincing arguments against their use. On the one hand, it is well known that many factors other than grammaticality might affect grammaticality judgements (see for instance Schütze 1996; Riemer 2008). On the other hand, the type of informant sought for in this study typically has a low competence in literacy, which adds an extra difficulty to the task.

Therefore, I decided to follow a technique often used in language documentation and to prepare a questionnaire based on video stimuli.⁷ I was first inspired by the stimulus kits available in the *L&C Field Manuals and Stimulus Materials* designed in the Max Planck Institute for Psycholinguistics,⁸ which consist of a number of video clips specifically designed to obtain data on several linguistic phenomena. While I was able to actually use some of these clips,⁹ it was necessary

7. Visual stimuli (drawings) have been also used in the study of reflexive constructions in the learning of Spanish as a second language (Zyzik 2006).

8. Available at <http://fieldmanuals.mpi.nl/>.

9. Specifically, I used 14 videos from the questionnaires *Reciprocal constructions and situation type* (N. Evans et al. 2004) and *Cut and break clips* (Bohnenmeyer, Bowerman & Brown 2001). They are videos 11, 12, 14, 19, 20, 70, 71, 72, 112, 144, 174, 175 y 181 in my questionnaire (see Appendix 2).

to create more clips that would meet the specific needs of this study. Because the data from this questionnaire are sporadically used in this study, I will briefly explain how it was designed.

In order to choose which verbs should be included in the questionnaire, I analysed the 1,071 most frequent verbs of Spanish (Almela Pérez, Cantos & Sánchez 2005). I introspectively categorised these verbs according to the possibility that the verb participated in any of the diathesis alternations identified in the literature. Table 3 shows the number of verbs that were identified for each category. I also codified the animacy of subject and object for these verbs, both in the basic (i.e. unmarked) construction and in the derived (often marked) diatheses, although this parameter was only used to distinguish further subtypes for verbs participating in the anticausative diathesis, which were the most productive.

Table 3. Number of frequent verbs per reflexive construction and diathesis alternation

Diathesis alternation	Number of verbs
Reflexive	218
Reciprocal	384
Anticausative	325
Conversive	20
Antipassive	26
De-objective	22
Non-reversible	12
No change of valency (transitive and intransitive)	108
Non-reflexive	281

Once all verbs were classified, a subset of verbs from each category was selected. This selection was made according to a practical criterion: the ease of recording a video stimulus to elicit a sentence with such a verb. In some cases, however, extra help for the recording was needed, and for these I created animated videos using the free software GoAnimate.¹⁰ Table 4 shows the number of videos per reflexive construction that were recorded for the questionnaire. The videos were ordered following a semantic criterion, grouping videos that depict similar actions. Animated clips, however, were all placed together at the end of the questionnaire, since they were quite demanding for informants; viewing them at the end of the sessions reduced the stress that they added to the interview. A list of the videos is provided in Appendix 2.

10. Available at <http://goanimate.com/>.

Table 4. Number of videos per reflexive construction and diathesis alternation

Diathesis alternation	Number of verbs
Reciprocal	26
Anticausative	66
Conversive	2
Antipassive	3
De-objective	2
Non-reversible	2
No change of valency	60
Non-reflexive	28

A total of 43 interviews were conducted using the questionnaire. I personally conducted 42 of them, while the remaining one (Santaella, Córdoba, QT 047) was conducted by a team of Spanish Philology students as part of the COSER campaign to interview Córdoba (Andalusia) in 2011.

The interviewed speakers share the same characteristics as the COSER informants: they are elderly rural speakers that have always or almost always lived in their native village, where they were interviewed. The same criteria on the chosen language as in the COSER apply – all speakers from Asturias chose to answer in Spanish, while some speakers from Galicia answered in Galician and some others did so in Spanish. As opposed to the COSER informants, most of the questionnaire informants had previously been contacted, since this kind of interview required a less flexible, and hence pre-organised, situation (both a table and a socket were needed, so that the videos could be comfortably played on a computer).

A further difference with the COSER interviews is that the speakers were aware of the linguistic nature of the task. They were told that it was part of a lexical enquiry to study the different words used across Spain for some concepts. Since the presence or absence of the reflexive construction is quite an abstract matter, it seems reasonable to assume that no informant “discovered” the real goal of the questionnaire (none suggested otherwise).

The task (simply describing the videos that the informants were shown) was explained at the beginning of the interview. The videos were played afterwards with no further explanation, unless the description provided by the informant did not contain the verbs or contexts sought. In such cases, I tried to avoid the use of the verb sought in the explanation or questions. Instances where this was not possible and the informant uses the same structure used by the interviewer have been discarded. The interview was recorded using the computer, which reduced the attention of the speaker to the fact that they were being recorded (although they were aware of it and have given their consent). When the informant did not accept to be video recorded, only audio was recorded.

The distribution of the interviews is shown in Map 3. The list of localities is given in Appendix 3. Map 3 shows that most interviews were conducted in the north-western area, which has the most differential situation from Standard Spanish according to the literature. Since no large differences were found between the south of this area and the rest of the territory, the original plan of interviewing at least one village per province was abandoned, in order to save time and funds.



Map 3. Localities interviewed with the questionnaire

Thanks to the questionnaire I collected 17,154 (potentially or actually) reflexive examples. In this study, however, I will only use the questionnaire data to supplement those areas where the COSER data lack sufficient information.

2.3.2 Written examples

In order to illustrate cases not attested in the COSER or the questionnaire data, I sometimes provide examples from written sources. I have tried to avoid made-up examples when possible.¹¹ In the same spirit, I have tried to avoid the use of the asterisk (*), unless it reflects an example taken from the literature or it indicates a theoretic prediction. Thus, I use either the number sign (#) to indicate that an example is semantically inappropriate or question marks (??) to indicate that I find a given example intuitively strange.

11. Probably doing what Itkonen (2005, apud Willems 2012) calls “inessential use of a corpus”, that is, using corpus data to illustrate clear examples. I prefer to do so, however, especially since so many examples deemed to be clear are subject to at least geographical variation.

The written examples that I cite have been obtained from Twitter. The advantage of Twitter as a data source is that it documents a linguistic register that is close to the spoken language (Estrada & De Benito Moreno 2016, 2018). The type of speaker whose language can be documented on Twitter, however, is completely different from the type of speaker documented in COSER and the questionnaire data, which must be taken into account.

I maintain the original spelling from the tweets, which is sometimes not only “original”, but also very innovative. Usernames and nicks are not provided, to preserve speakers’ privacy.¹²

2.4 Technical notes

In what follows I briefly describe the statistical methods (Subsection 2.4.1) and geographical representations (Subsection 2.4.2) used in this study.

2.4.1 Statistical methods

Quantitative methods used in this study involve, on the one hand, descriptive statistics, i.e. summary statistics that apply exclusively to the analysed sample and are not generalised to the population. Since most of the data used are categorical, these summary statistics are typically relative frequencies, given either as probabilities (from 0 to 1) or percentages (from 0 to 100). Sometimes these summary statistics are given in contingency tables, while at other times they are provided in stacked bar plots, which allow for a more graphic representation of the frequency differences between two levels of a factor (i.e. two values of an independent variable). At any rate, the absolute figures are always provided, so that the reader can evaluate the results together with the magnitude of the sample on which they are based. Moreover, I also give an explicit warning when the sample is too small.¹³ When large amounts of data are analysed, thus allowing for the aggregation of the data by verb lexeme (cf. Chapter 4), I have done so, obtaining in this way a numerical variable. The visualisations offered for such data typically aim to show the distribution

12. Together with the tweet, I cite the date of publication and, where possible, the geographical information of the tweet (for which I rely on the information users give in their public profiles or on the automatic geolocation of the tweet, where available).

13. Please note that sometimes the absolute figures of a given statistic might be lower than the absolute number of cases collected of a dependent variable, due to the fact that not all independent variables can always be codified (there are ambiguous cases, interruptions, etc.).

of the data, by using boxplots and dotplots – all such visualisations are explained in the relevant sections.

Whenever inferential statistics are used (i.e. statistics that aim to generalise the behaviour of the sample to the population), I rely mainly on generalised mixed linear models, which allow for the inclusion of the variables responsible for the lack of independence of the data. The reference level of the independent variable in these models is always the absence of the RM, which means that the values provided by the models always refer to the presence of the RM (compared to its absence). In order to properly interpret these models, it must be remembered that the values given in the summary of the model are always to be interpreted in comparison with the respective reference level. I consciously avoid the chi square test (and Fisher's exact test), since these are not suited for observational data such as those obtained from the exhaustive analysis of corpora.¹⁴ When mixed models are not suitable (often due to low quantities of data), I have resorted to these tests, but performed on 1000 random samples (bootstrapping) of the data, to preserve the independency criterion (see Sections 4.4 and 5.7).

Finally, statisticians have recently warned against the misuse of the concept of statistical significance in science and of p-values (Nuzzo 2014; Wasserstein & Lazar 2016; Colquhoun 2017), which do not measure the probability that the proposed hypothesis is true, as most authors seem to think, but rather measure how incompatible the data are with the null hypothesis. Thus, I avoid relying on p-values, which are typically misinterpreted, and provide instead confidence intervals for the effect size whenever inferential statistics are used.

All the data files and R-scripts used for generating the tables, plots and statistical models can be found in the Github repository found at https://github.com/Carlotadbm/Middle_voice_IberoRomance.

A final caveat regarding the size of the sample should be noted. Despite the fact that the whole sample of examples that this book deals with is quite large (see Chapter 2), the fine-grained level of detail of the analysis sometimes means that the number of examples is not enough to perform statistical analysis. This is acknowledged whenever it is the case. Most importantly, this does not impinge on the validity of quantitative methods or corpus data, which should never be considered a definitive tool on their own. Scientific knowledge cannot be built on the sole basis of quantitative analyses – it is crucial that interpretations of the data are backed up

14. These tests assume the independence of the data. Corpus data, however, are not independent, since the same speaker produces several of the same observations. Because corpus data involve free speech and not controlled experiments that allow for the pairing of observations, the independence requirement is not met. The same applies to other inferential statistics methods, such as confidence intervals.

by solid theories. Likewise, the exhaustive documentation of all possible linguistic forms is not granted by the examination of large corpora – low frequency forms might not be attested. Moreover, every single quantitative analysis must establish a dialogue with previous and future quantitative accounts based on similar data – replication is a pivotal tool of science, one that is too often disregarded in standard academic practices.

2.4.2 Geographical representation

Besides point maps to represent the location of a given variant, there are three further types of maps in this study that represent relative frequencies of the RM. Sometimes the Inverse Distance Weighted interpolation technique (IDW interpolation) is used, which projects the data from the points for which there is actual information to all other points on the map. This projection is made through an algorithm that takes into account the closest actual data-points and weighs them depending on their geographic distance. I have always used the closest seven points to produce the interpolations.¹⁵ The resulting frequencies are divided into several ranges and indicated through different colours or shades, as shown in the legend. An example of an interpolated map is offered in Map 4, where the points from which we actually have data are indicated in black.

A second technique to represent relative frequencies is through pie charts, as illustrated in Map 6. Each colour indicates one variable (as specified in the legend), while the size of the pie or the bars varies according to the total number of cases. This representation is not as smooth as IDW interpolated maps, but is more appropriate for those cases where no clear geographical pattern is found. It is also useful when few data are available, since the total of examples affect the validity of relative frequencies.

Finally, when absolute frequencies of the presence of the RM could not be compared to its absence, I have normalised the total values according to the duration of the interviews (see Subsections 6.2.2, 6.3). If no clear geographical pattern arises, I represent the normalised values with symbols whose size varies according to the normalised frequency value as shown in the legend (see Map 16).

15. Catalonia has been removed from the map of Peninsular Spain, since there is no data at all from this area in my sample.

CHAPTER 3

The middle voice in Spanish

3.1 Introduction

In this chapter I test the two main typological approaches to the middle voice (see Section 1.4) with my data. As noted above, some authors emphasise the relationship between middle voice and diathesis – the former tends to mark several kinds of the latter, typically the passive, conversive, anticausative, reflexive, reciprocal, antipassive and autobenefactive diatheses (Kulikov 2011, 2013). Kemmer (1993), however, adopts a different perspective. For her, the middle voice is semantically motivated, as it indicates the low elaboration of the event. A manifestation of this concept is subject affectedness, a feature traditionally associated with the middle voice (see Section 1.4).

According to Klaiman's (1991) terminology, then, Kulikov conceives of the middle voice as a derived voice, while Kemmer (1993) depicts it as a basic voice. What makes this theoretical contradiction especially interesting is the fact that both accounts aim to have general validity. Is the middle voice a derived or a basic voice? This is a difficult question to answer, since the two views of the middle are not independent from each other, as noted by García-Miguel (2001):

In general, transitivity is reduced every time that no energetic relationship is established between two distinct and clearly specified participants. Thus, it can be understood that the relative elaboration of the event is reduced in the detransitivising diatheses [Kemmer 1993]. (García-Miguel 2001: 379–380, my translation)

This chapter is devoted to testing which of these conceptions of the middle voice better fits the characteristics of the RM in the varieties under study. Section 3.2 deals with the relationship between middle marking and changes of diathesis (middle voice as a derived voice). Section 3.3 deals with the relationship between middle marking and the event's semantics (middle voice as a basic voice). A comparison of both accounts, which concludes that middle marking in these varieties is more strongly associated with diatheses changes in quantitative terms, is offered in Section 3.4.

3.2 The middle voice as a derived voice

The middle voice tends to encompass a number of different diatheses, i.e. the passive, conversive, anticausative, reflexive, reciprocal, antipassive and autobenefactive diatheses (Kulikov 2011, 2013). All these diatheses can be marked by the RM in the varieties under study. However, in order to compare Kulikov's account with that of Kemmer (1993), we need to measure how productively the RM marks these diatheses. In what follows I will only discuss the conversive, the anticausative, the antipassive and the de-objective diatheses. Kulikov groups the latter with the antipassive, but I will study it separately, following Geniušienė (1987).

The exclusion of the reflexive and reciprocal diatheses is justified on several grounds. First, they are fundamentally different from the other diatheses, since they are operational diatheses that preserve the thematic roles of the predicate (Kulikov 2011). Second, Kemmer (1993) aims to offer typological arguments for the distinction between the reflexive and the middle – in order to compare both accounts the same data must be analysed. Finally, these are the etymological values of the RM, and the RM preserves its pronominal characteristics, so an argument can be made to distinguish these two functions of the RM from the others, which are later and derived developments (see Section 1.3)

The exclusion of the passive and the autobenefactive is justified on similar grounds. First, Kemmer (1993) devotes little attention to these two diatheses in her account – she considers the passive a different category and has little documentation on indirect constructions. Moreover, the passive was the last development of the RM and is fundamentally different of other diatheses in being non-paradigmatic – RMs *me*, *te*, *nos* and *os* are excluded from this construction. Also, indirect reflexive constructions have a direct object which cannot be identified with the RM, which makes them crucially different to direct reflexive constructions. I will deal with indirect constructions in Chapter 6.

So, I will focus here on reflexive constructions where the RM is paradigmatic (i.e. inflects for all persons) and appears in intransitive constructions. Three types of intransitive constructions can be identified: (1) valency decreasing diatheses (such as the anticausative, 3.2.1, the absolute, 3.2.2, and the de-objective, 3.2.3); (2) valency demoting diatheses (such as the conversive, 3.2.4, and the antipassive, 3.2.5), and (3) constructions with no change of valency (3.2.6 and 3.2.7).

3.2.1 Subject deletion diathesis: Anticausative verbs

The causative–anticausative alternation is the change of diathesis most commonly associated with the RM in the literature. In anticausative verbs the subject of the transitive scheme is removed from the verbal structure and the object is promoted to subjecthood (Kulikov 2011). A typical example of this alternation is given in

(36), where the equivalence between the object in the causative structure and the subject in the intransitive construction is very clear. The highly productive attachment of the RM to anticausative verbs has motivated the claims that the RM is an intransitivising device.¹

(36) (Vicente Mateu 2001)

- a. Juan rompió la mesa.
Juan break.PST.3SG the table
'Juan broke the table.'
- b. La mesa se rompió.
the table REFL.3 break.PST.3SG
'The table broke.'

While it is clear that the table did not initiate the event in (36b), when the subject of the intransitive version is animate (37b) it is both the initiator of the event and the participant affected by it:

- (37) a. María escondió a Juan.
María hide.PST.3SG to Juan
'María hid Juan.'
- b. Juan se escondió.
Juan REFL.3 hide.PST.3SG
'Juan hid.'

Because of this difference, Geniušienė (1987) considers examples like (36) as cases of subject suppression, what she calls “decausative” verbs; but she classifies examples like (37) as object suppression, what she calls “autocausatives”. Although in these two cases these labels capture very well the semantic differences in the role of the subject, they fail to capture the similarities between examples like (38), since according to this classification (38a) and (38b) belong to different classes, despite being the same verb.

- (38) a. La tetera se movió con el temblor.
the teapot REFL.3 move.PST.3SG with the earthquake
'The teapot moved due to the earthquake.'
- b. María se movió incómoda en su asiento.
María REFL.3 move.PST.3SG uncomfortable in her seat
'María moved uncomfortably in her seat.'

1. As Haspelmath (1993) notes, the direction of morphological derivation of causative–anticausative pairs in the languages of the world varies greatly. As opposed to the tendency (found in the languages under study here) of marking the anticausative pair, there are also languages that display marking in the causative pair, (cf. the Basque pair *ikasi* ‘to learn’, *irakatsi* ‘to teach’, derived from the former through the causative affix *-ra-*, see Euskaltzaindia, 1987–2005). Some other languages show a strong tendency not to mark any of the members of the pair, producing the so-called labile verbs (cf. English, *John broke the table* / *The table broke*).

Furthermore, Geniušienė (1987) considers a class of “autocausative objective” verbs, where the subject is deleted and the original patient not only promotes to subjecthood but also becomes an actor who changes their own state (see (39)):

- (39) a. María desató al perro.
 María unleash.PST.3SG to.the dog
 ‘María unleashed the dog.’
 b. El perro se desató.
 the dog REFL.3 unleash.PST.3SG
 ‘The dog got unleashed.’
 c. *El perro desató al gato.
 the dog unleash.PST.3SG to.the cat
 ‘The dog unleashed the cat.’

I do not believe that this proliferation of classes is necessary, since the semantic differences in the role of the subject depend on its semantic characteristics: animate participants can be seen as responsible for the events they experience, while inanimate participants typically cannot. The contrast in (39) seems to be due to the fact that some verbs prefer agentive subjects in the transitive counterpart and not all animate participants can be conceived of as such. The “anticausative” label captures the fact that in all these cases the subject is affected by the action conveyed by the verb, with no external cause codified in the verb’s valency.

There is a large body of literature on the restrictions of the anticausative construction, most of which highlights the fact that only verbs that allow non-agentive subjects can show the alternation. In the literature on Spanish anticausatives, for instance, it has been noted that “causative [transitive] verbs with unaccusative uses are those where the event conveyed by the predicate might occur spontaneously, with no voluntary participation of an agent” (Mendikoetxea 1999a: 1591, my translation). In Schäfer & Vivanco’s (2015) words: “Only transitive verbs that do not restrict the θ -role of their external argument to agents enter the causative alternation” (Schäfer & Vivanco 2015: 207). This adequately predicts the contrast in (40), since transitive *cortar* requires an agentive subject (Mendikoetxea 1999a).² This contrast has led to the claim that *cortarse* cannot have an anticausative reading in Spanish (Mendikoetxea 1999a, Schäfer & Vivanco 2015).

- (40) (Mendikoetxea 1999a)
 a. El panadero cortó el pan.
 The baker cut.PST.3SG the bread
 ‘The baker sliced the bread.’

2. This is because it specifies the way the action is produced (see Mendikoetxea 1999a).

- b. *El pan se cortó (por sí solo).
 The bread REFL.3 cut.PST.3SG by REFL.3 alone
 ‘The bread cut.’

Such a claim, however, does not seem to hold for every instance of *cortarse*, as can be seen in Examples (41a), which is unambiguously anticausative, as the comparison with the purely reflexive reading of (41b) shows. This example does not necessarily refute the claim that transitive *cortar* requires an agentive subject (see (41c)), but does suggest that the properties of the transitive version of a verb (such as requiring an agent)³ do not necessarily apply to the anticausative verb itself or, at least, that these restrictions cannot be considered in a one-to-one comparison at the sentential level.⁴

- (41) a. No sé cómo me he cortado con el
 not know.PRS.1SG how REFL.1SG have.1SG cut.PTCP with the
 aspirador
 vacuum.cleaner
 ‘I don’t know how I’ve cut myself with the vacuum cleaner.’
 (December 4th 2019, Madrid, Spain. Tweet)
- b. @usuario tenia envidia de que me hayan
 @user have.PST.3SG jealousy of that DAT.1SG have.SBJV.3PL
 tenido que coser en el dedo y se ha
 have.PTCP that sew.INF in the finger and REFL.3 have.PRS.3SG
 cortado aposta para que le cosan
 cut.PTCP willingly for that DAT.3SG sew.SBJV.3PL
 ‘@user was jealous that they had to sew my finger and he’s cut himself so that they would sew him too.’ (September 13th 2012, Zaragoza, Spain. Tweet)
- c. #El cuchillo me ha cortado.
 the knife ACC.1SG have.PRS.3SG cut.PTCP
 ‘The knife has cut me.’

The anticausative examples of *cortar* are not limited to animate subjects, as (42) shows. Note that a transitive equivalent of these examples with a cause subject appears to be as unlikely as (41c).

3. As Mendikoetxea (1999a) notes, *cortar* also admits instrumental subjects, but those imply the presence of an agent.

4. Or to its ability of fitting in the anticausative construction. This fits well within Goldberg’s “surface generalization hypothesis”: “There are typically broader syntactic and semantic generalizations associated with a surface argument structure form that exist between the same surface form and a distinct form that is hypothesized to be syntactically or semantically derived from.” (Goldberg 2002: 329, apud García-Miguel 2010: 3).

- (42) Se cortó la cuerda de mi guitarra.
 REFL.3 cut.PST.3SG the string of my guitar
 ‘The string of my guitar broke.’ (October 3rd 2015, Lobos, Argentina. Tweet.)

This shows that a view that relies on a strict derivation process between the anticausative and the causative example is not always appropriate. Reflexive anticausatives are derived from unmarked causatives from a morphological point of view, which allows us to talk about derived intransitivity or derived diathesis. This concept is adequate from a historical point of view,⁵ but it does not necessarily entail the strict transfer of properties from the causative verb to its anticausative counterpart. The case of *cortar* proves that the anticausative construction has its own requirements (like rejecting external agents) and that some verbs can fit these requirements in the anticausative construction even if they do not show the same restrictions in the causative construction. As noted above, the most important distinctive feature of anticausative verbs is that they show a subject affected by the verbal action, without the direct intervention of external causes. Let us now turn to the analysis of the data.

I collected a total of 8,862 examples of anticausative verbs from the COSER interviews. These examples illustrate 480 different verbs, which can be found in the Appendix 4. An example of the same anticausative predicate, both marked and unmarked, is provided in (43), with an example of its causative counterpart given in (43c).

- (43) a. [the wool] Hay que tenderlo al sol para
 have.PRS.3SG that hang.out.INF-ACC.3SG to.the sun for
 que se seque.
 that REFL.3 dry.SBJV.3SG
 ‘One has to hang it out in the sun for it to dry.’
 (Madrigal de las Altas Torres, Ávila, COSER 0609)
- b. [the grass] Había que darle la vuelta, para que
 have.PST.3SG that give.INF-DAT.3SG the turn for that
secara por el otro lado.
 dry.SBJV.3SG for the other side
 ‘One has to turn it around for it to dry on the other side.’
 (Zas, La Coruña, COSER 2403)
- c. [chorizos] Nosotros los secábamos en el campo al aire.
 we ACC.3PL dry.PST.1PL in the field to.the air
 ‘We used to dry them in the field, outdoors.’ (Espera, Cádiz, COSER 1107)

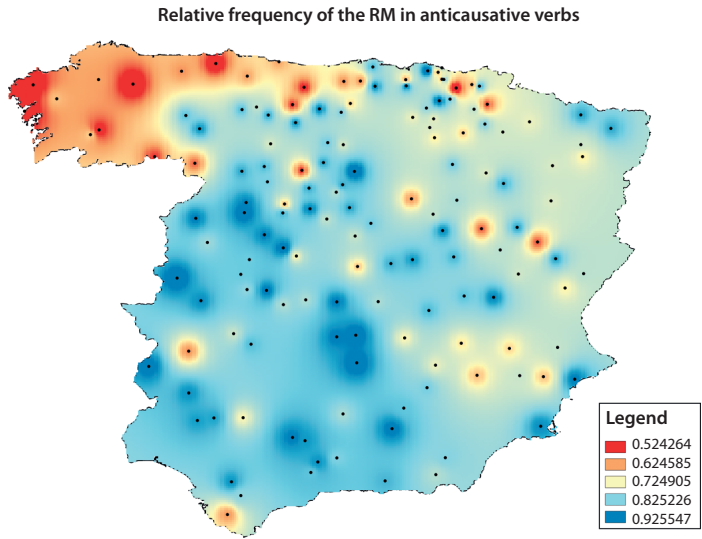
5. Although not necessarily in every case. As Cartagena (1970) notes, the productivity of these alternations enables causative versions to be created from non-reversible reflexive verbs (*se suicidó ~ la suicidaron*). Similar examples can be found in Galician (Domínguez Oroña 2015).

As shown in Table 5, the use of the RM is the most productive way of marking an anticausative predicate, being used in 77.5% of the cases, as opposed to the 22.5% of the examples that show a labile alternation. Most verbs provide marked examples in the corpus (452/480), while unmarked examples are attested for far fewer verbs (118/480).⁶

Table 5. Frequency of the RM with anticausative verbs in the COSER data

	Marked (RM)	Unmarked (No RM)	Total
Anticausative examples	6,868 (77.5%)	1,994 (22.5%)	8,862 (100%)
Anticausative distinct verbs	452	118	480

The geographical distribution of the total sample is quite insightful. Map 4 shows that unmarked (labile) anticausative verbs are far more frequent in the north-western area that includes Galicia and Asturias. However, it also shows that the presence of the RM is not systematic in the rest of the study area. A more detailed study of anticausative verbs, which was the most prominent class in our corpus, is offered in Chapter 4.



Map 4. Relative frequency of the RM with anticausative verbs in the COSER data

6. Because verbs, when considered as a whole, might show variable marking, the figures in the second row of Table 5 do not add up to the total. This is also why no percentages are offered. These observations are valid for the remainder of this chapter.

3.2.2 Object-deletion diatheses: Absolute constructions

In object-deletion diatheses, the direct object is completely removed from the verb structure and cannot be encoded even as an oblique object (as in the antipassive diathesis, see 3.2.5). Reflexive marking of verbs with null objects is found in Slavic languages, such as Russian (Kulikov 2011); and Baltic languages, such as Latvian and Lithuanian (Geniušienė 1987). In these languages, de-objective verbs marked with the RM seem to have a property reading (see also Nedjalkov 2007c):

- (44) Lithuanian (Geniušienė 1987: 83–84)
- a. Berniuk-as muša vaik-us
 Boy-NOM beat.PRS.3SG chils-ACC.PL
 ‘The boy beats children.’
- b. Berniuk-as muša-si
 Boy-NOM beat.PRS.3SG-REFL
 ‘The boy fights (is pugnacious).’

This use of the RM is not usually described for Romance languages, and none of the examples given above seem to have parallels in the varieties under study. In these varieties, on the contrary, it seems that the mere absence of the direct object (together with an imperfective tense) is enough to trigger the property reading (see (45)). These uses are normally called absolute uses (NGLE 2009: 34.4).

- (45) a. Sra si usted ve q su hijo pega, muerde o
 lady if you see that your child hit.PRS.3SG bite.PRS.3SG or
golpea llámele la atencion
 beat.PRS.3SG call.IMP-DAT.3SG the attention
 ‘Lady, if you see that your child hits, bites or beats [others], tell him off.’
 (March 16th 2013. Tweet.)
- b. #grandesmentiras mi perro no muerde
 big.lies my dog no bite.PRS.3SG
 ‘#biglies my dog doesn’t bite.’ (December 31st 2013. Tweet.)

Nonetheless, a few examples of reflexively marked verbs found in the COSER interviews appear in absolute contexts – that is, with non-explicit objects that can be recovered thanks to the context. Some are shown in (46) and (47). Because these uses of the RM were unexpected, only marked examples were collected, which is a methodological difference with all other diatheses studied here. Only 21 examples were documented, from ten different verbs. A total of six examples appear with perception verbs: *ver* ‘to see’, *mirar* ‘to look’ and *oír* ‘to hear’ (see (46)). We can even hint at a possible eastern ascription of these uses (which were documented in Valencia, Huesca and Soria, that is to say, all in eastern varieties). At any rate, the

fact that *mirarse* is recorded in two villages of Huesca does suggest that it may be a regional use, as does the fact that the two examples of *verse* (see footnote 35) were documented in the province of Valencia. Example (46a) is especially interesting, since it actually has a property reading – it means that the speaker has bad eyesight.⁷ This is reminiscent of the absolute uses which the RM is associated with in Slavic and Baltic languages described above. The remaining examples with perception verbs, however, do not show this property reading (46b), (c).

- (46) a. Pero ahora entre que no me veo bien y to, pues
 but now between that no REFL.1SG see.PRS.1SG well and all then
 no leo.
 no read.PRS.1SG
 ‘But now, since my sight is bad and all that, well, I don’t read.’
 (Enguera, Valencia, COSER 4310)
- b. [los niños jugaban fuera] Y mi madre haciendo la comida y
 and my mother making the food and
 las cosas de casa. Se miraba y: “Juega aquí, juega
 the things of house REFL.3 look.PST.3SG and play.INF.2SG here play
 aquí, hija mía”.
 here daughter mine
 ‘[the children were playing outside] And my mother was preparing the
 food and doing housework. She looked [at us] and: “Play here, play here,
 my daughter”’
 (Bandaliés, Huesca, COSER 2207)
- c. [E2: O sea que antes decían que el agua era mala para los enfermos.]
 Pero no te oyes, le privaron el agua
 but no REFL.2SG hear.PRS.2SG DAT.3SG take.away.PST.3PL the water
 al niño y se le murió el niño.
 to.the child and REFL.3 DAT.3SG die.PST.3SG the child
 ‘[E2: So they used to say then that water was bad for sick people.] But,
 aren’t you listening? They took water away from the child and the child
 died.’
 (Beratón, Soria, COSER 3924)

Other verbs include *aguardar* ‘to wait’, *apuntar* ‘to aim’, *dejar* ‘to leave’, *mamar* ‘to nurse’, *poner* ‘to put’ and *terminar* ‘to finish’. The examples in (47) do not show this property reading either. In (47a) it might be possible to hint at a semantic difference between *mamarse* ‘start nursing’ and *mamar* ‘nursing’, but more examples are needed. The case of *aguardar* ‘to wait’ (see (47b)) may be related to the presence of

7. Although this is the only example of *verse* with this property reading in the interviews I analysed for this study, I was lucky enough to participate in a COSER interview run in Pedralba (also in Valencia, COSER 4318), where a similar example was collected. In this interview, the speaker, referring to her husband and in order to explain his vision problems, said: *No se ve*.

the RM with *esperar* ‘to wait’, since they have very similar meanings.⁸ At any rate, note that there is intra-speaker variation, as shown by (47a), (b), (c). The example of *terminar* (in (47d)) is also unexpected and difficult to interpret, although it could be linked to the fact that *terminar* takes the RM in several contexts (both as an anticausative and a transitive verb). Unfortunately, the few examples we have of all these uses do not allow for an in-depth analysis, although they may be indicative of new trends in the development of the RM. At any rate, more research and data are needed to explain the presence of the RM in these examples.

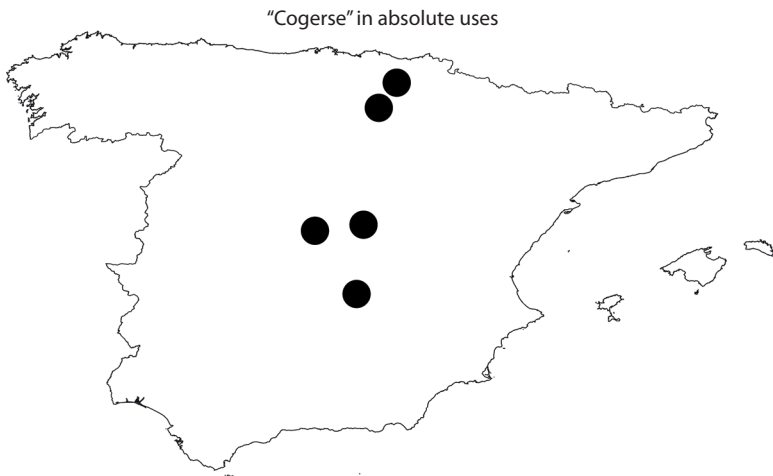
- (47) a. [había que quitarles los corderos a sus madres] Ya a un tiempo
already to a time
se juntaban otra vez, y ya no se
REFL.3 get.together.PST.3PL other time and already not REFL.3
mamaban, pero si no les apartabas, pues entonces
nurse.PST.3PL but if no DAT.3PL remove.PST.2SG well then
seguían mamando hasta que volvía a nacer
keep.PST.3PL nursing until that turn.PST.3SG to be.born.INF
otro hijo.
another son
‘[lambs need to be taken away from their mothers] After some time they got
back together [with their mothers] and they didn’t nurse anymore, but if you
didn’t take them away [from their mothers], well, then they kept on nursing
until another lamb was born.’ (Sieteiglesias, Madrid, COSER 2914)
- b. Aguárdate..., aguárdate.
wait.IMP.2SG-REFL.2SG wait.IMP.2SG-REFL.2SG
‘Wait, wait.’ (Muñoveros, Segovia, COSER 3707)
- c. Bueno, aguarde, hombre, que...
well wait.IMP.2SG man that
‘Well, man, wait, that ...’ (Muñoveros, Segovia, COSER 3707)
- d. Cuando ya se habían terminao las demás [mujeres]
when already REFL.3 have.PST3PL finish.PTCP the other women
hacían otros panes.
make.PST3PL other breads
‘When the women were already done, they baked other breads.’
(Jérica, Castellón, COSER 1307)

Finally, six examples of *coger* in absolute contexts were also documented. Interestingly, *coger* is used with a discursive meaning in these examples. This is a frequent use of this verb, used to introduce new actions in storytelling, as illustrated in

8. In this sense, it is interesting that the two examples documented are in the imperative, a context that favours the presence of the RM with *esperar* (see Section 5.7).

Example (48a). In these uses, *coger* is a transitive verb with a null object. Interestingly enough, the examples of *coger* with the RM in this construction are not completely exceptional, although definitely rare. There were six instances in the COSER data, scattered across five different localities (illustrated in (48b)). Although this is perhaps too speculative, it is noteworthy that the geographical distribution of these few examples show two clusters (see Map 5), at least pointing to a few areas where more research could be done on this topic.

- (48) a. Y cuando ya almorzaba y ya estaba yo tranquila
and when already eat.PST1SG and already be.PST1SG I calm
y los chiquillos y to, cogía y m'iba.
and the children and all take.PST1SG and REFL.1SG-go.PST1SG
'And when I had already eaten and I and the children had nothing else to
do, I used to leave.' (Almadén de la Plata, Sevilla, COSER 3806)
- b. [Si alguien de fuera llegaba al pueblo] Los viejos le hacían
the elderly DAT.3SG make.PST3PL
levantar a las tres de la mañana, pero si es hoy,
wake.up.INF at the three of the morning but if be.PRS.3SG today
se coge, se marcha y afuera.
REFL.3 take.PRS.3SG REFL.3 leave.PRS.3SG and away
'[if an outsider moved into the town] The elderly would awake him at three
in the morning, but if that were to happen today, he would go and leave.'
(Luzuriaga, Álava, COSER 0107)



Map 5. Discursive uses of *cogerse* in the COSER interviews

3.2.3 Object-deletion diatheses: De-objective verbs

Besides the rare absolute constructions described in the previous section, there is another class of reflexive verbs found in the varieties under study which can be related to their unmarked counterparts through the demotion of the direct object. They differ from the constructions just described in that the omitted object is not recoverable from the context, but it is lexicalised and therefore shows a fixed interpretation. These are the uses I will term “de-objective”. As opposed to those in Section 3.2.2, these contexts were collected exhaustively (that is, both marked and unmarked occurrences of these verbs were collected). Some examples are given in (49).

- (49) a. A la hora de pintar fijate en la orientación
at the time of paint.INF fixate.IMP.2SG-REFL.2SG in the orientation
de la habitación.
of the room
‘When painting, pay attention to the orientation of the room.’
(October 1st 2015, Venezuela. Tweet.)
- b. Explicate mejor. Escribiste tu eso?
explain.IMP2SG-REFL.2SG better write.PST.2SG you that
‘Explain yourself better. Did you write that?’ (June 15th 2015. Tweet.)
- c. Un día como hoy, pero hace 453 años, piratas ingleses se
a day like today but ago 453 years pirates English REFL.3
establecieron en la Isla de Tris
settle.down.PST.3PL in the island of Tris
‘A day like this, but 453 years ago, English pirates settled down in the Island of Tris.’
(October 26th 2011. Tweet.)

In these examples the role of the subject in the intransitive version of the verb is not the same as the role of the object in its transitive counterpart. They resemble the verbs that Geniušienė (1987) calls partitive object reflexive verbs (see (50)), in which an easily recoverable object is deleted from the structure. Similarly to (50), where the hypothetical objects are necessarily interpreted as *el pelo* ‘the hair’ and *la barba* ‘the beard’ respectively, the deleted object in (49a) can be interpreted as *atención* ‘attention’; as *ideas, conocimientos* ‘ideas, knowledge’ in (49b), and as *residencia* ‘residence’ in (49c).

- (50) a. Yo como que me cepillo y pa la cama porque bah
I like that REFL.1SG brush.1SG and to the bed because meh
‘Well, I brush my hair and I go to bed, because it’s not worth it.’
(June 14th 2015. Tweet.)

- b. Oye @usuario mira a ver si te
 hear.IMP.2SG @user look.IMP.2SG to see.INF if REFL.2SG
 afeitas que se te ve la barba desde aquí
 shave.PRS.2SG that REFL.3 DAT.2SG see.PRS.3SG the beard from here
 ‘Listen, @user, you should shave, because I can see your beard from here.’
 (June 8th 2015. Tweet. Reply to a tweet by another user,
 showing a picture of (a bearded) himself.)

The difference between the examples in (50) and the examples in (49) is that the deleted objects in (50) are easily recoverable for extralinguistic reasons (one normally brushes one’s own hair or shaves one’s own beard), while the recoverability of the deleted objects in (49) seems to be lexically determined, since one can fixate, explain and establish many things, but these de-objective verbs have specialised their meaning to refer to the subject’s attention, clarity and location respectively. Cano (1981) explains why the latter are not to be considered semantically reflexive in similar terms: “[In *explicarse* and *expresarse*] *se* could be analysed as a direct object, but it actually does not refer to the whole subject, but only to a part of it, such as “ideas”, “words”, etc.” (Cano 1981: 286, my translation).

All the possible “deleted objects” seem to belong to the personal sphere of the subject, being either mental entities or properties such as one’s residence. This, of course, explains the use of the RM, since the deleted object is something that belongs (broadly) to the subject. In this sense they are semantically connected to the notion of reflexive possessive datives (see Section 6.2.3). Geniušienė (1987) actually mentions that partitive-object reflexives may express psychological activities, and gives the following example: *restrain one’s emotion, tears* > *restrain oneself*. Nevertheless, although the semantic connection between partitive-object reflexives and de-objective verbs is clear, there are also reasons not to group all these examples in the same category, which clearly show that the examples in (49) do not belong in the reflexive class. On the one hand, they do not admit the *a sí mismo* reinforcement: see (51). On the other hand, they do not accept the expression of the object in an external possession construction (that is, with a possessive dative), as the indirect reflexive does – the explicit mention of the object requires an internal possession structure (with the possessive determiner, see (52)).

- (51) a. ??A la hora de pintar fíjate a tí misma
 at the time of paint.INF fixate.IMP.2SG-REFL.2SG to REFL.2SG self
 en la orientación de la habitación.
 in the orientation of the room
 ‘When painting, pay attention to the orientation of the room.’

- b. ^{??}Explicate mejor a ti misma. ¿Escribiste
explain.IMP2SG-REFL.2SG better to REFL.2SG self write.PST2SG
tú eso?
you that
'Explain yourself better. Did you write that?'
- c. ^{??}Un día como hoy, pero hace 453 años, piratas ingleses se
a day like today but ago 453 years pirates English REFL.3
establecieron a sí mismos en la Isla de Tris
settle.down.PST3PL to REFL.3 selves in the island of Tris
'A day like this, but 453 years ago, English pirates settled down in the Island
of Tris.'
- (52) a. Me cepillo el pelo ~ [?]Cepillo mi pelo.
REFL.1SG brush.PRS.1SG the hair ~ brush.PRS.1SG my hair
'I brush my hair.'
- b. Me afeito la barba ~ [?]Afeito mi barba.
REFL.1SG shave.PRS.1SG the beard ~ shave.PRS.1SG my beard
'I shave my beard.'
- c. [?]Me controlo las ganas ~ Controlo mis ganas.
REFL.1SG control.PRS.1SG the wishes ~ control.PRS.1SG my wishes
'I control my wishes.'
- d. [?]Me explico las ideas ~ Explico mis ideas.
REFL.1SG explain.PRS.1SG the ideas ~ explain.PRS.1SG my ideas
'I explain my ideas.'
- e. [?]Me establecí la residencia ~ Establecí mi residencia.
REFL.1SG settle.PST1SG the residence ~ settle.PST1SG my residence
'I settled my residence.'

The COSER interviews documented 839 examples of 46 different de-objective verbs, which can be found in Appendix 4. Table 6 shows that these contexts tend to be marked with the RM a great deal of the time (87.3%) and that all de-objective verbs show marking in the corpus, while only a few (9/46) provide unmarked examples.

Table 6. Frequency of the RM with de-objective verbs in the COSER data

	Marked (RM)	Unmarked (No RM)	Total
De-objective examples	736 (87.7%)	103 (12.3%)	839 (100%)
De-objective verbs	46	9	46

As discussed above, the deleted object can usually be identified with a (typically abstract) object from the personal sphere of the subject. Thanks to the abundance of verbs documented in the COSER, we can offer some clues about the semantics of such objects. First, the deleted object can be identified with the subject's mental processes, such as speech (*callarse* 'to shut up', *explicarse* 'to explain oneself', *explicotearse* 'to explain oneself', *dirigirse* 'to go', *expresarse* 'to express oneself', *clamarse* 'to profer', *confesarse* 'to confess'), thoughts (*aclararse* 'to clear one's ideas', *centrarse* 'to focus', *empeñarse* 'to insist', *recrearse* 'to take pleasure', *asimilarse* 'to assimilate'), perceptions (*fijarse* 'to note', *apercibirse* 'to notice') or feelings (*declararse* 'to declare one's love', *desahogarse* 'to give vent to'). Second, it may be linked to the subject's abstract skills, for instance its will or effort (*dedicarse* 'to do something intensively', *aplicarse* 'to do something thoroughly', *arrestarse* 'to dare, to do something difficult', *emplearse* 'to do something thoroughly', *arriesgarse* 'to take a risk'), its organisational skills (*administrarse* 'to organise one's money', *apañarse* 'to manage', *arreglarse* 'to manage', *descuidarse* 'to stop taking care of oneself', *gobernarse* 'to organise', *manejarse* 'to manage, to organise', *prepararse* 'to get ready', *organizarse* 'to organise oneself', *remediarse* 'to manage') or their limits (*aguantarse* 'to put up with something', *resistirse* 'to resist', *concretarse* 'to focus, to limit oneself', *restringirse* 'to limit oneself', *regularse* 'to regulate one's body', *propasarse* 'to go too far', *sobrepasarse* 'to go too far'). Third, it may be identified with notions related with the physical body of the subject, such as appearance (*lucirse* 'to shine') or health (*descomponerse* 'to feel sick', *recuperarse* 'to recover', *reponerse* 'to recover'). Finally, the deleted object may also refer to the (sometimes metaphorical) location of the subject (*establecerse* 'to settle', *instalarse* 'to settle', *situarse* 'to locate oneself, to understand oneself', *ubicarse* 'to locate oneself, to understand oneself', *asomarse* 'to show').

If we take a look at the most frequent verbs (see Figure 4),⁹ we can see that there is a lot of individual variation as regards the reflexive marking of the de-objective diathesis. It seems that verbs related to speech (*llamar*, *confesar*, *explicar*) especially favour the absence of the RM, but note that they are not the only ones (see *asomar* and *aguantar*).

Lastly, as regards the geographical distribution of the RM in de-objective verbs, since the situation seems to be determined by the verbal root to a high degree, it makes sense to look at them separately. Map 6 shows the five most frequent verbs that did not exhibit a systematic or almost systematic pattern, namely, *aguantar*,

9. In Figure 4, only verbs with ten or more occurrences were considered: *fijar* (257), *llamar* (111), *dedicar* (102), *asomar* (48), *confesar* (42), *arreglar* (41), *apañar* (33), *aguantar* (21), *descuidar* (20) and *explicar* (20).

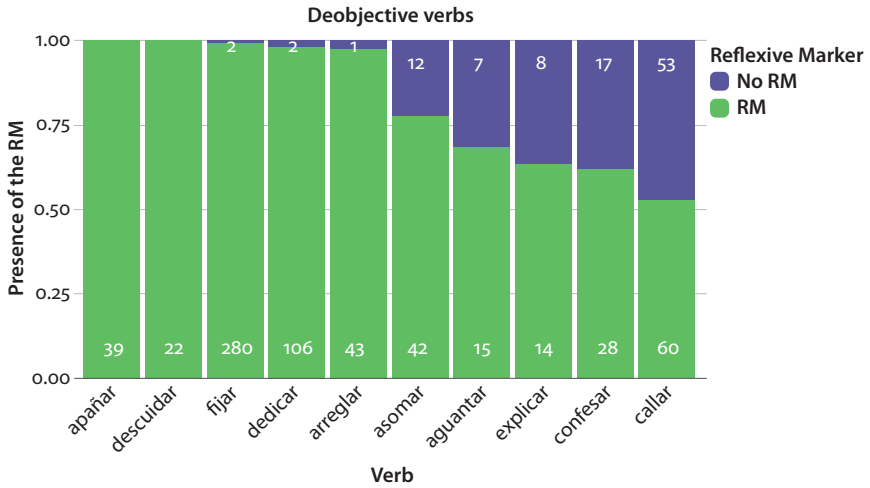
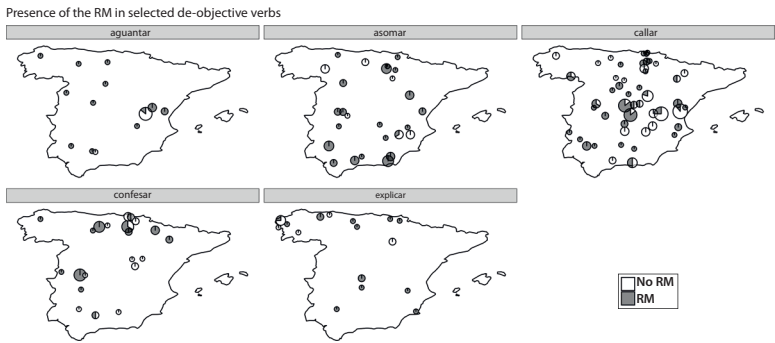


Figure 4. Frequency of the RM with individual de-objective verbs in the COSER data

asomar, *callar*, *confesar* and *explicar*. As can be seen, no clear nor consistent spatial pattern arises. While the same speaker is responsible for most of the cases of non-reflexive *aguantar*, the variation found for other verbs cannot be easily explained on the grounds of their geographical distribution.



Map 6. Frequency of the RM with selected de-objective verbs in the COSER data

The size of the pie charts is a function of the square root of the total number of occurrences of the verb at each point. The range of occurrences at each location is as follows: *aguantar* (1, 7), *asomar* (1, 5), *callar* (1, 9), *confesar* (1, 6), *explicar* (1, 4).

3.2.4 Subject-demoting diathesis: Conversive verbs

In conversive diathesis, the direct object in the basic structure is promoted to subjecthood, while the subject in the basic structure is demoted to the indirect object or to a high rank oblique object (Kulikov 2011):

- (53) a. Juan avergonzó a María.
 Juan embarrass.PST3SG to María
 ‘Juan embarrassed María.’
 b. María se avergonzó de Juan.
 María REFL.3 embarrass.PST3SG of Juan
 ‘María was ashamed of Juan.’

The fact that the RM often marks the so-called conversive diathesis has gone unnoticed in many studies, which have equated the verbs in this category with anticausatives (since in both of them the object is promoted to subjecthood). The main difference between these two diathesis alternations is the fact that in the latter the subject is completely deleted, while in the former it is demoted to an oblique object.

The conversive diathesis also shows some resemblance to the passive, since the participant coded as the object in the default transitive scheme is promoted to subjecthood in both of them (Kulikov 2011). However, while in conversive diathesis the subject is demoted to an object position, typically either the indirect object or a “high rank” oblique object, in the passive diathesis the subject in the active scheme is typically either removed or demoted to a low position (Kulikov 2011: 380). The passive and the conversive diatheses differ more radically in semantics terms, as this author also notes. While the typical passive occurs with verbs that have an agent and a patient in the active structure (see (54)), the conversive is especially common with verbs that take an experiencer and a stimulus, like perception and emotion verbs (see (55)).¹⁰

- (54) a. La policía arrestó al alcalde.
 the police arrest.PST3SG to.the mayor
 ‘The police arrested the mayor.’
 b. El alcalde fue arrestado (por la policía).
 the mayor be.PST3SG arrest.PTCP by the police
 ‘The mayor was arrested (by the police).’

10. Sánchez López (2002) highlights the fact that emotional reaction verbs take an experiencer subject in the reflexive version, as opposed to anticausatives, which take a patient subject.

- (55) a. El arresto del alcalde alegró a los ciudadanos.
 the arrest of.the mayor make.happy.PST3SG to the citizens
 'The arrest of the mayor made the citizens happy.'
- b. Los ciudadanos se alegraron (del arresto del alcalde).
 the citizens REFL.3 be.happy.PST3PL of.the arrest of.the mayor
 'The citizens were happy about the arrest of the mayor.'

The connection of the converse with the anticausative lies in the fact that the oblique object encoding the stimulus argument can remain unspoken in the converse. In this case, the structure is superficially identical to anticausative verbs, in which the object is promoted to the subject position and the agent or cause is removed from the structure. Nevertheless, besides the possibility of encoding the first argument as an oblique object (a possibility which anticausative verbs lack), there is another important difference between them in the varieties under study. The transitive counterpart of a verb with converse diathesis does not necessarily always code the second argument (the experiencer) as a direct object, since it may receive dative morphology. Several authors have noticed this alternation, exemplified in (56), which seems to depend on several linguistic parameters (such as animacy, verb aspect, or the position of the subject) (Fernández-Ordóñez 1999: 1323–5, Rodríguez Ramalle 2005).

- (56) a. A María le asustan las tormentas.
 to María DAT.3SG frighten.PRS.3PL the storms
 'María is scared of storms.'
- b. María iba andando tan tranquila cuando Juan la
 María go.PST.3SG walking so calm when Juan ACC.3SG
 asustó.
 frighten.PST3SG
 'María was calmly walking by when Juan frightened/scared her.'

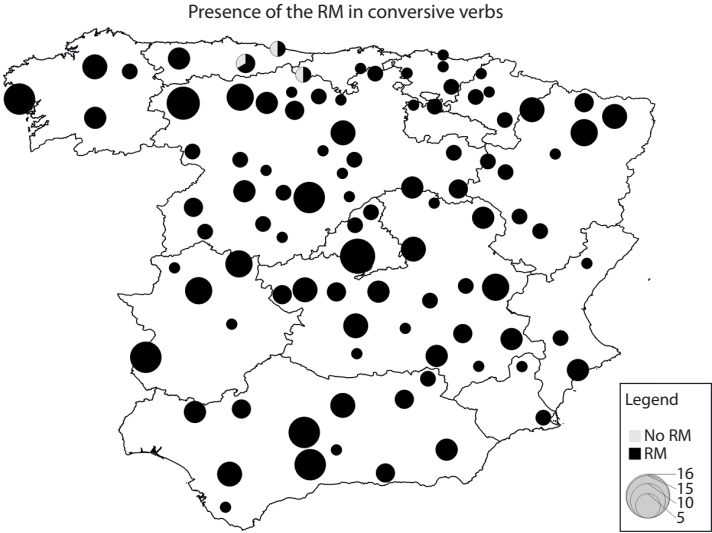
The COSER interviews contain 455 examples from 39 different converse verbs, which can be found in Appendix 4. As pointed out in the literature, most of these verbs are emotion verbs, but not exclusively – ten of them (25.6%) cannot be considered as such.

Turning to the productivity of the RM in converse diathesis, only three (0.7%) of these 455 examples are unmarked in the COSER interviews (see Table 7) and virtually all verbs show marking in the corpus (38/39), while very few also provided unmarked examples (3/39). Two unmarked examples are given in (57). Despite the paucity of unmarked examples, it is worth mentioning that they are all found in the north-western area, in a few villages in Asturias and Cantabria, as can be seen on Map 7. Galicia also shows consistent marking of converse verbs, with no exception in our data.

- (57) a. Pero si [el oso] no va herido pues suele
 but if the bear not go.PRS.3SG injured well use.to.PRS.3SG
espantar de los animales.
 get.scared.INF of the animals
 ‘But if the bear is not injured, it tends to be afraid of other animals.’
 (Ledantes, Cantabria, COSER 1212)
- b. Sí, no, animar sí me animo, qué remedio
 yes no cheer.up.INF yes REFL.1SG cheer.up.PRS.1SG what remedy
 tengo, pero... cuesta, ¿eh?
 have.PRS.1SG but be.hard.PRS.3SG huh
 ‘Yes, sure, cheering up I do cheer up, what else can I do, but it is hard, huh?’
 (Alea – Linares, Ribadesella, Asturias, COSER 0506)¹¹

Table 7. Frequency of the RM with conversive verbs in the COSER data

	Marked (RM)	Unmarked (No RM)	Total
Conversive examples	452 (99.3%)	3 (0.7%)	455 (100%)
Conversive verbs	38	3	39



Map 7. Frequency of the RM with conversive verbs in the COSER data

11. Example (57b) is especially interesting, as it is an infinitive in a focus position, a context that may favour the absence of the RM – movement of the verb to the focus position can allow for the absence of verbal objects (Rodríguez Molina 2014). Note that this observation would imply that the RM is interpreted as a verbal object and is hence less grammaticalised (at least in these varieties), an idea worth exploring further in future works with both more diachronic and diatopic data.

3.2.5 Object-demoting diathesis: Antipassive verbs

The term “antipassive” was coined by Silverstein (1972, apud Polinsky 2013) and underlines the fact that the antipassive mirrors the passive construction. In the antipassive construction the object of a transitive verb is demoted – not the subject, as in the passive construction (see (58)). It is common to group the cases of demotion to an oblique argument position with the cases of complete deletion of the object (Kulikov 2011; Polinsky 2013), since the oblique object in the antipassive may not be explicitly coded, hence producing an absolute reading. However, I have kept them apart in this study, since not all verbs that mark object demotion with the RM take an oblique object. For this, I have followed a lexical criterion – verbs that may take the oblique object (that is, that can undergo the antipassive change of diathesis) are still considered antipassive verbs if the oblique object is absent, and thus are dealt with here. Other cases of object deletion were discussed above (Sections 3.2.2 and 3.2.3).

(58) (Bogard 1999a)

- | | | |
|----|---|---------------|
| a. | Juan aprovechó tu experiencia. | (Transitive) |
| | Juan profit.PST.3SG your experience | |
| | ‘Juan used your experience.’ | |
| b. | Juan se aprovechó (de tu experiencia). | (Antipassive) |
| | Juan REFL.3 profit.PST.3SG from your experience | |
| | ‘Juan profited (from your experience).’ | |
| c. | Tu experiencia fue aprovechada (por Juan). | (Passive) |
| | your experience be.PST.3SG profit.PTCP by Juan | |
| | ‘Your experience was used (by Juan).’ | |

The fact that the term “antipassive” has been used mostly in typological studies and in works on ergative languages explains why it has been applied to Spanish only recently.¹² However, the label has become popular and is used in a number of modern studies (Bogard 1999a; Sánchez López 2002; Portilla 2007). Antipassive constructions are not equally productive in all the languages in which they are found (Polinsky 2013). In fact, productive antipassive constructions are restricted to certain language families (Mayan, Salishan, Nakh-Daghestanian, Austronesian and Australian languages), all of these found either in North America or in Australian and the South Pacific islands. Geniušienė (1987), who calls this diathesis “deaccusative”, indicates that these reflexive verbs “are usually a small peripheral set of verbs”

12. Although it is almost a cliché to point out that antipassives are typically found in ergative languages, this is actually not the case, as Polinsky (2013) notes. She found no correlation between ergativity and the existence of antipassive structures in her sample, that is to say, antipassive constructions are equally common in accusative languages. The status of the antipassive in accusative languages has been extensively discussed by Janic (2016).

(Geniušienė 1987: 256). She also notes that antipassive reflexive verbs in Lithuanian and Latvian belong to a large lexical range and that they show notable variation regarding the case marking taken by the demoted object.¹³

Polinsky also observes that it is not uncommon for antipassive markers to be syncretic with other detransitivisers, especially RMs. This is the case of many verbs in the varieties under examination in the current study. However, it should be noted that not all verbs that display antipassive diathesis in these varieties are marked with the RM, as can be seen from the contrast in (59).¹⁴

- (59) a. ¡Tengo que pensar un plan!
 have.PRS.1SG that think.INF a plan
 ‘I have to think of a plan!’
 b. ¡Tengo que pensar en un plan!
 have.PRS.1SG that think.INF in a plan
 ‘I have to think of a plan!’

This is why a brief note on the examples I collected is in order. I collected exhaustively instances of those verbs that are known to take the antipassive construction and show reflexive marking, in order to see whether the presence of the RM in such constructions was general. However, I was not exhaustive in the collection of other verbs that are typically unmarked in the antipassive construction – although I collected reflexive marked examples of those, if they appeared at all.

As regards the semantic function of the antipassive construction, Polinsky (2013) claims that it cancels out the implication of the prototypical transitive verb that “the event denoted by that verb causes a change of state in the object participant” (Polinsky 2013: 4.3). She notes that the antipassive construction is associated with the low individuation of the object in several languages (hence being required or preferred when the object is plural, indefinite, non-specific, generic or an implicit argument). Along similar lines, Cooreman (1994), in her study of the antipassive construction in ergative languages, concludes that antipassive constructions tend to appear in low transitivity contexts.

13. She lists the three main lexical types (represented by at least ten verbs) of antipassive verbs in these languages: verbs meaning ‘throw’ or ‘splash’; verbs meaning ‘dig (in), burrow (in), splash (in)’; and verbs meaning ‘catch, get hold of’, ‘snatch at’. Only the latter seem to be marked with the RM in the varieties under study, with examples such as *agarrarse a la barra* ‘to grab the bar’ and *cogerse de las agarraderas* ‘to grab the handgrip’.

14. The NGLE (2009: 36.3) lists a large number of verbs in which either a direct object or an oblique object can be coded with no marking whatsoever on the verb: *acertar* (*con*), *hablar* (*en*), *contestar* (*a*), *saber* (*de*), *pagar* (*por*), *recurrir* (*contra*), etc. Rodríguez Ramalle (2005) actually considers them to be in separate groups: unmarked verbs fall under what she terms “the prepositional alternation”, while marked verbs appear to participate in the antipassive alternation. However, she does not specify the reasons why each verb takes one alternation or the other.

However, Bogard (1999a, 2006) emphasises how the agent is presented in the antipassive: “The antipassive expresses an agent with an even higher topical value and its contrast with the patient, which loses its topical value and, as a consequence, also loses its formal manifestation as an argument of the sentence” (Bogard 2006: 766, my translation).

Concerning the specific case of reflexive antipassive verbs in Spanish, Maldonado (1999) claims that the RM in these constructions decreases the degree of control of the subject over the action. That is, in his account, demoting the patient to an oblique object entails that the subject is not interacting directly with the object. This can be seen both in reflexive and non-reflexive antipassives (see (59)). The addition of reflexive marking then adds an additional meaning, this time related to the decrease of control over the action and not the subject.

For Bogard (1999a), however, the presence of the RM is not a matter of control, but a matter of affectedness – a notion related to the semantic nature of the RM, originally reflexive (Bogard 1999a: 324–5). According to him, agreement between the RM and the subject adds a nuance related to the subject’s involvement in the action.

Although different authors do not agree as to the specific semantic content of the antipassive alternation in Spanish, they all link it to a reduction in *semantic* transitivity (that is, in Hopper & Thompson’s (1980) sense, see Section 1.4.2). The COSER data show that verbs with a broad range of meanings enter into this alternation in the varieties under study, making it difficult to propose a unified semantic account.

A total of 28 different verbs, which can be found in Appendix 4, participate in the antipassive construction in the COSER interviews. Besides the relatively large number of cognition and speech verbs (*confundir* ‘to confuse’, *acertar* ‘to be right’, *olvidar* ‘to forget’, *acordar* ‘to remember’, *recordar* ‘to remember’, *decidir* ‘to decide’, *entender* ‘to understand’, *negar* ‘to deny’, *tratar* ‘to deal with’, *referir* ‘to refer’), not many generalisations can be made, although there are some small subgroups with similar meanings. Both *aprovechar* ‘to use’ and *disfrutar* ‘to enjoy’ have an autobenefactive meaning; *cambiar* ‘to change’ and *mudar* ‘to change’ are partial synonyms, and there are some verbs related to grabbing (*agarrar* ‘to grasp’, *coger* ‘to take’, *sujetar* ‘to hold’, and its antonym *desprender* ‘to remove’) and (un)loading (*acarrear* ‘to carry’, *cargar* ‘to load’, *descargar* ‘to unload’, and a more neutral version, *contener* ‘to contain’). Verbs like *atravesar* ‘to cross’, *dejar* ‘to leave’, *empezar* ‘to begin’, *encargar* ‘to order’, *encontrar* ‘to find’, *cuidar* ‘to take care’, however, do not fit into any of these clusters.

The productivity of the RM with these verbs is quite high in the COSER interviews, as shown in Table 8. From 274 examples of the verbs which were exhaustively collected, only 53 (19.3%) are unmarked (see Table 8). When we look at verbs, it is interesting to note that all of them accept marking (11/11), but almost all of them also provide unmarked Examples (9)/(11).

Table 8. Frequency of the RM with antipassive verbs in the COSER data (only exhaustively collected verbs)

	Marked (RM)	Unmarked (RM)	Total
Antipassive examples	221 (80.7%)	53 (19.3%)	274 (100%)
Antipassive verbs	11	9	11

The lack of a clear semantic association between the antipassive and a specific verb semantics, together with the fact that, as stated above, many verbs that also display the antipassive alternation do not take the RM, makes the reflexive antipassive a somewhat special category among the reflexive constructions analysed thus far. Reflexive anticausatives, for instance, also lack a strong semantic association, whereas reflexive marking is very consistent across all the members of the category. On the other hand, reflexive conversive verbs show both constant reflexive marking across the category and a strong semantic association with emotion events. As regards de-objective verbs, the RM is not productive in all object-deletion contexts, but seems to be reduced to a few semantic contexts related to the personal sphere of the subject. Reflexive marking in antipassive constructions, however, is neither consistent across all the members of the alternation nor can it be associated with a specific semantic content.

In this sense, it is worth noting that the productivity of the RM with antipassive verbs seems to be dependent to a great degree on the lexical items even within verbs that show reflexive marking. Figure 5 shows the relative distribution of marked and

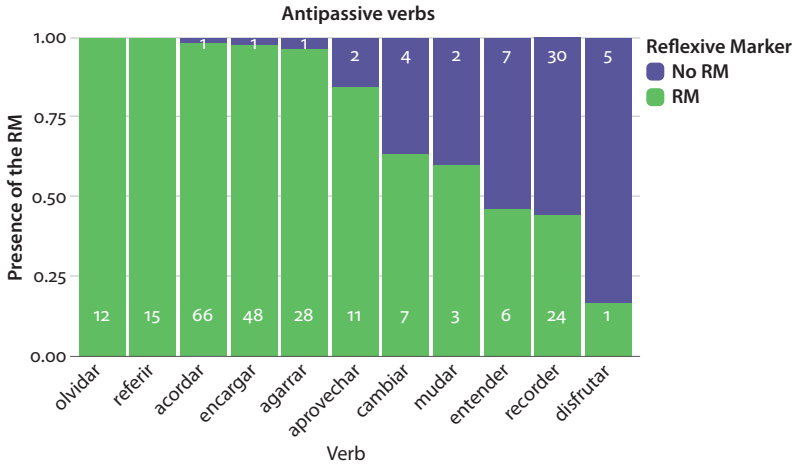
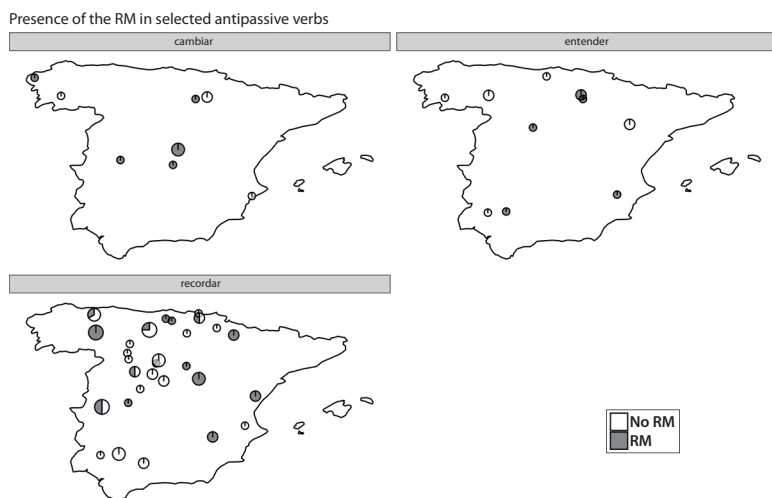


Figure 5. Frequency of the RM with antipassive verbs in the COSER data

Acordar(se) ‘to remember’ here refers only to cases of this verb in Galicia, where it alternates with a transitive unmarked counterpart. In the rest of the territory it does not, so it has been considered a non-reversible reflexive verb. This is discussed in more detail in Section 3.2.7.

unmarked forms in antipassive verbs with five or more occurrences in the corpus. This figure shows that the presence of the RM varies noticeably from verb to verb, with some of them showing systematic marking and others showing a high proportion of unmarked occurrences.

The geographical distribution of the frequency of the RM with the three most frequent and variable antipassive verbs shows no clear geographical pattern either (see Map 8), but it needs to be highlighted that the data are rather scarce. In the case of *recordar* ‘to recall’ our data seem to suggest that *recordarse* is particularly frequent in non-central varieties (north to south), while the opposite could be inferred from the case of *entender* ‘to understand’, which shows very few occurrences.



Map 8. Frequency of the RM with selected antipassive verbs in the COSER data

The size of the pie charts is a function of the square root of the total number of occurrences of the verb at each point. The range of occurrences is as follows: *cambiar* (1, 3), *entender* (1, 2), *recordar* (1, 4).

3.2.6 No change of diatheses: Reflexive intransitive verbs

As Kemmer observed, purely syntactic approaches to reflexive verbs or the middle voice, which attribute a valency-changing function to the reflexive/middle marker, encounter a number of problems. One of these is the general tendency found in the middle systems of the world for the reflexive/middle marker to attach to verbs with no effect on their original valency, this in a non-negligible number of cases:

M[iddle] M[arked] classes whose verbs do not generally have corresponding unmarked transitive verbs, and which have received little attention in the literature mentioned above, are the emotive speech action, translational motion, and cognition middle classes. Particularly for the latter two classes, where unmarked forms exist, they are often intransitive, in keeping with the natural one-participant semantics of the root verb meanings (e.g. Spanish *ir-se* ‘go away’, *caer-se* ‘fall’, formed on the roots *ir* ‘go’ and *caer* ‘fall’ respectively). (Kemmer 1994: 185–6)

In this section I discuss intransitive verbs that might take the RM, which henceforth I will call reflexive intransitive verbs. As on some previous occasions, when addressing the study of the presence of the RM in verbs with no change of valency we are faced with the problem of exhaustivity. Since many verbs do not show the RM at all, it makes little sense to collect every instance of every intransitive verb. For this reason, exhaustive data were only collected for some verbs which are known to take the RM. Some verbs were collected exhaustively in all the COSER interviews, while others were only collected in subcorpus E, due to their high token frequency. The latter category includes the verbs: *estar* ‘to be’,¹⁵ *salir* ‘to go out’, *volver* ‘to go back’, *ir* ‘to go’, *venir* ‘to come’, *morir* ‘to die’, *vivir* ‘to live’, *nacer* ‘to be born’, *dormir* ‘to sleep’, *escapar* ‘to escape’, *valer* ‘to be worthy’, *andar* ‘to go by foot’, *llegar* ‘to arrive’, *quedar* ‘to stay’ and *entrar* ‘to go in’, while the former includes the verbs: *pasar* ‘to pass by’, *caer* ‘to fall’, *esperar* ‘to wait’, *marchar* ‘to leave’, *reír* ‘to laugh’, *cagar* ‘to shit’, *montar* ‘to climb’, *arder* ‘to burn’, *crecer* ‘to grow’, *mear* ‘to piss’. I also collected exhaustively the instances of *quedar* ‘to stay’, *entrar* ‘to go in’ or *llegar* ‘to arrive’ in those localities of subcorpus NE that showed either causative instances of *quedar* ‘to stay’ or *entrar* ‘to go in’ or marked examples of *entrar* ‘to go in’ and *llegar* ‘to arrive’,¹⁶ since they are frequent verbs with an infrequent marked counterpart. Only marked examples were collected from all other intransitive verbs. A total of 62 intransitive verbs that took the RM at least once were documented. These verbs can be found in Appendix 4.

If we take into consideration only the 25 verbs that have been exhaustively collected, we can check the productivity of the RM as a marker for this class.

15. *Estar* was not collected in Galicia, where no instances of the marked version (*estarse*) were found.

16. Western Spanish dialects show a productive use of *quedar*, *entrar* and *caer* as causative intransitives (a description of the current geographical distribution can be found in Lara Bermejo 2020). However, the existence of such a causative use does not seem to affect the frequency of the RM with these verbs in this area (that is, this frequency is similar to the frequency of the RM in speakers that do not show examples of the causative version) (De Benito Moreno 2015: 7.3.1.6.4, 7.3.1.7).

A total of 14,343 examples were found, out of which only 2,620 (18.3%) showed the RM (see Table 9). This eloquently shows that the norm is for intransitive verbs to be unmarked. Obviously, marked and unmarked examples were attested for all verbs (25/25).

Table 9. Frequency of the RM with exhaustively collected intransitive verbs in the COSER data

	Marked (RM)	Unmarked (No RM)	Total
Intransitive examples	2,620 (18.3%)	11,723 (81.7%)	14,343 (100%)
Intransitive verbs	25	25	25

If we take a closer look at each of these 25 verbs separately (presented in two separate figures for the sake of a clear visualisation, see Figures 6 and 7), we see that there is a lot of internal variation, suggesting that this is a very heterogeneous class. Although some verbs with similar meanings (like the pairs *cagar* ‘to shit’/ *mear* ‘to piss’) show similar proportions, this is not the case for more abstract groupings – movement verbs like *ir* ‘to go’, *venir* ‘to come’, *salir* ‘to go out’, *marchar* ‘to leave’ and *volver* ‘to go back’ show notable differences in the proportions of their marking.

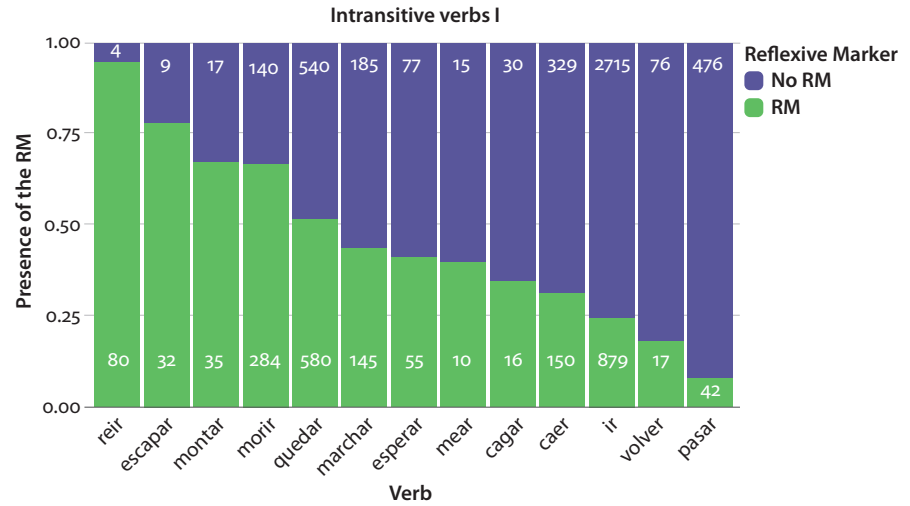


Figure 6. Frequency of the RM by verb in exhaustively collected intransitive verbs in the COSER data (most frequently reflexive verbs)

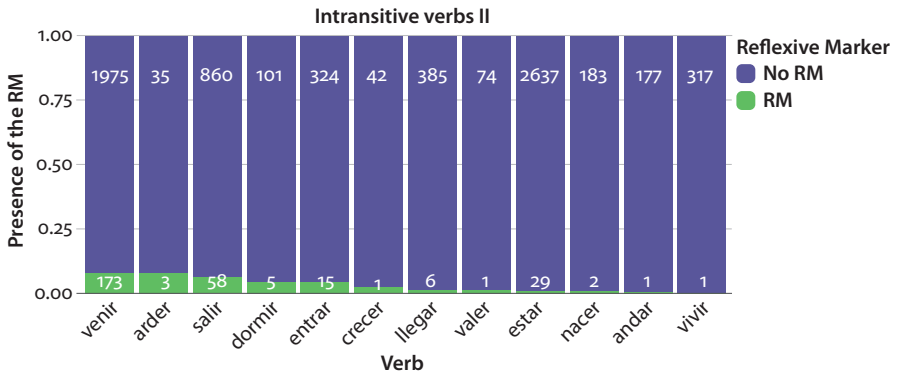
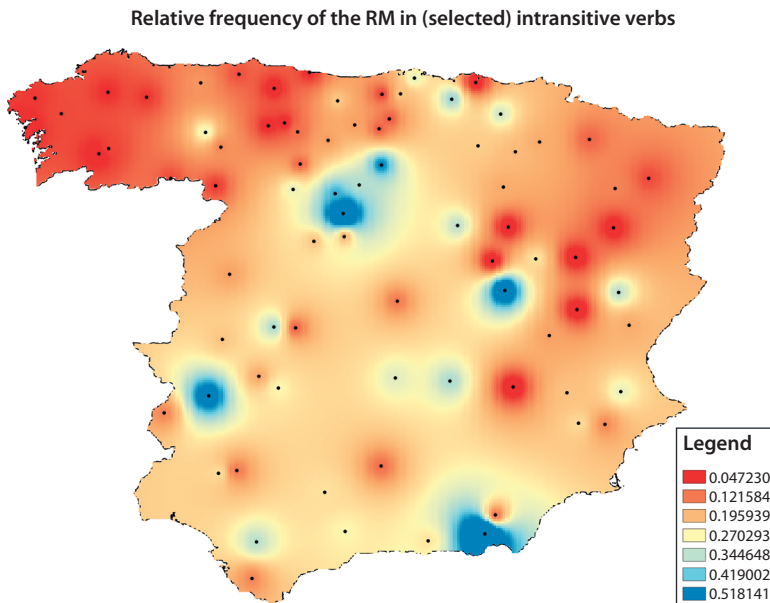


Figure 7. Frequency of the RM by verb in exhaustively collected intransitive verbs in the COSER data (least frequently reflexive verbs)

Let us now turn to the geographical distribution of reflexive intransitive verbs, for which I have only considered these 25 verbs. Because they show very different frequencies, I have calculated first the relative frequency of each verb in each interview and then calculated the mean of all verbs documented in a given interview. In order to avoid extreme, non-representative figures, I have only considered verbs with ten or more occurrences per interview. Map 9 shows a rather clear pattern, according to which north-western varieties are the most resistant to marking verbs with the RM.



Map 9. Relative frequency of reflexive intransitive verbs in the COSER data

3.2.7 No change of diatheses: Non-reversible verbs

A second problem that purely syntactic approaches to reflexive verbs or the middle voice have to face has to do with verbs that lack a derivational base, either because they have no unmarked counterpart or because the meaning of the unmarked counterpart is too distant and hence cannot be interpreted as the derivational base (Kemmer 1993). Nevertheless, these are also common in the middle systems of the world – Inglese (forthcoming) even considers them as a definitory element of middle voice systems. They have received several names: “non-reversible” reflexive verbs (Geniušienė 1987), deponents (Kemmer 1993), non-oppositional (Inglese 2021, forthcoming) and *reflexiva tanta* or *media tanta*. In the literature on reflexive constructions in Spanish they have sometimes been called inherently reflexive, which is somewhat confusing, since the term does not refer unequivocally to formal properties and can be interpreted as a semantic label. In what follows, I will use Geniušienė’s terminology, which most transparently refers to the lack of an unmarked counterpart, the crucial parameter investigated here.

It has been noted that non-reversible verbs can be explained in diachronic terms, by the loss of the unmarked counterpart (Cartagena 1972; Martín Zorraquino 1979; Cidrás Escáneo 1991; Portilla 2007). The historical study of a number of these verbs in Spanish shows that they had unmarked counterparts in previous stages of the language, and that the relationships between the marked and unmarked forms in those older stages fit into the alternations found in the modern language (Cartagena 1972). The same findings have been reported for Galician – non-reversible verbs come from old alternations where the unmarked verb has disappeared (Cidrás Escáneo 1991). Globally, this suggests that they are not to be seen as a special category within reflexive constructions, but that each non-reversible verb belongs to one of the categories mentioned above. Actually, Spanish non-reversible verbs have sometimes been classified as antipassive verbs on these grounds, since they very often take oblique complements that convey the patient (Portilla 2007; Bogard 1999a).

Although it has been claimed that Spanish non-reversible verbs always take an oblique complement (see Hernández Terrés 1984), the fact is that some do not (Otero 1999). Actually, some non-reversible verbs were in fact “born” non-reversible, due to their inherent reflexive meaning (*suicidarse* ‘to commit suicide’) (Cartagena 1972, see also de Benito Moreno accepted).¹⁷

There have been several proposals towards providing a coherent semantic account of the class of non-reversible verbs. Hernández Terrés (1984) argues that

17. Cartagena notes that *suicidar* actually shows a tendency to be inserted in the alternations typical of the RM, as shown by ironical examples like *Yo la suicidé*. Cidrás Escáneo (1991) makes similar claims about (*d*)*espadirse* and *omillarse* in Galician, which “were born” reflexive and later developed an unmarked counterpart.

they all have animate subjects and convey processes or activities that only affect the agent, by changing its state. He claims that the general meaning described by Martín Zorraquino (1979) for many reflexive constructions, namely, “the assimilation to the subject’s sphere”, is stronger in these verbs.

Maldonado (1999) notes that all non-reversible verbs express an emotion or an emotional response and claims that they can be explained in the same terms as all the other uses of the RM: according to him, the RM produces energetic as opposed to absolute construals, that is, it foregrounds energy in the conceptualisation of the event. In this case, the RM is attached to events with a middle experiencer, which shows “an intermediate level of responsibility as regards the emotion that it expresses” (Maldonado 1999: 114, my translation). Middle experiencers contrast with passive experiencers, which show no control at all over the emotion. This would therefore explain the examples in (60), since verbs that do not admit the RM, like *llorar* ‘to cry’ and *sollozar* ‘to sob’, also reject adverbs that require high control, while *quejarse* ‘to complain’ admits them:

(60) (Maldonado 1999)

- a. Ese hombre es de los que se quejan poco a poco.
that man is of the that REFL.3 complain.PRS.3PL little to little
‘That is one of those men that complain little by little.’
(Middle experiencer)
- b. *Ese hombre es de los que lloran / sollozan poco a poco.
that man is of the that cry.PRS.3PL / sob.PRS.3PL little to little
‘That is one of those men that cry/sob little by little.’ (Passive experiencer)

According to this author, the middle experiencer always has a complex representation – it takes both the original thematic role required by the verb and an active meaning component that triggers some kind of expressive nuance. For instance, *vanagloriarse* ‘to boast’ can be interpreted as an intensification of *halagar* ‘to flatter’ or *glorificar* ‘to glorify’. On this view, the existence of non-reversible verbs is a natural result of the fact that some verbs have a complex base and hence require a complex experiencer (i.e. a middle experiencer).

It has been noted that there is a large number of non-reversible verbs in Spanish, which suggests that they are not some sort of rarity (Cartagena 1972). This probably has to do with the fact that, for some mental verbs, the transitive form with no RM seems to be more marked (in terms of frequency) than the middle form with the RM (Maldonado 1999). This is illustrated in (61), where the transitive form with an agent and a patient (in (61c)) is dispreferred. The forms with the RM being the unmarked ones, and thus more frequent, would explain that their unmarked counterparts are lost in the evolution of the language, while they themselves survive.

- (61) (Maldonado 1999)¹⁸
- a. Este tipo de respuestas me enojan.
this type of answers ACC.1SG upset.PRS.3PL
‘This kind of answers upsets me.’
 - b. Me enojo con este tipo de respuestas.
REFL.1SG upset.PRS.1SG with this type of answers
‘I get upset with this kind of answers.’
 - c. *Los niños enojaron al maestro.
the children upset.PST.3PL to.the teacher
‘The children irritated the teacher.’
 - d. Los niños hicieron enojar al maestro.
the children make.PST.3PL upset.INF to.the teacher
‘The children made the teacher upset.’
 - e. El maestro se enojó con los niños.
the teacher REFL.3 upset.PST.3SG with the children
‘The teacher got upset with the children.’

The COSER interviews yielded a total of 770 examples of 57 different non-reversible reflexive verbs, which can be found in Appendix 4.¹⁹ This clearly confirms Cartageña’s claim as to the quantitative importance of class.²⁰ The productivity of the RM in this class is, by definition, maximal (although there are a few interesting exceptions discussed below – see also Table 10). This is why its semantic composition is of interest to us in the present study.

Table 10. Frequency of the RM with non-reversible verbs in the COSER data

	Marked (RM)	Unmarked (RM)	Total
Non-reversible examples	764 (99.2%)	6 (0.8%)	779 (100%)
Non-reversible verbs	3	57	57

This large list of verbs allows us to update previous descriptions. First, as is clear from Table 11, not all non-reversible verbs are emotion verbs, although they represent a large part of the class (24.1%) – even larger (43.1%) if we group all mental events together by adding cognition verbs. There are also a few verbs related to the body domain which are either body processes, non-translational motion events, or translational motion events.

18. The star in (61c) is also added by Maldonado – such an example sounds grammatical to me.

19. Note that I consider verbs that completely lack an unmarked form and also verbs whose meaning is very distant from their marked counterpart.

20. Inglese (forthcoming) notes that there are languages, such as Old Hittite, where non-reversible verbs have the highest type frequency.

Table 11. Non-reversible verbs by semantic class in the COSER data*

Body processes 3 (5.2%)	<i>herniarse</i> 'to get a hernia', <i>constiparse</i> 'to catch a cold', <i>entumirse</i> 'to get stiff'
Cognition 11 (19%)	<i>acordarse</i> 'to remember', <i>antojarse</i> 'to feel like doing something, to have cravings', <i>conformarse</i> 'to be satisfied', <i>encapricharse</i> 'to get something into one's head', <i>enrollarse</i> 'to jabber', <i>esforzarse</i> 'to try hard', <i>fiarse</i> 'to trust', <i>figurarse</i> 'to imagine', <i>hacerse cargo</i> 'to understand', <i>ocurrirse</i> 'to have an idea', <i>resignarse</i> 'to resign oneself to'
Emotion 14 (24.1%)	<i>achicarse</i> 'to be intimidated', <i>apegarse</i> 'to become attached', <i>arrepentirse</i> 'to regret', <i>atreverse</i> 'to dare', <i>burlarse</i> 'to make fun of', <i>cachondearse</i> 'to make fun of', <i>cebarse</i> 'to vent one's anger', <i>desbocarse</i> 'to fall out of control', <i>desgañitarse</i> 'to shout oneself hoarse', <i>engarrancharse</i> 'to get angry', <i>guasearse</i> 'to make fun of', <i>pitorrearse</i> 'to make fun of', <i>quejarse</i> 'to complain', <i>revestirse</i> 'to get angry'
No trans. motion 3 (5.2%)	<i>acurrucarse</i> 'to curl up', <i>espanzarse</i> 'to fell down', <i>espatarrarse</i> 'to fell down'
Spontaneous 7 (12.1%)	<i>agusanarse</i> 'to become infested with worms, maggots', <i>alunarse</i> 'to become infested with worms, maggots', <i>mullarse</i> 'to clot', <i>musirse</i> 'to become moldy', <i>regustarse</i> 'to get more flavour', <i>revenirse</i> 'to go bad, rancid (food)', <i>triar</i> 'to curdle'
Trans. motion 4 (6.9%)	<i>esvolarse</i> 'to fly off', <i>fugarse</i> 'to flee', <i>pirarse</i> 'to leave', <i>recorgajarse</i> 'to swing'
Other 16 (27.6%)	<i>agravar</i> 'to entrust', <i>comportarse</i> 'to behave', <i>desenvolverse</i> 'to perform well', <i>deshacerse</i> 'to get rid of', <i>desliarse</i> 'to get out of control', <i>escabullirse</i> 'to escape, to slip through', <i>esmerarse</i> 'to put a lot of effort', <i>eternecerse</i> 'to go on forever', <i>hacerse (con algo)</i> 'to acquire', <i>lucrarse</i> 'to make a profit', <i>portarse</i> 'to behave', <i>preciarse</i> 'to pride oneself on', <i>prestarse</i> 'to be suitable for', <i>remontarse</i> 'to go back to (a point in the past)', <i>suscribirse</i> 'to take out a subscription', <i>terciarse</i> 'to arise the opportunity'

* This semantic classification follows Kemmer's (1993) categories, anticipating Section 3.3.

A non-negligible number of spontaneous events are also documented. Interestingly, most of these have a similar meaning, related to 'to rot, to go bad' (such as *alunarse* 'to become infested with worms, maggots', *agusanarse* 'to become infested with worms, maggots', *revenirse* 'to go bad, rancid (food)', *triar* 'to curdle', *musirse* 'to become moldy'), but also other meanings related to changes of state that food may undergo (such as *regustarse* 'to get more flavour', *mullarse* 'to clot'). These examples show that a further two generalisations that have been made for this class, i.e. that they always take animate subjects and oblique objects (Hernández Terrés 1984) do not hold for the entire class. There is also a notable number of verbs that cannot be ascribed to any of Kemmer's (1993) semantic types. In all, these observations suggest that this class is not as uniform as some authors have claimed.

Finally, it is worth noting that some unmarked forms of these verbs were documented, which of course is unexpected. On the one hand, *acordar* ‘to remember’ is actually “reversible” in Galician (and was considered accordingly in interviews in Galicia). This is an excellent example of how the use of different varieties and languages can be interpreted in diachronic terms, since these examples of Galician show the equivalent of an older stage of Spanish.²¹ Our data show that unmarked instances of *acordar* appear in the transitive construction, whereas the RM appears in the antipassive construction, both with and without the oblique object (see (62)). Therefore, the examples of this verb in Galicia were counted as antipassive.

- (62) a. Non había máis que un coche, mire que tempos eu
no have.PST.3SG more than a car look.IMP.3SG what times I
acordo.
remember.PRS.1SG
‘There was just one car, look what old times I remember.’
(Orellán, La Coruña, COSER 2401)
- b. Non s’acordaban de ningunha cousa.
no REFL.3-remember.PST.3PL of none thing
‘They didn’t remember anything.’ (Orellán, La Coruña, COSER 2401)

Three more examples of supposedly non-reversible verbs without the RM were found. There are two instances of *revenir* (in the same speaker from Huesca, see (63)). According to the DRAE, when *revenir* means ‘to go bad, rancid (food)’, it is compulsorily reflexive. The only definition of *revenir* which the DRAE does not classify as non-reversible is ‘to go back to one’s normal state’, which is labelled as intransitive.

- (63) [la masa] Y ya cuando había revenido que decían,
and already when have.PST.3SG get.soft.PTCP that say.PST.3PL
entonces se hacía los panes
then REFL.3 make.PST.3SG the breads
‘And then, when [the dough] had become soft, as they said, then the breads were made.’
(Oliván, Huesca, COSER 2222)

There was also an unmarked example of *quejarse* ‘to complain’, one of the most cited non-reversible verbs. The example is extremely interesting, since it is clear that the absence of the pronoun is due to the construction in which the verb appears (see (64)). *Quejar* ‘to complain’ appears in an infinitive clause governed by a perception verb (*oír* ‘to hear’). In this construction the subject of the infinitive is coded in the

21. The transitive use of *acordar* was already less frequent than *acordarse* in the first Castilian texts (Corominas & Pascual 1981).

accusative, looking superficially identical to the object of the main verb (Hernanz 1999: 2236–47). It seems that the fact that the infinitive subject is identical to the expected RM (*me* ‘me’) prevents it from appearing, in order to avoid the close repetition of two identical pronouns.

- (64) *Estuvo toa la mañana el médico aquí oyéndome quejar.*
 be.PST.3SG all the morning the doctor here hearing-ACC.1SG complain.INF
 ‘The doctor stayed here the whole morning, listening how I complained.’
 (Rábano, Valladolid, COSER 4417)

3.3 The middle as a basic voice: Semantics of the middle domain

I now turn to the analysis of the distribution of the RM in the varieties under study from a semantic perspective. As noted above (Section 1.4.4), the semantic categories to be analysed in this chapter follow Kemmer’s (1993) typological account, where she identifies the main semantic categories that middle systems tend to mark across the world. Her account of the middle voice relies on the relative distinguishability of events, a concept involving “the degree to which the facets in a particular situation, i.e. the participants and conceivable component subevents in the situations, are distinguished” (Kemmer 1993: 208). The distinguishability of events is related to the distinguishability of its participants; that is, when the participants of an event refer to the same entity, they can be presented as showing different degrees of individuation. On this view, subject affectedness, traditionally considered the core semantic component of the middle voice, is just one aspect of the relative distinguishability of participants.

Kemmer claims that the middle domain lies in the intermediate categorial space between pure one-participant events and reflexive constructions, which convey two-participant events but differ from prototypical two-participant events in that they are coreferential:

[T]he reflexive marker in general has the function of designating events in which the Initiator and Endpoint participants are to some extent distinct. [...] In reflexive events, the Initiator acts on itself just as it would on another entity; the reflexive marker is there simply to signal the unusual fact that the different participant roles happened to be filled by the same entity. The middle marker, on the other hand, has the basic function of indicating that the two semantic roles of Initiator and Endpoint refer to a single holistic entity. (Kemmer 1993: 66)

In what follows I will explore how adequately Kemmer’s approach fits the behaviour of the RM in the varieties under study here. In order to do so, I will map the situation types described by Kemmer (see also Section 1.4.4) into lexical verbal categories and analyse the frequency of the RM in each of them, this to test two

specific predictions. First, if the middle domain relies heavily on semantic notions (i.e. situation types), we would expect higher middle marking (reflexive marking in the varieties under study) within these situation types, especially in those that are semantically closer to the direct reflexive (the source of the middle marker in these varieties). Second, we would also expect the semantic distance between these types to be mapped by the relative frequency of marking in these categories – the shorter the distance from the reflexive, the higher the frequency of the RM and vice versa.

Section 3.3.1 addresses the five types of the body action middle, Section 3.3.2 deals with the cognition middle, and in Section 3.3.3 I analyse spontaneous events. In Section 3.3.4 I address briefly (some) naturally reciprocal events, and lastly, Section 3.3.5 focusses on verbs which cannot be ascribed to any of these semantic classes.

Before the analysis of each category, a few general observations about the data used in this chapter are in order. First, only data from subcorpus E of the COSER were used here. The reason for this is that unmarked intransitive and transitive verbs were not collected exhaustively in the interviews excluded from subcorpus E, so the inclusion of only marked forms of those verbs would distort the quantitative results.

Finally, it should be noted that the classification of semantic categories is not always straightforward. Polysemy is a widespread property of natural languages and borders between two different meanings of a given verb are often fuzzy, making semantic classifications difficult. I have used several criteria to reduce the risk of inconsistencies in the classification. First, when one verb has two or more different meanings that can be ascribed to different situation types, I have classified each example accordingly. Second, when two different meanings of a verb are clearly related in metaphorical terms (typically one being more abstract than the other), I have classified the metaphorical uses in the category of the non-metaphorical uses. That is, I have assumed that metaphorical extensions of a given verb inherit the middle-marking of the source construction.

3.3.1 Body action middles

Under this term, Kemmer discusses the behaviour of some verbs that refer to actions normally “carried out on or through one’s own body” (Kemmer 1993: 53). Some are usually treated as purely reflexive in the literature, but she presents convincing evidence that they constitute a different class from a typological perspective. On the one hand, whereas these verbs are typically marked with the RM in one-form languages (cf. Spanish *lavarse*, *sentarse*, *estirarse*, *moverse*), they usually appear as unmarked intransitive verbs (cf. English *wash*, *sit down*, *stretch*, *move*) or in passive

or change-of-state constructions (cf. English *get washed*, *get dressed*) in languages without a middle marker. In two-form languages, they tend to be marked with the middle marker (the light form), rather than with the reflexive (the heavy form).

According to Kemmer, these actions are similar to reflexives in that they are carried out on one's body, but their heterogeneous marking patterns in the languages of the world is due to the fact that "inherent in their meaning is the lack of expectation that the two semantic roles they make reference to will refer to distinct entities" (Kemmer 1993: 58). On the other hand, purely reflexive verbs are formally marked because the expectation is precisely the opposite: they refer to actions normally carried out by one entity on a different entity. This contrast explains why body action middles can be unmarked (and/or conveyed by intransitive verbs) in many languages. Actually, in some languages it is the transitive counterpart of these verbs that is derived from the intransitive (Kemmer 1993: 58–59). In the varieties under study here, we find both intransitive verbs formally derived from a transitive counterpart, often by means of the RM, and purely intransitive verbs.

Kemmer (1993) distinguishes five situation types within the body-motion middle: grooming or body care actions, body processes, changes in body posture, non-translational motion and translational motion. She proposes that these can be ordered according to the degree of the participation of the endpoint – in grooming and body care actions the endpoint is not involved so centrally, while its degree of participation is much higher in translational motion events. The degree of the participation of the endpoint is related to the expectation of coreferentiality between the endpoint and the initiator – the higher the degree of participation, the lower the expectation of coreferentiality:

Moreover, the greater the participation or involvement of the Endpoint-self, the more likely the event is to be viewed as one that *cannot* be directed towards others. [...] Thus the cline describes a gradual progression from actions for which there is merely the *expectation* that the Endpoint will be the same entity as the Initiator, to, at the other extreme, the *requirement* that the Endpoint must be the same entity.

(Kemmer 1993: 60–61, emphasis in original)

In the following sections, I will analyse the frequency of the RM in these five situation types in the varieties under study.

3.3.1.1 *Grooming or body care actions*

Most grooming verbs in the varieties under study can enter in semantically and syntactically reflexive constructions, where the RM is a referential pronoun, as shown by the admission of the emphatic phrase *a sí mismo* (see Section 1.2). Grooming verbs in the direct reflexive construction show high rates of reflexive marking (93% in COSER), very similar to other verbs in the direct reflexive construction (100% in

COSER) (De Benito Moreno 2015).²² The lack of RM in some cases is attributable to the special behaviour of *vestir(se)* ‘to dress’. This verb admits the absence of the RM in imperfective contexts, as noted by Fernández Ramírez (1986: 395). That is to say, when it expresses an imperfective event (equivalent to *to dress* in English) it admits either the absence or the presence of the RM, as shown in (65), while when it expresses a perfective action (equivalent to *to get dressed* in English) it clearly prefers the RM, as illustrated in (66) (De Benito Moreno 2015).

- (65) a. Yo he vestío siempre de pana.
I have.PRS.1SG dress.PTCP always of corduroy
‘I have always dressed in corduroy.’ (Povedilla, Albacete, COSER 0222)
- b. Había quien se vestía de blanco, había quien
have.PST.3SG who REFL.3 dress.PST.3SG of white have.PST.3SG who
se vestía de negro, con mantilla o sombrero.
REFL.3 dress.PST.3SG of black with mantilla or hat
‘There were people who dressed in white, there were people who dressed in white, with a mantilla or a hat.’ (Narros del Puerto, Ávila, COSER 0614)
- (66) Pues yo creo que me vestí yo sola.
well I believe.PST1SG that REFL.1SG dress.PST.1SG I alone
‘Well, I believe that I got dressed on my own.’
(Navalmoral de la Mata, Cáceres, COSER 1015)

Since the direct reflexive construction was excluded from the analysis of the middle as a derived voice (Section 3.2), these examples must also be excluded from the analysis of the middle as a basic voice. However, grooming actions are not only expressed by means of the direct reflexive construction in the varieties under study. In subcorpus E of the COSER there are eight other grooming verbs, in either anti-causative or intransitive constructions, all of which can be found in Appendix 4 (see one example in (67)). There are 87 such examples, only 13 of which (14.9%) showed the RM, as seen in Table 12. In other words, their behaviour differs diametrically from direct reflexive grooming verbs. Note, however, that only three verbs provide unmarked examples – the patterns observed here have to do with the high token frequency of the intransitive verb *ir* ‘to go’, which is typically unmarked.

- (67) A novia non iba de blanco.
The bride no go.PST.3SG if white
‘The bride did not use to dress in white.’ (Orellán, A Coruña, COSER 2401)

22. The main difference between the two categories is related to the frequency of reflexive emphasisers like *a sí mismo*, since no examples were found with grooming verbs. This corroborates Kemmer’s claim that self-directness is the expected case for grooming verbs (De Benito Moreno 2015).

Table 12. Frequency of the RM with grooming verbs in subcorpus E

	Marked (RM)	Unmarked (No RM)	Total
Grooming examples	13 (14.9%)	74 (85.1%)	87 (100%)
Grooming verbs	8	3	8

3.3.1.2 *Body processes*

Kemmer also mentions that verbs designating natural body processes (‘to scratch’, ‘to sneeze’) are sometimes middle-marked in the languages of the world. Although she does not develop this idea much, she includes them within the body action middle class. It seems safe to assume that they are semantically closer to grooming events and events of bodily care than to motion ones. She notes that they “require self-affectedness, since one cannot perform these actions (or anything resembling them) on another” (Kemmer 1993: 61), which suggests that their distance from the direct reflexive is greater than for grooming verbs. That is, we would expect them to show less marking than the latter.

Only 36 instances from eight different verbs (see Appendix 4) designating body processes were found in subcorpus E of COSER, of which 18 (50%) showed the RM, as seen in Table 13). Only three of these eight verbs attest unmarked cases, while all of them provide marked examples. Anticausative, de-objective, intransitive and non-reversible verbs are attested in this category.

- (68) Claro, si eres changa y te meas en la cama.
 Clearly if be.PRS.2SG dirty and REFL.2SG piss.PRS.2SG in the bed
 ‘Of course, if you are dirty and piss in bed.’ (Almajano, Soria, COSER 3901)

Table 13. Frequency of the RM with body process verbs in subcorpus E

	Marked (RM)	Unmarked (No RM)	Total
Body process examples	18 (50%)	18 (50%)	36 (100%)
Body process verbs	8	3	8

3.3.1.3 *Change in body posture events*

Verbs denoting changes in body posture are similar to reflexives in that “they denote actions in which a volitional entity acts on its own body” (Kemmer 1993: 55). They differ from them, however, in that the endpoint of the action is not the body itself but its position. Although these verbs are not typically singled out in the literature on reflexives or anticausatives in Spanish, it is noteworthy that at least some anticausatives that convey changes in body posture seem to be more permissive with the absence of the RM in Standard Spanish, but only in the affirmative imperative (see (69a)). In this sense, they contrast both with reflexive and anticausative grooming verbs (see (69b), (c)) and with anticausative translational motion verbs (see (69d)).

- (69) a. ¡Quinto, levanta, tira de la manta!
conscript stand.up.IMP.2SG pull.IMP.2SG of the blanket
‘Conscript, wake up, pull your blanket away!’ (Spanish reveille song)
- b. ??¡Quinto, peina y a formar!
conscript brush.IMP.2SG and to march.INF
‘Conscript, brush your hair and stand!’
- c. ??¡Quinto, calienta al fuego, que estás empapado!
conscript warm.IMP.2SG to.the fire that be.PRS.2SG soaking
‘Conscript, warm up at the fire, you are soaking!’
- d. ??¡Quinto, mueve, que estás en medio!
conscript move.IMP.2SG that be.PRS.2SG in middle
‘Conscript, move, you are in the way!’

Subcorpus E yielded 382 examples from 16 different change of body posture verbs (see Appendix 4). All of these were anticausative verbs. The proportions of reflexive marking in these verbs strongly contrasts with the situation of the grooming and body process verbs analysed in the above section, since 363 examples (95%) showed the RM, as seen in Table 14. Of the 16 attested verbs, six showed unmarked examples, while 15 show marked examples.

Table 14. Frequency of the RM with verbs of changes in body posture in subcorpus E

	Marked (RM)	Unmarked (No RM)	Total
Change in body posture examples	363 (95%)	19 (5%)	382 (100%)
Change in body posture verbs	15	6	16

3.3.1.4 Non-translational motion events

This label refers to “actions of motor manipulation of the body or a body part, without any particular change of the body” (Kemmer 1993: 56), including verbs meaning ‘to turn around’, ‘to stretch one’s body’ and ‘to bow’. According to Kemmer, these verbs are similar to the reflexive in that the endpoint is the subject’s body. However, these actions are not exactly performed *on* someone’s body, but simply *by* someone’s body; a distinction that relates directly to the relative distinguishability of participants.

A total of 98 examples from 17 different non-translational verbs (see Appendix 4) were recorded in subcorpus E of the COSER interviews. Anticausative, de-objective, conversive, intransitive and non-reversible verbs are all attested in this category. These verbs were marked around half the time (50 marked examples, i.e. 51%), as shown in Table 15. From 17 verbs, 14 showed marked examples, while only six provided unmarked examples.

- (70) Se ha acurrucao ahí.
REFL.3 have.PRS.3SG curl.up.PTCP there
‘S/he curled up there.’ (Malagón, Ciudad Real, COSER 1414)

Table 15. Frequency of the RM with verbs of non-translational motion in subcorpus E

	Marked (RM)	Unmarked (No RM)	Total
Non-translational motion examples	50 (51%)	48 (49%)	98 (100%)
Non-translational motion verbs	14	6	17

3.3.1.5 *Translational motion events*

“Actions involving motion of an animate entity under its own power through space” (Kemmer 1993: 69) also display middle marking cross-linguistically. In Kemmer’s (1993) approach, this category only includes originally intransitive verbs. She considers reflexive translational motion verbs that have a causative unmarked counterpart to be true reflexives, since they usually take a second participant (distinct from the first):

When the reflexive marker is added, there is no greater degree of conceptual fusion than with any other transitive verb. In support of this, it may be noted that many languages without reflexive-derived middle systems can reflexivize such verbs in the normal way, e.g. English *I threw myself into a chair* or Latin *mē ferō* ‘I proceed or walk in procession’, literally ‘I carry myself’. (Kemmer 1993: 69)

However, I disagree with this characterisation, and believe that these examples are clear cases of anticausatives, for several reasons. In the first place, many translational motion anticausatives admit inanimate subjects, and hence cannot be considered reflexive (see (71)). From a semantic point of view, the characteristics of the movement can be identical regardless of the animacy of the subject (compare (71) and (72)), suggesting that they describe the same type of events.

- (71) Pero como le picas, se sale el aire, y por
but as DAT.3SG poke.PRS.2SG REFL.3 go.out.PRS.3SG the air and for
ahí se suele meter algo de aceite.
there REFL.3 tend.PRS.3SG get.in.INF some of oil
‘But since you poke it, the air goes out and some of the oil typically get in through there.’ (Valle de Cerrato, Palencia, COSER 3426)
- (72) Tus chicos han estao allí en la huerta y se
your.SG guys have.PRS.3PL be.PTCP there in the orchard and REFL.3
han metío en la huerta...
have.PRS.3PL get.in.PTCP in the orchard
‘Your guys have been there, in the orchard, and have got in the orchard...’
(Argamasilla de Alba, Ciudad Real, COSER 1404)

Furthermore, in these verbs, the actions of the subject in the causative version (see (73)) are not identical to the actions of the subject in the anticausative version (see (72)), while the actions performed by the subject in true reflexive constructions are identical to the non-reflexive variant.

- (73) He metido a tus chicos en la huerta.
have.PRS.1SG put.in.PTCP to your guys in the orchard
'I have made your guys get into the orchard.'

Another crucial piece of evidence supporting the claim that these verbs should be considered anticausatives and not true reflexives is that they do not admit the emphatic reinforcement *a sí mismo*:

- (74) ^{??}Tus chicos han estao allí en la huerta y se
your guys have.PRS.3PL be.PTCP there in the orchard and REFL.3
han metío a sí mismos en la huerta...
have.3PL get.in.PTCP to REFL.3 same.PL in the orchard
'Your guys have been there, in the orchard, and have got in the orchard...'

For all these reasons, the current analysis also includes translational motion events conveyed by verbs that show transitive-intransitive alternations. Subcorpus E yielded 8,610 examples from 49 different translational motion verbs (see Appendix 4), which were anticausative, antipassive, intransitive and non-reversible. The vast majority (80.7%, some 6,948 examples) were unmarked, as shown in Table 16. Most verbs attest marked examples (43/49), but a large number of them show also unmarked examples (29/49).

Table 16. Frequency of the RM with verbs of translational motion in subcorpus E

	Marked (RM)	Unmarked (No RM)	Total
Translational motion examples	1,662 (19.3%)	6,948 (80.7%)	8,610 (100%)
Translational motion verbs	43	29	49

3.3.2 The cognition middle

Mental events also tend to be marked by the middle marker in the languages of the world, making up what Kemmer calls the *cognition middle*. According to Maldonado (1999), if a language has a middle system, it will always mark mental and motion events.

Mental events usually have two participants; the experiencer, in whose mind the event takes place; and the stimulus, which causes the event to take place. Kemmer notes that the latter, however, is not necessarily always encoded in the sentence.

Although these events do not relate directly to the reflexive or reciprocal prototype, Kemmer claims that their relationship with the middle comes from the fact that the experiencer is understood as being both the initiator of the event (since it directs its attention to the stimulus) and the endpoint (since it is affected by the stimulus). In her words:

Mental events are semantically middle by virtue of the following properties: The chief participant involved, an Experiencer or Mental Source, is both the Initiator and the Endpoint of the event; there is no distinguishability of participants; and affectedness of the Initiator is part of the lexical semantics of verbs designating such events. (Kemmer 1993: 141)

Maldonado's (1999) account of middle mental events is somehow more elaborate. As mentioned above (Section 3.2.7), he argues for the existence of an intermediate category between active and passive experiencers, that is, middle experiencers. Middle experiencers do not show control over the action expressed by the verb, but they participate in it. This explains a contrast like the one in (75a), (b): in the unmarked variant (75a), the experiencer is the subject and the event has a more volitional reading than in the marked variant (75b), where the experiencer is encoded in the dative and shows no control over the event. That is, in this case, the presence of the RM reduces the global transitivity of the event. According to Maldonado, this is typically the case for cognition or perception events. The contrast in (75c), (d) shows the opposite situation. In the unmarked variant (75c), the experiencer is the object, which shows no control over the event. However, the global transitivity of the event is increased with the presence of the RM, as in (75d), since in this case the experiencer is also the subject, showing increased participation in the event. In order to distinguish between middle and reflexive constructions, the difference between two types of control (either over the object or the action) is crucial – the RM reduces control over the object in reflexive constructions (see (75e)), while it decreases control over the action in middle constructions (Maldonado 1999).²³

(75) (Maldonado 1999)

- | | | | | |
|----|------------------------|----------------|----------------|----------------------|
| a. | Te | olvidé. | | (Active experiencer) |
| | ACC.2SG | forget.PST.1SG | | |
| | 'I forgot you.' | | | |
| b. | Se | me | olvidó. | (Middle experiencer) |
| | REFL.3 | DAT.1SG | forget.PST.3SG | |
| | 'I forgot (about) it.' | | | |

23. This is in line with Kailuweit's claim that "The pronominal construction is – paradoxically – the more causative one of the two anticausative constructions" (Kailuweit (2011: 123).

- c. Lo que dijiste me entristeció. (Passive experiencer)
 the that say.PST.2SG ACC.1SG sadden.PST.3SG
 ‘What you said to me made me sad.’
- d. Me entristecí con lo que dijiste. (Middle experiencer)
 REFL.1SG sadden.PST.1SG with the that say.PST.2SG
 ‘I got sad with what you said to me.’
- e. ¡Yo me entristezco cuando se me pega la (Active experiencer)
 I REFL.1SG sadden.PRS.1SG when REFL.3 DAT.1SG hit.PRS.3SG the
 gana!
 wish
 ‘I get sad whenever I want to!’

Kemmer distinguishes three different types of mental events, namely: emotion events, cognition events, and perception events.²⁴ In the following, I analyse the presence of the RM in these three categories in the varieties under study.

3.3.2.1 *Emotion events*

Emotion events (such as ‘become frightened’, ‘be angry’, ‘hate’, ‘fear’, ‘be happy’, ‘worry’) show a high degree of affectedness of the experiencer and, compared to other mental events (cognition and perception), they are typically lower in volitionality. All these characteristics lead Kemmer to hypothesise that they are likely to be middle marked in the languages of the world.

She also notes that “[i]n some languages, notably modern Indo-European languages with middle systems (Romance, Germanic, Slavic languages), the middle marker serves in a number of middle situation types to mark causative-inchoative oppositions, e.g. Spanish *asustar* ‘frighten’, *asustar-se* ‘become frightened’” (Kemmer 1993: 132). It has already been noted above (Section 3.2.4) that a strong association between the conversive diathesis and emotion events is found in the varieties under study.

A further distinction between emotion verbs and “emotional response” verbs can be made (Maldonado 1999). In the former, the experiencer is encoded as the subject (*amar* ‘to love’, *detestar* ‘to hate’, *odiar* ‘to hate’), while in the latter the experiencer is coded as the object (*asustar* ‘to frighten’, *preocupar* ‘to worry’, *maravillar* ‘to amaze’). While the attachment of the RM to emotion verbs typically produces a reflexive construction, its attachment to emotional response verbs typically generates a middle construction, although elaborate contexts can create purely reflexive

24. Kemmer makes a further distinction between simple and complex mental events. In the latter, the stimulus is a dependent event, typically encoded by an embedded clause. I will not pursue this distinction here, since, as she admits, “in practice, it is difficult to investigate the marking properties of the two types independently” (Kemmer 1993: 127).

readings (Maldonado 1999). Most verbs addressed in this section fit within the emotional response category.

Before turning to the analysis of our data, it should be noted that I have also considered the verbs that Kemmer calls “emotive speech actions” (such as *quejarse* ‘to complain’, *lamentarse* ‘to complain’) within this class, as well as some related events that cannot strictly be considered speech actions (*reír(se)* ‘to laugh’, *llorar* ‘to cry’).

A total of 312 examples from 36 emotion verbs were documented in subcorpus E (see Appendix 4). Anticausative, conversive, de-objective, antipassive, intransitive and non-reversible verbs are attested in this category. While most of them show marked examples (33/36), just a few attest unmarked examples in the corpus (7/36). The RM is present in 94.2% of the cases (294 examples), while only 18 examples (5.8%) were unmarked, as can be seen in Table 17. An example is given in (76).

- (76) Es la abuela que se preocupe de ese nieto.
be.PRS.3SG the grandmother that REFL.3 worry.SBJV.3SG of that grandson.
‘Is the grandmother who worries about the grandson.’
(Humanes de Madrid, Madrid, COSER 2907)

Table 17. Frequency of the RM with verbs of emotion in subcorpus E

	Marked (RM)	Unmarked (No RM)	Total
Emotion examples	294 (94.2%)	18 (5.8%)	312 (100%)
Emotion verbs	33	7	36

3.3.2.2 Cognition events

Two typological generalisations regarding the middle marking of cognition events are in order. First, Kemmer notes that the cross-linguistic tendency towards middle marking is not as strong with cognition verbs as with emotion events. This fits with Maldonado’s (1999) claim that, since cognition verbs typically have an experiencer subject, the attachment of the RM tends to produce reflexive constructions (instead of middle ones). Second, despite the fact that middle marking of cognition events is less productive than middle marking of emotion events, Kemmer points out that there are no languages in her sample that only attest either marked emotion verbs or marked cognition verbs, which suggests that they are strongly connected.

There were 765 examples of 37 different cognition events (see Appendix 4) in subcorpus E, of which 701 (91.6%) showed the RM, as opposed to 64 unmarked examples (8.4%), as seen in Table 18. No intransitive verbs are found in this category, but all other syntactic categories (anticausative, conversive, de-objective, antipassive and non-reversible) are attested. All verbs show marked examples in the corpus, while only 7 provide unmarked examples.

- (77) Por ahí el gobierno se ha equivocado
for there the government REFL.3 have.PRS.3SG mistake.PTCP
‘The government has made a mistake with that.’(Zufre, Huelva, COSER 2122)

Table 18. Frequency of the RM with verbs of cognition in subcorpus E

	Marked (RM)	Unmarked (No RM)	Total
Cognition examples	701 (91.6%)	64 (8.4%)	765 (100%)
Cognition verbs	37	7	37

3.3.2.3 Perception events

Of all the situation types considered by Kemmer in the cognition middle, perception verbs are the most resistant to middle marking, maybe because their experiencer presents a low degree of affectedness, as she notes. Another cause that she proposes is the fact that these verbs

are subject to exceptionally strong pressure to be formally subsumed under the two-participant prototype (which translates in terms of marking to transitive morphosyntax with an unmarked verb root) because of the relatively higher volitionality of such verbs as against the cognition and emotion verbs.

(Kemmer 1993: 137)

Depending on the role of their subject, perception verbs can be divided into stimulus-based (*sonar* ‘to sound’) and experiencer-based. The latter can be further divided according to whether the experiencer is passively involved (*oír* ‘to hear’) or actively involved (*escuchar* ‘to listen’).²⁵ Kemmer hypothesises that stimulus-based perception verbs are the more relevant to the middle, on the basis of evidence from German (*es fühlt sich gut* ‘it feels good’, *es hört sich gut an* ‘it sounds good’) and she proposes that they are related to passive uses.

A total of 275 examples from five perception verbs (see Appendix 4) were documented in corpus E of the COSER interviews. Absolute and de-objective verbs are represented. Four verbs attest marked examples in the corpus, while three provide unmarked examples. An example is given in (78). Around two thirds of these examples (176, 64%) were marked, as opposed to the 99 examples (36%) that did not show the RM. This is illustrated in Table 19.

- (78) “Tú, cuando se arranque el toro, apercíbete.”
you when REFL.3 start.PRS.3SG the bull be.careful.IMP.2SG-REFL.2SG
‘When you see that the bull starts running, be careful.’
(Malagón, Ciudad Real, COSER 1414)

25. However, see Maldonado (1999) for a different interpretation of the contrast between *oír* and *escuchar*.

Table 19. Frequency of the RM with verbs of perception in subcorpus E

	Marked (RM)	Unmarked (No RM)	Total
Perception examples	176 (64%)	99 (36%)	275 (100%)
Perception verbs	4	3	5

3.3.3 Spontaneous events

Spontaneous events are “situations which designate changes of state of an entity, but in which no Agent entity receives coding” (Kemmer 1993: 142). They are often physical processes or changes in shape, size, colour, etc., that do not normally need an agent (‘die’, ‘grow’, ‘evaporate’, ‘vanish’) and that often correspond to the anticausative version of a causative-anticausative pair (Kemmer 1993). The relationship between these situation types and the middle lies in the fact that only one of the two potential participants (the endpoint) is encoded. This participant is also conceptualised as the initiator – this accounts for a low degree of elaboration of the participants (and therefore the event). However, spontaneous events differ from all middle situation types described thus far in that they typically involve inanimate entities. This accounts for another related difference – spontaneous events are characterised by “the complete lack of volitional initiation by the Patient of the event” (Kemmer 1993: 146). These differences justify considering spontaneous events as somehow being a distant relative to the other middle categories (Kemmer 1993), although they have also been considered as one of the central domains of the middle, in that they show a strong association with middle marking in Cushitic languages (Mous 2001).

They have also been placed on a continuum that includes passives and impersonals, since all these constructions have an affected object and a schematic inductive force (Maldonado 1999). The different degrees in the schematicity of the force determines the two poles of the continuum – in spontaneous events there is no energy in profile, while impersonal constructions convey a rather complex force, which must necessarily be a non-specified human. These events differ from other middle situations: in the constructions included on this continuum, the RM also focusses on the change of state of a thematic participant, but simplifies the event by reducing the representation of the inductive force:

Note that while [other] middle uses of the *se* form [...] are related to the recipients of the energy transmission, be it the direct or the indirect object, the passive / impersonal contrast only changes the degree of elaboration of the inductive force.
(Maldonado 1999: 272, my translation)

Kemmer (1993) also proposes a continuum, between events that are likely to be seen as produced by an external agent and those that are unlikely to have such an agent. She hypothesises that this distinction may be relevant to middle marking in some languages (see also Section 4.3).²⁶ She illustrates this hypothesis with Modern Icelandic, where the middle marker (-*st*) is preferred for the former type of events, while the latter are marked by an ancient Germanic affix that indicates change of state (-*n*-). This relates to the degree of responsibility/control of the subject of the spontaneous event – the more likely the interpretation of an external agent, the less likely the subject will be perceived as responsible. A similar split has been proposed for Spanish, with the suggestion that the presence of the RM decreases the volitionality of the subject. Dealing with the verb *caer(se)*, Arce Arenales, Axelrod & Fox note:

In other words, with *se*, in conjunction with some intransitive verbs that already suggest that the subject is affected by the action, the affectedness of the subject is increased, to the extent that any volitionality on the part of the subject seems to be ruled out. This case represents an intriguing illustration of the interaction between the middle diathesis domain and the lexical meaning of particular verbs.

(Arce Arenales, Axelrod & Fox 1994: 6)

A similar claim has been made by other authors, such as Portilla:

[the] fundamental difference between non-reflexive constructions (both intransitive and transitive) and optional reflexive constructions lies in the fact that the latter emphasise the anticausativeness of the participant of the event.

(Portilla 2007: 133, my translation)

However, the opposite claim has also been made for Romance languages. For anticausatives in Brazilian Portuguese, French, German and Spanish and on the basis of verbs like *empobrecer(se)* ‘to get poor(er)’ (see (79)), Kailuweit (2011) proposes that:

If a contrast of responsibility is coded, the subject of the pronominal construction of a single verb will denote an argument more responsible for the outcome of the action than the argument denoted by the subject of the bare intransitive construction of the same verb.

(Kailuweit 2011: 114)

(79) Juan ^{??}(*se*) empobreció por su propia culpa.

Juan REFL.3 impoverish.PST.3SG for his own fault

‘It was Juan’s own fault that he became poorer.’

(Kailuweit 2011: 115)

²⁶ A very similar distinction is made by Haspelmath (1993), who claims that the probability of using the anticausative expression (i.e. marking on the intransitive pair) or the causative expression (i.e. marking on the transitive pair) depends on the “probability of an outside force bringing about the event” (Haspelmath 1993: 103). Maldonado (1999) calls the two possibilities internal and external energy.

I will return to this question in Section 4.3. As regards our data, 2,657 examples from 167 different spontaneous verbs were documented in subcorpus E of the COSER interviews (see Appendix 4). Anticausative, antipassive, intransitive and non-reversible verbs are included in this category. As illustrated in Table 20, 1,107 of these examples (41.7%) showed the RM, as opposed to the 1,550 unmarked examples (58.3%). Most verbs attest marked examples (137/167), while a few less than half provide unmarked examples (71/167).

Table 20. Frequency of the RM with verbs expressing spontaneous events in subcorpus E

	Marked (RM)	Unmarked (No RM)	Total
Spontaneous examples	1,107 (41.7%)	1,550 (58.3%)	2,657 (100%)
Spontaneous verbs	137	71	167

3.3.4 Naturally reciprocal events

Just as some transitive verbs encode events that tend to be performed on oneself, there are also verbs that encode typically reciprocal events: “events that are either necessarily (e.g. ‘meet’) or, or [sic] else very frequently (e.g. ‘fight’, ‘kiss’) semantically reciprocal” (Kemmer 1993: 102).²⁷ Kemmer notes that these verbs tend to show middle-marking (as opposed to reciprocal marking, used in reciprocal constructions with other kinds of verbs) in many languages of the world, which justifies its inclusion as one of the typical middle situation events. As happens with grooming events, naturally reciprocal events in the varieties under study here are mostly encoded by syntactically reciprocal constructions, where the RM is a referential pronoun that can be doubled by a PP such as *el uno al otro* ‘one another’. For the same reason, these examples will be excluded from the following analysis.

27. Kemmer (1993) actually identifies a number of typologically recurrent meanings of these verbs in the sample she examined: “[A]ntagonistic action (‘fight’, ‘quarrel’, ‘wrestle’), what might be called affectionate actions (‘kiss’, ‘embrace’, ‘make love’), encountering and associated social actions (‘meet’, ‘greet’, ‘shake hands’), actions denoting unintentional physical contact (‘bump into’, ‘collide’) and physical convergence or proximity (‘touch’, ‘join’, ‘unite’, ‘be close together’) [...] acts of exchanging (‘trade’), of sharing and dividing (‘share’, ‘split hunting catch’), of agreement/disagreement, of interlocution (‘converse’, ‘argue’, ‘gossip’, ‘correspond’ [via letters]), and predicates of similarity and dissimilarity, among others” (Kemmer 1993: 104–5). See Knjazez (2007) for a different (but similar) semantic classification of lexical reciprocals. Other authors that have dealt with naturally reciprocal events or the similar concept of symmetric predicates are Nedjalkov (2007a) and Dimitriadis (2008).

The COSER interviews document a total of 844 of examples from 38 naturally reciprocal verbs (see Appendix 4) not found in syntactically reciprocal constructions. On the contrary, they are either anticausative or intransitive verbs. One example of a naturally reciprocal anticausative verb is given in (80). As seen in Table 21, most of these examples showed the RM (809, 95.1%) and most verbs provide marked examples (34/38), while only a few show unmarked examples (9/38).

- (80) Yo tengo una hija que no se ha casao.
I have.PRS.1SG a daughter that not REFL.3 have.PRS.3SG marry.PTCP
'I have a daughter who hasn't got married.' (Luzuriaga, Álava, COSER 107)

Table 21. Frequency of the RM with verbs expressing naturally reciprocal events in subcorpus E

	Marked (RM)	Unmarked (No RM)	Total
Reciprocal examples	803 (95.1%)	41 (4.9%)	844 (100%)
Reciprocal verbs	34	9	38

3.3.5 Verbs with no semantic ascription in Kemmer’s account

A considerable number of verbs (a total of 215, see Appendix 4)²⁸ showed the RM in our corpus despite the fact that they do not fit into any of Kemmer’s categories. In order to be as exhaustive as possible, I will consider all the COSER interviews in this section.

Table 22 shows that most of the verbs that do not fit into Kemmer’s categories (180; 83.7%) show one of the diathesis alternations discussed in Section 3.2, while only 35 (16.3%) are classified as either non-reversible or intransitive.²⁹ This suggests that the presence of the RM with these verbs is motivated syntactically and not semantically. Table 22 also shows the analogue figures for verbs that do fit into Kemmer’s categories, in order to analyse the productivity of the syntactic parameters. The antipassive and the de-objective diatheses show that syntactic factors play an important role in explaining the presence of the RM, since verbs with no semantic ascription outnumber those with semantic ascription in these categories. While this is not so for anticausative verbs, the total number of verbs with no semantic

²⁸. Only 197 verb lexemes appear in Appendix 4, since here I have duplicated lexemes that appear in more than one change of diathesis, for the sake of the analysis below.

²⁹. As noted above, the question as to whether non-reversible verbs belong to a syntactic alternation or not can only be determined through a historical study.

ascription in this case is so high that the relevance of the syntactic alternation needs to be acknowledged. The converse is exceptional among syntactic alternation, showing that almost the whole class can be ascribed to Kemmer’s categories (due to the strong association between emotion verbs and the converse diathesis, as observed in Section 3.2.4). This makes it virtually impossible to distinguish which parameter (the semantic or the syntactic) plays a stronger role in the presence of the RM in this category.

Table 22. Verbs by ascription to Kemmer’s semantic categories and diathesis alternation in the COSER data

Syntactic type	Verbs outside Kemmer’s categories	Verbs inside Kemmer’s categories	Total
Absolute	7 (70%)	3 (30%)	10 (100%)
Anticausative	122 (25.9%)	349 (74.1%)	471 (100%)
Antipassive	17 (58.6%)	12 (41.4%)	29 (100%)
Converse	5 (12.8%)	34 (87.2%)	39 (100%)
De-objective	29 (61.7%)	18 (38.3%)	47 (100%)
Non-reversible	16 (27.6%)	42 (72.4%)	58 (100%)
Intransitive (no change of valency)	19 (30.6%)	43 (69.4%)	62 (100%)

Reflexive intransitive verbs with no semantic ascription are the most interesting category, since neither the presence of either a diathesis alternation nor a semantic notion associated with the middle domain can account for their middle marking. In Chapter 5, which focusses on reflexive intransitive verbs, a possible explanation in terms of analogy is offered.

Finally, my data show that a number of verbs that cannot be ascribed to any of Kemmer’s categories can be grouped into two functional domains,³⁰ namely pseudo-copulative and auxiliary verbs. These will be discussed in Section 3.3.6 and 3.3.7 respectively.

30. I prefer to use the term “functional” rather than “semantic” here, since their similarities in content are rather abstract and they are better described in terms of grammatical content.

3.3.6 Pseudocopulative or attributive verbs

Pseudocopulative verbs are not included in Kemmer's account. Nonetheless, they have been singled out in several analyses of Spanish reflexive constructions (Alcina & Blecua 1975; Fernández Ramírez 1986; Gómez Torrego 1992). Pseudocopulative or attributive verbs are characterised by taking a compulsory secondary predicate (Rodríguez Arrizabalaga 2005) and are semantically homogenous to some extent, with meanings similar to 'to be', 'to become' and 'to find, to consider'. The most productive of these verbs (*hacerse*, *volverse*, *ponerse*, *quedar(se)* 'to become') are also change-of-state verbs – their secondary predicate specifies the final state of that change (see (81)) (Van Gorp 2014). That is, these verbs, in combination with their secondary predicate, could fit within Kemmer's category of spontaneous events.

- (81) a. Lo otro se queda como líquido.
 the other REFL.3 remain.PRS.3SG like liquid
 'The other stuff becomes sort of liquid.'
 (Madrigal de las Altas Torres, Ávila, COSER 0609)
- b. Y si yo me volviera joven, no me
 and if I REFL.1SG return.PST.SBJV.1SG young not REFL.1SG
 casaba con un hombre del campo.
 marry.PST.1SG with a man of.the country
 'If I were young again, I wouldn't marry a man from the countryside.'
 (Portillo de Solórzano, Cantabria, COSER 1219)

However, in some cases their meaning is closer to denoting a mere state (see (82)) – most of these cases could be considered cognition events. I have, however, disregarded these semantic differences and analysed them together, following the tradition of Spanish and Galician grammarians of grouping pseudocopulative or attributive constructions together as a class of their own.

- (82) a. Yo no quería novio, yo me encontraba mu niña.
 I not want.PST.1SG boyfriend I REFL.1SG find.PST.1SG very child
 'I didn't want a boyfriend, I thought I was too young.'
 (Antequera, Málaga, COSER 3001)
- b. Yo nunca me he sentío más que a otras, ni
 I never REFL.1SG have.PRS.1SG feel.PTCP more than to other nor
 na.
 nothing
 'I have never felt I was more special than other girls or anything like that.'
 (Antequera, Málaga, COSER 3001)

Note also that the degree of distinguishability of participants is variable. For instance, examples like (82a) are close to a reflexive construction. A clear-cut division between reflexive attributive constructions and middle attributive constructions has been proposed (Rodríguez Arrizabalaga 2005). In the former, the RM is a reflexive pronoun, with referential meaning, while in the latter it is a verbal particle (with intransitivising value). According to this account, reflexive attributive constructions are formed with (physical or mental) perception verbs or speech verbs (*ver* 'to see', *sentir* 'to feel', *encontrar* 'to find', *saber* 'to know', etc.), admit the *a sí mismo* PP and may show the alternation illustrated in (83). On the other hand, middle attributive constructions show verbs that simply are aspectual variants of the copula *ser* (*volverse*, *ponerse*, *hacerse*, *quedar(se)* 'to become', etc.) and participate in the anticausative alternation (except for *quedar(se)*).

- (83) (Rodríguez Arrizabalaga 2005)
- a. Sabía que ella era diferente.
 know.PST.3SG that she be.PST.3SG different
 'She knew that she was different.'
 - b. Se sabía diferente.
 REFL.3 know.PST.3SG different
 'She knew herself to be different.'

However, the examples documented in COSER do not fit as nicely into this picture. First, the productivity of both constructions is very uneven. A total of 1,246 examples of pseudocopulative constructions were documented in subcorpus E of the COSER interviews, with only ten different verbs. Six of them belong in the middle attributive construction according to Rodríguez Arrizabalaga: *quedar* (937), *poner* (194), *hacer* (82), *volver* (6), *tornar* (2) and *echar* (1) 'to become'.³¹ The remaining examples would seem to fit within the reflexive attributive construction, since their verbs convey physical and mental perception: *ver* 'to see' (12), *encontrar* 'to find' (8), *dar* 'to give' (2) and *sentir* 'to feel' (2). However, most of these examples do not seem to admit the *a sí mismo* PP (see (84)). That is, although some perception and cognition verbs appear to admit the *a sí mismo* PP more easily in this construction, the verbs documented in the COSER interviews show a high degree of lexicalisation and do not fit well into the reflexive construction. This justifies grouping all such examples as middle constructions:³²

31. The total number of occurrences in subcorpus E are indicated in brackets.

32. Note that some of the examples considered in this section were also considered in Section 3.3.4, since they are symmetrical predicates, such as *hacerse novios* 'to become a romantic pair'.

- (84) a. Se acercaba y se dio (??a sí
REFL.3 approximate.PST.3SG and REFL.3 give.PST.3SG to REFL.3
misma) ya tan perdida que apretó el gatillo, bum, y
self already so lost that press.PST.3SG the trigger boom and
lo mata. (Jérica, Castellón, COSER 1307)
ACC.3SG kill.PRS.3SG
'And he was getting closer and closer and she thought her situation was so
desperate that she pressed the trigger and, boom!, she kills him.'
- b. Yo soy mayor de edad, y me encuentro (??a mí
I be.PRS.1SG older of age and REFL.1 find.PRS.1SG to REFL.1SG
misma) en mi derecho de venir cuando quiera.
self in my right of come.INF when want.SBJV.1SG
'I've come of age and I have the right to come whenever I want.'
(Arjona, Jaén, COSER 2301)

If we look at the presence of the RM, we see that most of them (62.2%, 775 examples) were marked, as opposed to the remaining 471 unmarked examples, which account for the remaining 37.8% (Table 23). Almost all verbs show marked examples (9/10), while only three verbs provide unmarked examples.

Table 23. Frequency of the RM with pseudocopulative verbs in subcorpus E

	Marked (RM)	Unmarked (No RM)	Total
Pseudocopulative examples	775 (62.2%)	471 (37.8%)	1,246 (100%)
Pseudocopulative verbs	9	3	10

A look at their distribution according to their changes of diathesis is rather interesting, since it makes clear that intransitive verbs are to blame for the unmarked examples. This is even more interesting when we look at the specific verbs in each class, since the only intransitive verb in the sample is *quedar* ‘to become’, which was marked half of the time (see Table 24). In the anticausative pseudocopulative class, however, there were nine different verbs and they were marked 99% of the time. Even more interestingly, the only three unmarked examples were all found in Galicia (see (85)).

Table 24. Frequency of the RM with pseudocopulative verbs in subcorpus E

	Anticausative	Intransitive	Total
Marked (RM)	306 (99%)	469 (50.1%)	775 (62.2%)
Unmarked (No RM)	3 (1%)	468 (49.9%)	471 (37.8%)
Total	309 (100)	937 (100%)	1,246 (100%)

- (85) a. Hai que metelo no congelador. Se non
 have.PRS.3SG that get.in.INF-ACC.3SG in.the freezer if not
torna todo amarillo e despois non nada, non
 turn.PRS.3SG all yellow and after no nothing no
 vale.
 worth.PRS.3SG
 'It has to be put in the freezer. Otherwise it turns all yellow and it is of no
 use.'
 (Orellán, La Coruña, COSER 2401)³³
- b. E cunha pota, [A-PIIn: dar barreno] hasta que ponía espeso
 and with.a pot give.INF drill until that put.PST.3SG thick
 de todo.
 of all
 'And with a pot, one had to stir until it thickened completely'.
 (O Malladoiro, Orense, COSER 3303)

3.3.7 The RM in auxiliary verbs

Another interesting functional class surfaced in my data. Thirteen verbs (see Appendix 4) were found in the reflexive form while functioning as auxiliaries of aspectual periphrasis (see (86)), accounting for 167 examples in the whole corpus.³⁴ Most of these verbs show reflexive usages in other contexts, suggesting that their reflexive auxiliary uses have been grammaticalised out of these reflexive contexts. Note, however, that the original interpretation of the verb is ruled out: their only meaning is aspectual. Moreover, they can also be the source of other uses, since the omission of the main verb in the discourse allows them to lexicalise or to spread to paratactic periphrasis (see (87)).³⁵

- (86) a. Yo era pequeno ya, me agarraba a segar y
 I be.PST1SG little already REFL.1SG grab.PST1SG to reap.INF and
 eso.
 that
 'I was little and I started reaping and all that.'
 (Aguaviva de la Vega, Soria, COSER 3923)

33. Unfortunately, only two examples of *tornar(se)* were found in the data. However, it is worth noting that Larochette (1939) observes that both *tornar* and *tornarse* are possible in Old Spanish, while only *ponerse* and *fazerse* are documented in the intransitive construction.

34. Only marked auxiliaries are taken into account in this section, which is why no quantitative analysis of the frequency of marking is provided.

35. For a discussion of paratactic periphrasis in European languages, such as the ones in (87a), see Coseriu (1977).

- b. [Las gallinas] las quitamos porque luego ya los
 ACC.3PL remove.PST.1PL because later already the
 muchachos se dieron en tirarlas
 boys REFL.3 give.PST.3PL in remove.INF-ACC.3PL
 ‘We removed [the hens] because afterwards the boy decided to remove
 them.’ (El Barco de Ávila, Ávila, COSER 0607)
- c. ¡No cuentas cosas... de pena, déjate de contar
 no tell.SBJV.2SG things of pity leave.IMP.2SG-REFL2SG of tell.INF
 cosas...!
 things
 ‘Don’t say sad things, stop telling us things...!’
 (La Nava de Ricomalillo, Toledo, COSER 4214)
- (87) a. Porque se agarraban cuatro o cinco hombres y lo
 because REFL.3 grab.PST.3PL four or five men and ACC.3SG
 echaban en una mesa.
 throw.PST.3PL on a table
 ‘Because four or five men got ready and threw it on the table.’
 (Moraleja, Cáceres, COSER 1014)
- b. [estudiar] Luego él se puso y por correspondencia,
 later he REFL.3 put.PST.3SG and for correspondence
 estando trabajando se sacó..., porque él empezó y
 being working REFL.3 obtain.PST.3SG because he start.PST.3SG and
 estuvo hasta el segundo.
 be.PST.3SG until the second
 ‘Afterwards he got started and by mail, while working, he got..., because
 he started and stayed until second grade.’
 (Higueruela, Albacete, COSER 0211)

3.4 The middle voice in Spanish and Galician: A basic or a derived voice?

The data examined in this chapter provide sufficient evidence to evaluate the question stated in Section 3.1: is the middle voice in the varieties under study a derived voice based on changes of diathesis or, on the contrary, is it a basic voice based on semantic parameters? If the former option is true, we should expect consistent middle-marking across constructions that show a change of diathesis. On the contrary, if the latter option is true and middle-marking is strongly associated with situation types, consistent middle-marking across the verbs that convey such situation types is expected.

Although our data show that none of these extreme situations applies, semantic classes seem to be a poorer predictor of the presence of the RM than diathesis alternations. This is shown by the generalised linear mixed model provided in Table 25. As can be seen, the greater effect is found for the presence of a diathesis alternation, while belonging to one of Kemmer's classes has virtually no effect on the presence of the RM.

Table 25. Generalised linear mixed model: Presence of the RM (subcorpus E) depending on the presence of a diathesis alternation [reference level: No change of diathesis], the belonging to one of Kemmer's classes [reference level: Not in Kemmer's classes] (fixed factors), and the place (random factor)

Effect	Term	Estimate	Std. error	Statistic	Conf. low	Conf. high
fixed	(Intercept)	-1.716	0.087	-19.71	-1.887	-1.546
fixed	Derived voice: with change of diathesis	2.941	0.043	68.071	2.856	3.025
fixed	Basic voice: in Kemmer's classes	-0.029	0.043	-0.684	-0.114	0.055
random	sd (Intercept) (COSERID)	0.626	NA	NA	NA	NA

The same conclusion can be drawn from Figure 8, which shows the negative linear correlation between the probability of the presence of the RM in a given semantic category and the probability that the examples found in the same semantic category show no change of diathesis.³⁶ That is, the presence of the RM is inversely

³⁶ Because the data considered in this study cannot be seen as independent, I calculated the probabilities in Figure 8 by obtaining the mean of the probabilities in each of the two large areas that arose from the geographical distribution of anticausative verbs: north-western varieties (Galicia and Asturias) and the rest of the territory. This compensates to some degree for the inequalities between the sample. Theoretically it would have been more accurate to calculate the mean either by verb or by location, but the low raw numbers for some verbs and/or locations (when grouped by semantic category) would have resulted in more noise in the data, rather than in less independent results.

proportional to the number of examples with no change of valency in a given category. Pearson's R, which measures both the direction and the strength of a correlation, is -0.75 , indicating a rather large negative correlation. R^2 is 0.57 , which means both that the prediction error of our regression line is 57% smaller than when using the mean, and that the probability of a semantic class of not showing a valency change explains 57% of the variance in our data, also quite high values. This confirms that the presence of the RM is, to a great extent but not entirely, an indicator of one of the derived diathesis analysed in Section 3.2 (anticausative, de-objective, antipassive or conversive). Therefore, it also suggests that marking situation types which show a low degree of elaboration of events is not the main function of the RM in these varieties.

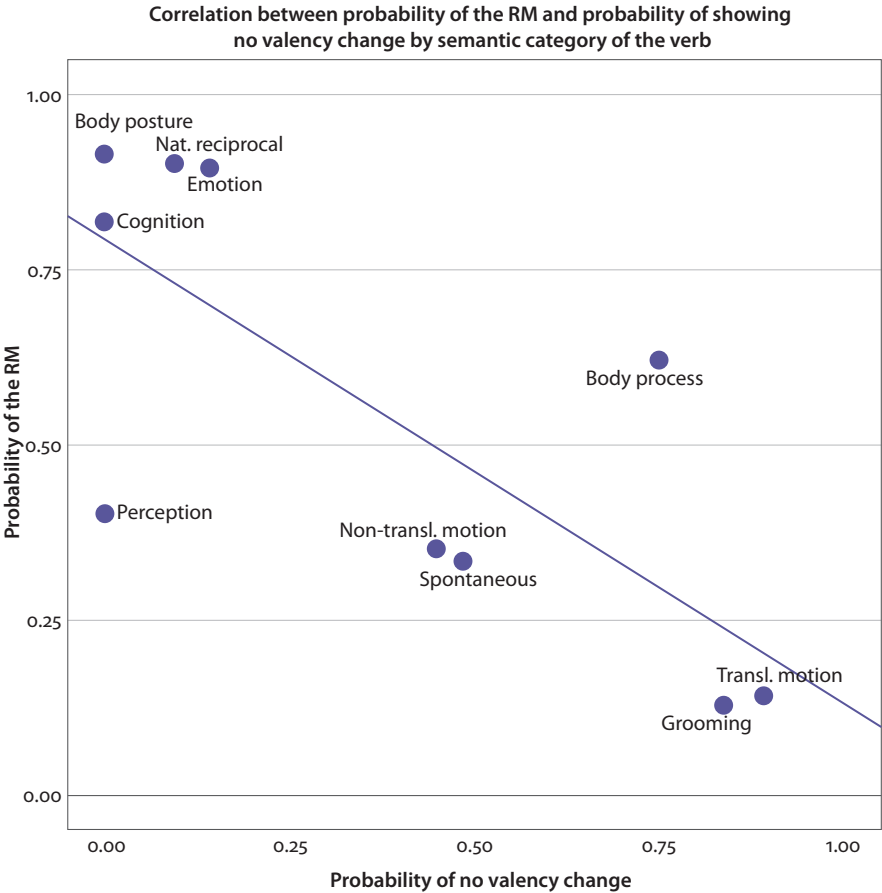


Figure 8. Correlation between the relative frequency of the RM and number of verbs with no change of diathesis by semantic category of the verb in subcorpus E
Pearson's $R = -0.75$, $R^2 = 0.57$

We can safely conclude that the main function of the RM is to indicate a change of diathesis. This, of course, is not new – the intransitivising function of the RM has been noted in several studies (Babcock 1970; Cartagena 1972; Martínez 1981; Vera Luján 1996).³⁷ However, analyses concluding that the RM affects the verb's diathesis have also been rejected by some scholars, especially by those aiming to give a comprehensive account of all reflexive constructions (Maldonado 1999; García 1975). On this view, the fact that the RM attaches to intransitive verbs (and also to transitive ones, as will be seen in Chapter 6) with no change of valency is evidence against considering intransitivisation as a possible uniting function of the RM.

The variationist and diasystematic approach I take here, however, does not require the existence of a uniting function for the RM – it only requires the existence of one or more links between different functions or contexts of the RM. From this point of view, our data strongly suggest that the intransitivising function(s) of the RM is not only older than its attachment to non-derived intransitive (and transitive) verbs, since it is more frequent in the former, but also different. Such a difference is crucial. The highly frequent presence of the RM in intransitivising contexts, as opposed to its low frequency and heterogeneity in contexts where no change of valency is produced, is an indicator that these two categories are not to be treated equally. This relates to Kailuweit's (2011) remark that situations where variation is found should not be equated with situations with no variation. Although variation was found in every category investigated, their frequencies (and hence productivity) are considerably different.

This is not to say that semantic approaches to the middle voice have nothing to say. The COSER data show that the situation types which Kemmer (1993) considers to be related to the middle domain are well-represented within the middle-marked structures documented in the varieties under study. This is nicely illustrated in Figure 9, which represents the semantic categories of the verbs that also show some kind of change of diathesis. Less than one third (27.6%) of the verbs were not ascribed to any middle situation type.

37. Other authors have noted that the RM “produces” not only intransitive sentences but also transitive sentences where the subject position is the one that appears to be blocked (cf. reflexive impersonal constructions: *Se castigó a los ladrones*). On the basis of this evidence, Bogard (1999b) prefers to classify the RM as an “argumental blocking” device – very similar to Cartagena's (1972) proposal that the RM is a “non-obliquity marker”.

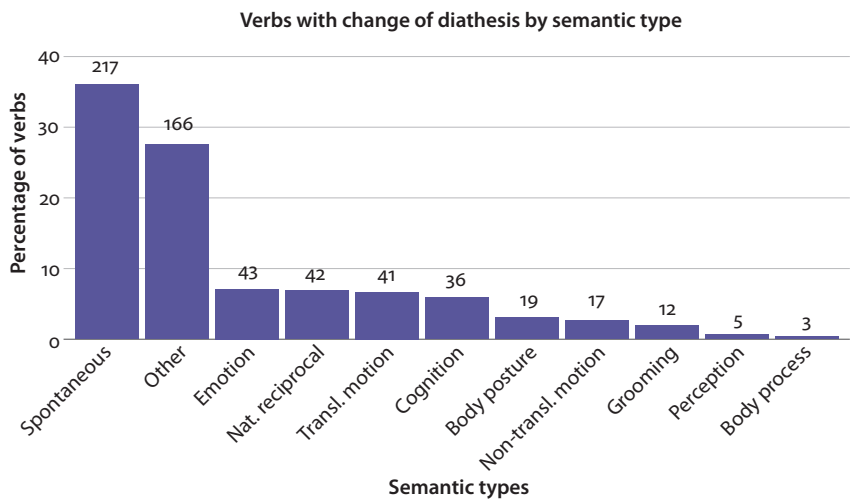


Figure 9. Semantic categories of marked verbs with a change of diathesis in the COSER data

Our data also allow for testing another one of Kemmer’s predictions, regarding the semantic distance to the reflexive prototype of each situation type. According to her account, we expect that the shorter the semantic distance from the reflexive prototype, the higher the frequency of reflexive marking. In order to test this prediction, I will only consider examples in the anticausative or the conversive alternation in the NE corpus, since these are the two diatheses that most affect the role of the external agent, which is either deleted or demoted, and also the ones that were most exhaustively collected in my data. Figure 10 shows the relative frequency of the RM in these examples, by semantic class. Semantic classes have been ordered according to the relative frequency of marking, so as to facilitate their visualisation.³⁸

If we compare these results with Kemmer’s semantic map (see Figure 2, Section 1.4.4 above) we see that the semantic distances she hypothesized coincide in many cases with our data. On the one hand, translational motion events ($N = 1,829$) are the situation type of the body middle that favour the least middle-marking, as her account predicts, being outnumbered by non-translational events ($N = 55$) and change in body posture verbs ($N = 870$). However, the relative order of non-translational motion and change in body posture events does not fit Kemmer’s account, since the latter show reflexive marking most often. Another of Kemmer’s

38. Perception events are not included because no anticausative or conversive verb conveyed such events. Body processes were also excluded because only 7 examples of anticausative verbs were documented in this category.

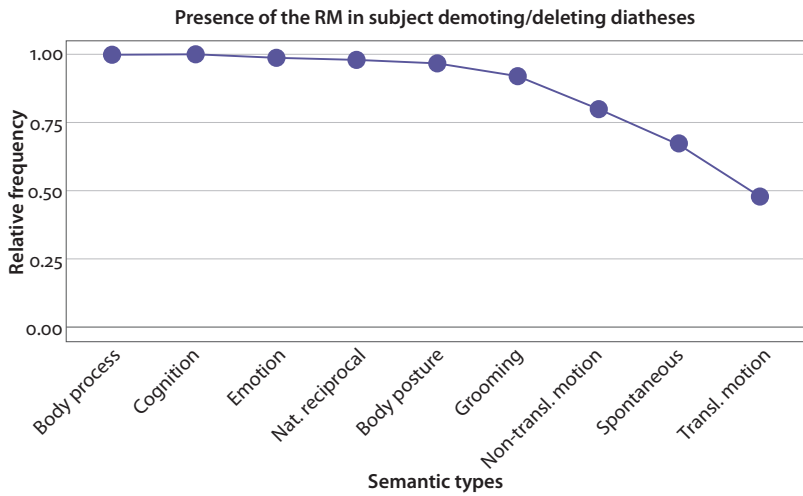


Figure 10. Relative frequency of the RM with anticausative and conversive verbs by semantic category in the COSER data (corpus NE)

predictions borne out in our data is the fact that spontaneous events ($N = 2,841$) are somehow distant from the prototype, since they show a rather low frequency of marking in Figure 10. To sum up, the most relevant conclusion drawn from Figure 10 is the relative low-frequency of marking with spontaneous and translational motion events, which fits Kemmer's account.

Anticausative verbs in Spanish

4.1 Introduction

This chapter presents a fine-grained analysis of the restrictions that affect the presence of the RM in (some) constructions with a change of diathesis. In Chapter 3 I argued that the main function of the RM is to indicate a change of diathesis. However, the presence of the RM in this function only reaches 82% of cases, that is, it is not systematic. This chapter will seek to explore which linguistic factors determine this variation. The analysis confirms that the verb lexeme, the probability of the verb appearing in transitive contexts, the presence of a dative, the animacy of the subject and the geographical area, as well as some specific constructions have an effect in the probability of the appearance of the RM. I discuss all these factors in the light of previous approaches and suggest some alternative explanations that do not entail giving such a prominent role to aspectual explanations.

As already mentioned, most authors who have looked at this variation have focussed on anticausative verbs. I will also restrict myself to the anticausative alternation (although I will refer to the other diathesis when sufficient data are available), this for two reasons. First, anticausative verbs are the most abundant in my data. Second, my data suggest that variation in marking in the antipassive and de-objective diatheses is lexically conditioned, while the conversive diathesis does show virtually absolute reflexive marking in my data. Since the productivity of the RM as a marker of valency reduction is dependent on the geographical area, as shown in Chapter 3, I will distinguish two different varieties for the analysis. Following the eloquent distribution in Map 4, I will group the interviews from Galicia and Asturias under the label “north-western varieties”, as opposed to the (less catchy) label “Rest of the territory”.

This chapter begins with a general overview of the different verbs that enter into the alternation, plus their marking, in Section 4.2, in that it has been claimed that there are three types of anticausative verbs in terms of their marking patterns (Schäfer 2008, Cennamo & Jezek 2011). In Section 4.3 I explore on the notion of externally and internally caused events, which some authors consider a relevant parameter in understanding the marking patterns of anticausatives (Mendikoetxea 1999a, 2000; Cidrás Escáneo 2003; Kailuweit 2011). Sections 4.4 and 4.5 are devoted to analysing the effect of aspectual parameters and event structure respectively, both

of which have been said to have an effect on the presence of the RM (Kailuweit 2011; Vivanco 2016). In Section 4.6 I discuss the role of the animacy of the subject, which has been considered crucial in the studies which adopt a diachronic perspective (Monge 1955; Waltereit 2000). In order to provide a joint analysis of these factors, I conduct a generalised linear mixed model, the results of which are discussed in Section 4.7, followed by a summary in Section 4.8.

4.2 Different kinds of anticausative verbs

In the recent literature, it is common to classify anticausative verbs into three groups, depending on their marking properties (Schäfer 2008, Cennamo & Jezek 2011; Vivanco 2016). Class A anticausatives are always marked with the RM, class B anticausatives (also called labile anticausatives) are never marked with the RM, and class C anticausatives show variable marking. Identifying which verbs belong in which class in Spanish, however, is problematic due to the important dialectal, social and probably even stylistic variation found in this language. As far as I am aware, no study has addressed this question with a corpus-based methodology. Thus, it is not unusual to find contradictory claims in the literature. For instance, while Maldonado (1999) considers that *enfermarse* is obligatorily reflexive, Vivanco (2016) classifies it as a class B anticausative. The fact that the former author is Mexican and the latter Spanish accounts for this divergence of judgement. However, one may wonder about the degree of applicability of these accounts if they do not address such variation systematically. I hope to overcome some of these shortcomings by examining an exhaustively analysed corpus, and by restricting my generalisations to the varieties under study. In what follows I investigate the validity of this three-part division for the two main varieties considered here.

In order to classify verbs depending on their marking possibilities, we need to select verbs that appear a number of times in the sample, so that we can estimate the frequency with which they take the RM. I have selected verbs that showed ten or more occurrences in each variety and calculated the proportion of marked uses for each of these verbs. Figure 11 shows two boxplots, which graphically depict the dispersion of this probability distribution by area.¹ First, it provides further

1. Boxplots provide a visual summary of the statistics of a sample, with special attention to its dispersion: (1) the bold horizontal line gives the median; (2) the two horizontal lines that delimit the box represent the values of the first and third quartile, i.e. the minimum and maximum value of the central 50% of the data; (3) the vertical lines (the so-called whiskers) represent the highest and lowest value that are within 1.5 interquartile ranges in the box – the interquartile range is

confirmation of our interpretation of Map 4, above, showing that these two varieties behave in completely different ways regarding the frequency of the RM. It is clear that this frequency is far lower in the north-western area. In north-western varieties, the mean of the relative frequency of the RM is 0.4, as compared to 0.8 in the rest of the territory. The median, a more representative measure, differs even more strongly in these two areas: 0.15 in north-western varieties (i.e. almost never marked) and 0.991 in the rest of the territory (i.e. almost always marked). Second, it also shows that the dispersion of our sample is far larger in the northwest than in the rest of the area. However, note that this dispersion is also affected by the fact that only 21 verbs occurred ten or more times in the north-western area (as opposed to the 96 verbs documented in the rest of the study area).

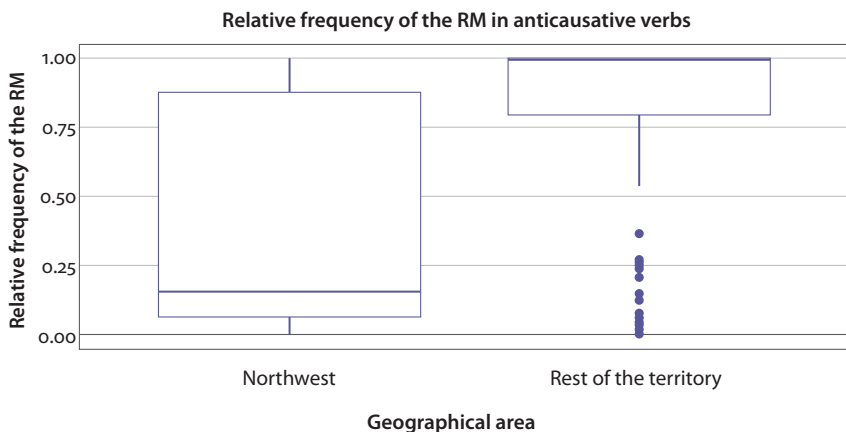


Figure 11. Boxplots for the relative frequency of the RM with anticausative verbs by geographical area in the COSER data

North-western area: Min. = 0; 1st Qu. = 0.05882; Median = 0.15; Mean = 0.39819; 3rd Qu. = 0.86957; Max. = 1

Rest of the territory: Min. = 0; 1st Qu. = 0.7941; Median = 0.991; Mean = 0.8238; 3rd Qu. = 1; Max. = 1

the difference between the first and third quartile; (4) the second and third measures just described are measures of sample dispersion that justify the representation of the remaining values as outliers, indicated by the dots outside the whiskers. The numerical values of both the mean, the median, the first and third quartiles, and the lowest and highest values, are given below the boxplot.

Less synthetic ways of representing the dispersion of our data also provide interesting insights. Figure 12 plots the relative frequency of the RM in each of these verbs individually. That is, each dot is one of the verbs considered in our sample and its position along the x-axis indicates the relative frequency with which it was marked by the RM. From the perspective of the distinction between class A, B and C anticausatives, Figure 12 shows opposite situations depending on the variety. In the northwest, class B is larger than class A, although most verbs show variation (that is, they belong to class C). However, the extreme gap in the middle suggests that a more sensible distinction would group verbs that almost always take the RM and verbs that almost never take it. The situation in the rest of the territory is rather different. A large majority of the 96 verbs considered belong to class A, very few to class B, and a non-negligible number of verbs show variation (class C). The gap between the two poles is not so extreme (although note again the smaller sample size of the north-western area), but still apparent, with a noticeable area in the centre of the plot with very few or no verbs at all.

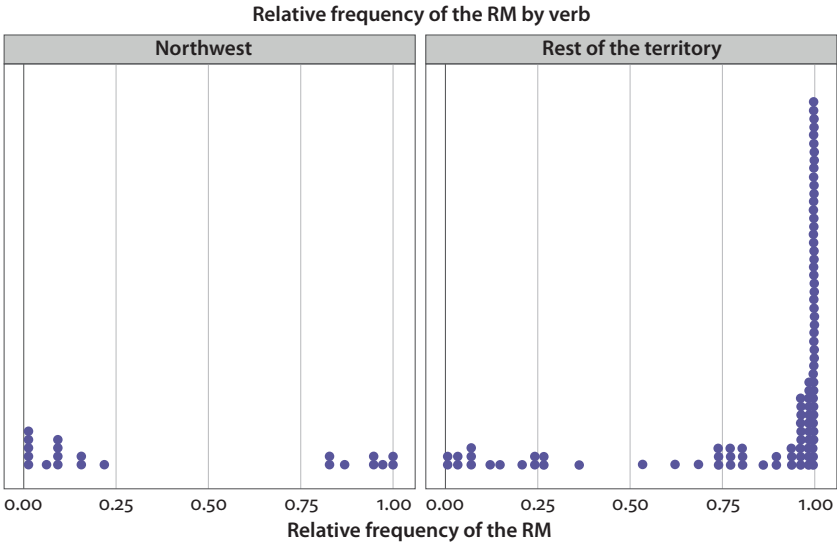


Figure 12. The 96 most frequent anticausative verbs in the COSER data plotted by relative frequency of the RM (rest of the territory)

This is seen more clearly in Figure 13, which represents the empirical cumulative distribution of our data. The x-axis represents the relative frequency of the RM, while the y-axis represents the percentage of data of one specific variable value plus all the smaller values. Every angle indicates a specific value that was found in the sample. An example will clarify this: the angle located approximately above the 0.23

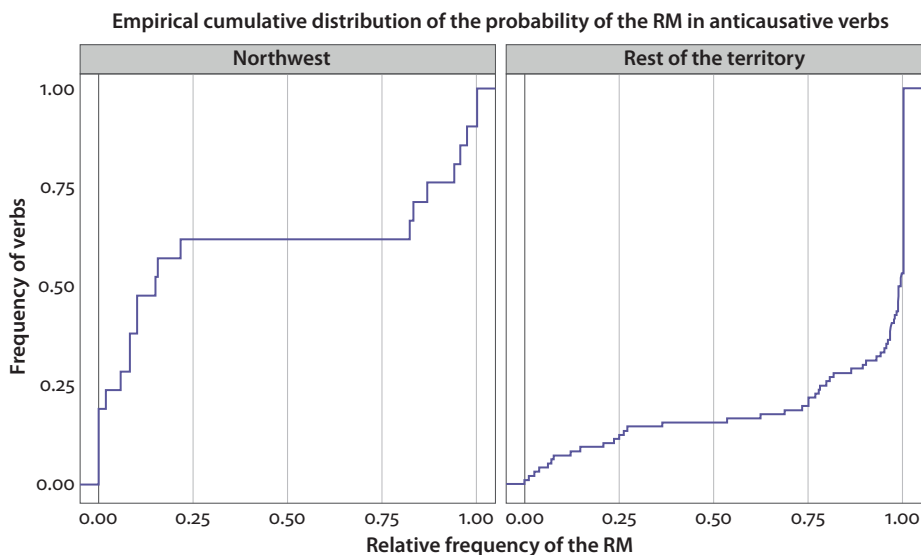


Figure 13. Empirical cumulative distribution for the relative frequency of the RM with the 96 most frequent anticausative verbs in the COSER data (rest of the territory)

value on the x-axis in the graph for the northwest represents one variable value. Its value is indicated on the x-axis: it represents a frequency of the RM of about 0.23 (23%). Its position on the y-axis indicates that around 63% (0.63) of the verbs in our north-western sample showed the RM as often or less than 23% of the time. The large horizontal segment shows the large gap identified above. As we can see, in the rest of the territory this gap is not so clear, being “interrupted” by some angles – those are the few verbs that take the RM with a frequency between 26% and 74%.

It is clear from both representations that there is a break between verbs that tend to be marked with the RM (more than 60%–75% of the time) and verbs that clearly prefer to be unmarked (marked in less than 25–30% of the cases). Figure 13 also shows that the former account for about 40% of verbs in the northwest and around 80% of verbs in the rest of the territory, where in fact 50% of verbs are always marked by the RM. Conversely, verbs that tend to be unmarked amount to more than 60% of the most frequent verbs found in the northwest, but only 20% of the most frequent verbs in the rest of the territory.

In the next two Subsections I take a more detailed look at these verbs in both varieties.

4.2.1 Rest of the territory

The 15 verbs that take the RM less than 40% of the time are: *bajar* ‘to go down’, *cambiar* ‘to change, be different’, *cocer* ‘to boil’, *desaparecer* ‘to disappear’, *empezar* ‘to start’, *engordar* ‘to gain weight’, *escurrir* ‘to drain’, *fermentar* ‘to ferment’, *hervir* ‘to boil’, *mejorar* ‘to improve’, *pasar* ‘to pass’, *pasear* ‘to go for walk’, *sangrar* ‘to bleed’, *subir* ‘to go up’ and *variar* ‘to change, be different’. The remaining 81 verbs take the RM in more than 50% of cases (all except *parar* ‘to stop’ take it more than 60% of the time) and can be found in Appendix 4. As noted above, the break between the two groups regarding the relative frequency of the RM suggests that they behave like two different classes. In this sense, the distinction between class A, B and C anticausatives might not be adequate for analysing real speech, where a strict classification of each verb might not be possible. Actually, all of these verbs (except *variar* ‘to change’) showed marked examples in the COSER interviews. With this evidence, it is safer to assume that there are only two classes of anticausatives (or that there are two types of C anticausatives), namely, those that tend to be marked most of the time, and those that tend to be unmarked most of the time. Classifying all verbs that show some variation under the same category (i.e. class C anticausatives) ignores the fact that they do not behave as a homogenous class.

Let us take a closer look at the verbs that tend to be unmarked. Figure 14 provides an overview of the frequency of the RM with each of these verbs. Some of them (*engordar* ‘to gain weight’, *hervir* ‘to boil’, *cambiar* ‘to change, be different’) are typically considered class B (labile) anticausatives (Vivanco 2016, see also Mendikoetxea 1999a, although the latter does not follow a three-fold classification of anticausative verbs). At first sight, most of these verbs fit the aspectual classification of Vivanco’s (2016) class B verbs, that is, they are either activities (*engordar* ‘to gain weight’, *pasar* ‘to pass’, *sangrar* ‘to bleed’, *bajar* ‘to go down’, *pasear* ‘to go for walk’, *subir* ‘to go up’) or achievements (*cocer* ‘to boil’, *desaparecer* ‘to disappear’, *hervir* ‘to boil’). However, our group also includes accomplishments (*empezar* ‘to start’, *escurrir* ‘to drain’, *fermentar* ‘to ferment’, *mejorar* ‘to improve’, *cambiar* ‘to change, be different’, *variar* ‘to change, be different’). It should be said, of course, that the aspectual characteristics of each verb cannot be defined without the specific context in which it appears. For instance, verbs like *hervir* or *cocer* ‘to boil’ are typically understood as achievements if their subject is *el agua* ‘water’, but they can also be accomplishments if their subject is food, as with *las patatas* ‘the potatoes’.

It is also worth mentioning that some of the verbs that are only rarely marked are very atypical anticausatives. *Bajar* ‘to go down’, *pasar* ‘to pass’, *pasear* ‘to go for walk’, *subir* ‘to go up’ and are not change-of-state verbs, as most anticausatives are (although they convey a change of location). Furthermore, *pasar*, which in the uses considered in this section encodes a translational movement, can show both causative and anticausative uses, yet the characteristics of the object in the former

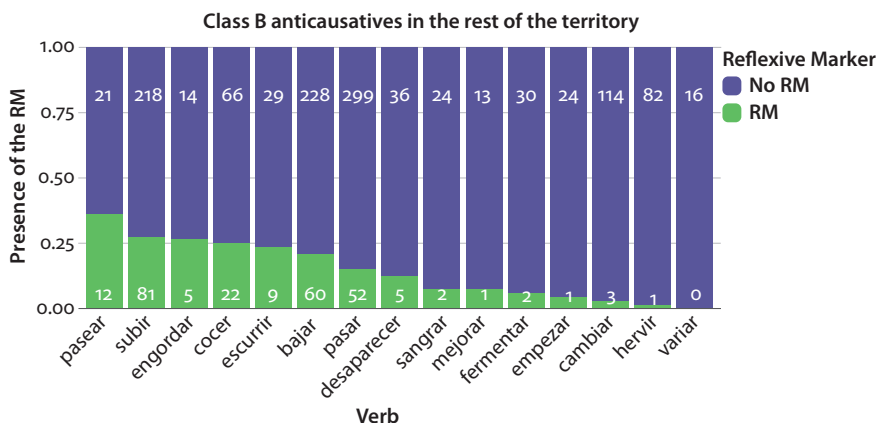


Figure 14. Frequency of the RM in the 16 most frequent verbs that tend to be unmarked in the COSER data

(typically inanimate) do not typically coincide with the characteristics of the subject in the latter (typically animate), although all possibilities exist, see (88). This relates to the fact that the causative-anticausative alternation does not apply on the sentential level, but on the verb level (see Section 3.2.1 above). Ultimately, what is relevant for allowing the anticausative alternation is the fact that the verb admits the causative interpretation in the transitive form (even if it is mostly agentive, as with *cortar* ‘to cut’) and that the intransitive form allows for an external cause.

- (88) a. Cuando ya pasaron la bandeja.
 when already pass.PST.3PL the tray
 ‘When they passed with the tray.’ (Enguera, Valencia, COSER 4310)
- b. Cuando pasó la bandeja.
 When pass.PST.3SG the tray
 ‘When the tray passed.’
- c. Si queréis los paso al huerto.
 if want.PRS.2PL ACC.2PL pass.PRS.1SG to.the orchard
 ‘If you want to, I can lead you into the orchard.’
 (Povedilla, Albacete, COSER 0222)
- d. Pasamos al huerto
 pass.PST.1PL to.the orchard
 ‘We went into the orchard.’

Finally, it should be noted that *bajar* ‘to go down’ and *subir* ‘to go up’ are often considered as intransitive verbs in studies on the so-called aspectual *se*. Accordingly, I will address them in more detail in Chapter 5, together with *pasar* ‘to pass’ and *pasear* ‘to go for a walk’ which shows a similar behaviour. However, since other intransitive verbs typically subsumed in that category, such as *ir*, *venir* or *salir*, do

not show causative alternations, I decided also to include them here, for the sake of systematicity.

Although in most cases the reflexive variant has the same meaning as its unmarked counterpart, the case of *hervirse* (see (89a)) and *sangrarse* (see (89b)) are interesting, since a small semantic nuance is added in these cases. *Sangrarse* means ‘to bleed out’ in the two examples in which I documented it (cf. *desangrarse* ‘to bleed out’) and *hervirse* means ‘to boil over’, the same meaning that Kailuweit (2011) documents in speakers from Argentina and Venezuela.² In both cases the use of the RM seems to be connected with a higher affectedness of the subject, one of the typical functions of the middle voice. *Sangrarse* also differs aspectually from *sangrar* – while the marked form is telic, the unmarked form is not.

- (89) a. Tengo eso ahí al sol y al sol está mal.
have.PRS.1SG that there at.the sun and at.the sun be.PRS.3SG bad
Aquí, al sol se me van a cabrear y se
here at.the sun REFL.3 DAT.1SG go.PRS.3PL to upset.INF and REFL.3
me van a hervir.
me go.PRS.3PL to boil.INF
'I have that thing in the sun and that's not good. Here, in the sun they are
going to get upset and boil over.'
(Santervás de la Vega, Palencia, COSER 3423)
- b. [¿Y la madre murió?] Sí. Na más nacer. Se
yes nothing more be.born.INF REFL.3
sangró.
bleed.PST.3SG
'Yes, she did. Right after giving birth, she bled to death.'
(Valeria (Las Valeras), Cuenca, COSER 1634)

It is fair to ask why these verbs and not others show such resistance to reflexive marking. One possible explanation might lie in their etymology, since the RM might be interpreted as a marker of *derived intransitivity* (cf. Haspelmath 1993). Accordingly, verbs which show the anticausative alternation now, but which were intransitive in Latin, might be seen as not showing derived intransitivity, but rather *derived transitivity*. Indeed, some of these verbs were primarily intransitive in Latin. This is the case with *escurrir* 'to drain' (<EXCŪRRŌ, ĒRE, intr., Segura Munguía 1985) and *hervir* 'to boil' (<FERVĒŌ, ĒRE, intr., Segura Munguía 1985). Similarly, *desaparecer* 'to disappear' is a derived form of the intransitive verb APPĀRĒŌ, ĒRE (Segura Munguía 1985). However, *cocer* 'to boil' (<CŌQUŌ, ĒRE), *sangrar* 'to bleed' (<SANGUĪNŌ, ĀRE) and *subir* (SŪBĒŌ, ĪRE) 'to go up' were all transitive and intransitive

2. Although (89a) is the only example of *hervirse* found in the COSER interviews (and in the questionnaire), it seems to be a rather common form in some American varieties (see also Kailuweit 2012). An informal search of Twitter shows that the emphasised meaning ‘to boil over’

in Latin, meaning that they should not be seen as cases of derived intransitivity in Romance. The fact that we actually find reflexive examples is a sign of the strong tendency of some of these varieties to generalise the pattern “unmarked = transitive/causative verb – marked = intransitive/anticausative verb”.

Some of these verbs are Romance creations (*bajar* ‘to go down’, *empezar* ‘to start’, *engordar* ‘to gain weight’, *pasar* ‘to pass’ and *pasear* ‘to go for a walk’). Although intransitive uses of *empezar* are documented from the earliest records, Corominas & Pascual (1981) note that the higher frequency of the transitive usage indicates that this is the basic structure of the verb. No indication is given of the transitivity characteristics of *engordar*, *pasar* or *pasear* by these authors.

The remaining verbs were transitive in Latin, so they are clear cases of derived intransitivity. This is the case with *mejorar* ‘to improve’ (<MĚLIÖRÖ, ĀRE, tr., Segura Munguía 1985), *fermentar* ‘to ferment’ (<FERMĚNTÖ, ĀRE, tr., Segura Munguía 1985 – the related intransitive was FERMENTERSCÖ, OLD 1982) and *cambiar* ‘to change’ (<CAMBIĀRE, tr., Segura Munguía 1985; OLD 1982).

These mixed results suggest that the historical origin of these verbs is not enough to explain the resistance of such verbs to reflexive marking, so they cannot be considered fossilised from an earlier stage in which the productivity of the RM to mark change of valency was not as high (see also Vivanco 2016: 178).

4.2.2 North-western varieties

As regards north-western varieties, Figure 15 shows the frequency of reflexive marking in each of the 21 most frequently attested verbs. The break between verbs that tend to be marked and verbs that tend not to be so is very obviously illustrated

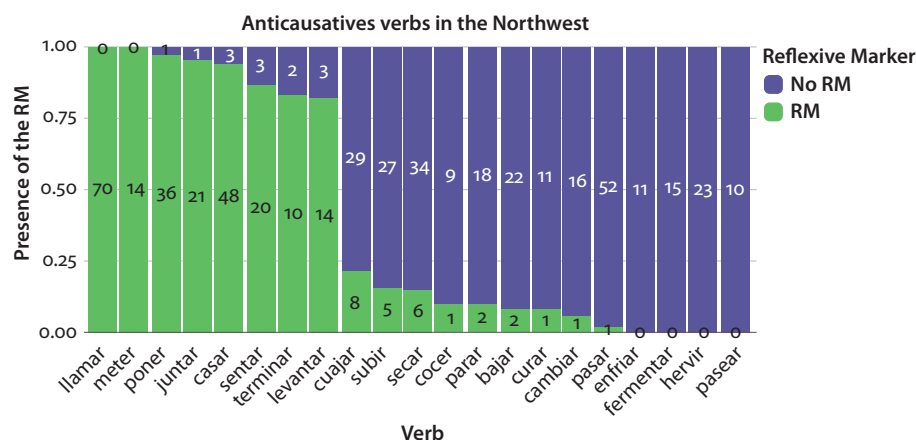


Figure 15. Frequency of the RM with the 21 most frequent anticausative verbs in the north-western area in the COSER data

by the strong contrast between *levantar* 'to get up' and *cuajar* 'to clot'. Figure 15 also shows that some verbs that are typically marked in the rest of the territory, like *enfriar* 'to cool', *curar* 'to heal/to cure', *secar* 'to dry' or *cuajar* 'to clot' are typically unmarked in northwestern varieties (a few examples are given in (90)). This difference lies at the root of the contrast between the two varieties and suggests that they represent different stages of a similar evolutionary process, in which north-western varieties represent an earlier stage.

- (90) a. [la sangre] se pon gorda y entós hay que
REFL.3 put.PRS.3SG thick and then have.PRS.3SG that
lo revolver pa que vaya enfriando un poco.
ACC.3SG shake.INF for that go.SBJV.3SG cooling a little
‘[The blood] gets thick, so then you need to move it so that it cools down
a little bit.’ (Alea – Linares (Ribadesella), Asturias, COSER 506)
- b. [la hierba] Darle la vuelta, para que secara por
give.INF-DAT.3SG the turn for that dry.SBJV.3SG for
el otro lado
the other side
‘Turn ‘[the grass] around, so that it dries on the other side.’
(Zas, La Coruña, COSER 2403)

This idea of the two areas representing different stages of a similar evolutionary process is further supported by the data in Table 26, where the frequency of the RM with these verbs in north-western varieties is compared to its frequency in the rest of the territory. As can be seen, the frequency of the RM is consistently higher in the rest of the territory. Disregarding a few exceptions where the difference is rather small (*cambiar*, *llamar*, *meter*), the only salient exception to this pattern seems to be *terminar*. This is consistent with a process of lexical diffusion which is more advanced in the rest of the territory.

Table 26. Frequency of the RM in the COSER data with the 21 verbs that both areas have in common

Verbo	Northwest	Rest of the territory
<i>bajar</i>	8.3%	20.8%
<i>cambiar</i>	5.9%	2.6%
<i>casar</i>	94.1%	98.7%
<i>cocer</i>	10%	25%
<i>cuajar</i>	21.6%	79.9%
<i>curar</i>	8.3%	75%
<i>enfriar</i>	0%	81.7%
<i>fermentar</i>	0%	6.2%
<i>hervir</i>	0%	1.2%

Table 26. (continued)

Verbo	Northwest	Rest of the territory
<i>juntar</i>	95.5%	97.5%
<i>levantar</i>	82.4%	95.9%
<i>llamar</i>	100%	99.5%
<i>meter</i>	100%	99.7%
<i>parar</i>	10%	53.5%
<i>pasar</i>	1.9%	14.8%
<i>pasear</i>	0%	36.4%
<i>poner</i>	97.3%	99.4%
<i>secar</i>	15%	77.7%
<i>sentar</i>	87%	98.8%
<i>subir</i>	15.6%	27.1%
<i>terminar</i>	83.3%	62.4%

Let us turn now to a discussion of those semantic parameters deemed in the literature to be most relevant in explaining the alternation between marked and unmarked anticausatives.

4.3 Externally and internally caused events

The concept of causation requires greater discussion, in that it has been linked to the variable marking patterns of anticausatives. In what follows, I briefly discuss proposals that suggest that the semantics of causation determines whether an anticausative verb is marked or not, and test these using the corpus data.

By definition, every verb that exhibits anticausative alternation might be interpreted as having an external cause, since it can be also expressed by a transitive verb with causative meaning. However, in his account of marked and unmarked anticausatives in German, Schäfer (2008) argues strongly against this, proposing that only the latter admit causers with an unintentional reading. In my opinion, his arguments are valid inasmuch as they are used to defend a specific semantic decomposition of anticausative verbs (which have a causative subevent, even if it lacks an external argument). However, I believe his arguments do not prove that anticausatives cannot take external forces, as he also claims. In the following I discuss and reject the two main arguments that he provides.

The first test he uses to reject the claim that anticausatives (both marked and unmarked) have an implicit external argument involves focussing on a causing entity, as shown in (91). (91b) is not too surprising, since anticausatives reject external agents, as noted in Section 3.2.1. However, verbs that exhibit anticausative

alternation normally accept a cause subject in the transitive version, so why is (91c) non-felicitous? I believe this is so because the cause is encoded in the nominative, which clashes with the intransitivity of the anticausative verb. The sentence improves (both in German and Spanish) if the cause is introduced by a preposition such as *durch* 'by' or *por* 'by' (see (92)). Schäfer (2008) actually notes that anticausatives admit causes introduced by *durch*, which he takes as a test to suggest that they have a causative event (which in his analysis does not necessarily imply having an external cause).

(91) (Schäfer 2008)

- a. Der Teller zerbrach
'The plate broke.'
- b. ??Aber es war nicht Peter, sondern Maria
'But it was not Peter but Mary.'
- c. ??Aber es war nicht das Erdbeben sondern der Wind
'But it was not the earthquake but the wind.'

- (92) a. [Der Teller zerbrach] Aber es war nicht durch das Erdbeben sondern durch der Wind
'[The plate broke] but it was not by the earthquake but by the wind.'
- b. El plato se rompió
the plate REFL.3 break.PST.3SG
'The plate broke.'
- c. Pero no fue por el terremoto, sino por el viento
'But it was not by the earthquake but by the wind.'

The second test that Schäfer (2008) uses to argue for the lack of an external argument in anticausative events involves the addition of *von selbst* 'by itself', as in (93). All German anticausatives can take this phrase, which "stresses the fact that no external force can be identified" (Schäfer 2008: 56). However, this does not constitute proof that external forces *cannot* be identified with anticausative verbs, but proof that internal forces can be identified. In fact, Mendikoetxea (1999a) uses the same proof to argue for the opposite claim: that Spanish verbs taking the phrase *por sí solo* 'by itself, on its own' accept an external cause, because this phrase excludes an external force that could be interpreted in another way, by stressing the fact that the event was internally caused.

- (93) Der Teller zerbrach von selbst
the plate break.PST.3SG by self
'The plate broke by itself.'
- Der Geige verstimmte sich von selbst
the violin go.out.of.tune.PST.3SG REFL.3 by self
'The violin went out of tune by itself.'

Actually, the situation that Schäfer is trying to explain, i.e. the fact that (some) anticausatives might take datives with an unintentional causer reading, also shows that anticausatives can be understood as having an external cause, since this is what unintentional causers are:

- (94) a. Die Vase zerbrach dem Hans
 The vase break.PST.3SG the.DAT Hans
 'Hans broke the vase unintentionally.'
- b. El jarrón se le rompió a Juan
 The vase REFL.3 DAT.3SG break.PST.3SG to Juan
 'Juan broke the vase unintentionally.'

From this, I conclude that all anticausatives can be interpreted both as having an external cause or as excluding it (i.e. having an internal cause). The difference between externally and internally caused events has been used in several studies to explain the different productivity of the RM. It has been claimed that there are two types of intransitive change-of-state verbs, namely, events with an external cause and events with an internal cause (see Mendikoetxea 1999a, 2000 and García 1975 for Spanish, and Cidrás Escáneo 2003 for Galician). The former have three prominent characteristics: (1) they show the causative-anticausative alternation (which explains the fact that they are externally caused); (2) the anticausative variant is marked by the RM; and (3) the anticausative variant is clearly unaccusative. On the contrary, intransitive change-of-state verbs with internal causation do not exhibit causative-anticausative alternation, are not marked by the RM, and it is not always clear whether they are unaccusative or unergative. The concept of "internal cause" actually refers to the idea that the change of state is produced because of internal and inherent properties of the subject, but the subject is not interpreted as intervening in the causation of the event. Thus, changes of state with an internal cause cannot take the phrase *por sí solo* 'on their own', which is admissible with changes of state with an external cause:

- (95) (Mendikoetxea 1999a)
- a. El barco se hundió por sí solo.
 the ship REFL.3 sink.PST.3SG by REFL.3 alone
 'The ship sank on its own.'
- b. La puerta se abrió por sí sola.
 the door REFL.3 open.PST.3SG by REFL.3 alone
 'The door opened on its own.'
- c. ?El niño creció por sí solo.
 the child grow.PST.3SG by REFL.3 alone
 'The child grew up on his own.'

- d. ??Juan empeoró por sí solo.
 Juan worsen.PST.3SG by REFL.3 alone
 'Juan got worse on his own.'

The distinction between externally and internally caused events has been used to explain the variable marking patterns of the so-called class C anticausatives, like *aclarar* and *oscurecer* (Mendikoetxea 1999a).³ When these verbs can take the transitive alternation and admit *por sí solo*, they can also take the RM in their intransitive version. The opposite is true when they refer to the weather, for instance (see (96)).

(96) (Mendikoetxea 1999a)

- a. El día ha aclarado.
 the day have.PRS.3SG clear.PTCP
 'The weather has cleared up.'
- b. ??El sol ha aclarado el día.
 the sun have.PRS.3SG clear.PTCP the day
 'The sun has cleared the day/weather up.'
- c. El jersey se ha aclarado por sí solo.
 the jersey REFL.3 have.PRS.3SG clear.PTCP by REFL.3 alone
 'The jersey has faded on its own.'
- d. {Juan / La lejía / La lavadora / El sol} ha
 Juan the bleach the washing.machine the sun have.PRS.3SG
 aclarado el jersey.
 clear.PTCP the jersey
 '{Juan / Bleach / The washing machine / The sun} has faded the jersey.'

Class B anticausatives, however, are more problematic in this account, since they exhibit anticausative alternation but do not take the RM. *Hervir* 'to boil', for instance, seems to resist the presence of the RM, while *cocerse* 'to boil', a close synonym, does take it. Being so similar in meaning, it is hard to argue that *hervir* is an internally caused verb, whereas *cocer* is externally caused. Moreover, since both verbs show the causative alternation, (see (97)), the distinction would only hold if we were to accept that events with an internal cause could enter the causative alternation. Mendikoetxea argues this is only so if "the situation can be manipulated so as to create the conditions to initiate the change-of-state process, as happens with the verb *hervir*" (Mendikoetxea 1999a: 1601–2, my translation).

3. The notion of external and internal causes is reminiscent of Kemmer's (1993) concept of low elaboration of the participants (see Section 1.4.4). Externally caused events show a higher degree of distinguishability of both participants and events than internally cause events. Within Kemmer's account this means that they are closer to the reflexive and are thus more likely to show reflexive marking, as predicted above.

(97) (Mendikoetxea 1999a)

- a. Juan ha hervido la leche.
Juan have.PRS.3SG boil.PTCP the milk
'Juan has boiled the milk.'
- b. La leche (*se) ha hervido.
the milk REFL.3 have.PRS.3SG boil.PTCP
'The milk has boiled.'
- c. Juan ha cocido las patatas.
Juan have.PRS.3SG boil.PTCP the potatoes
'Juan has boiled the potatoes.'
- d. Las patatas se han cocido.
the potatoes REFL.3 have.PRS.3PL boil.PTCP
'The potatoes have boiled.'

Similarly, the existence of verbs with different marking behaviours in related languages, like Fr. *fondre* and Sp. *fundir(se)* (both of which admit the causative alternation), is dealt with by postulating that *fondre* is a verb of change of state with an internal cause, while *fundirse* is a verb of change of state with an external cause (Mendikoetxea 1999a). However, this conclusion seems not to be based on the semantic characteristics of the event, but on the presence of the RM, which makes the description of the distinction somewhat circular (see also Kailuweit 2012).

A different syntactic argument has been put forward to justify the behaviour of class B anticausatives (Sánchez López 2002, Zribi-Hertz 1982). On this view, the transitive-intransitive alternation displayed by these verbs (in (97a, b) and (98a, b)) is not exactly causative. The fact that class B predicates can enter the periphrastic causative construction (98c) would show that the change of diathesis of labile verbs like *subir*, *aumentar*, *hervir*, etc., does not correspond to a true causative alternation, since class A predicates like *cerrar* do not allow this construction (99c) (Sánchez López 2002). However, the alleged unacceptability of (99b) seems to be due to the fact that, since *cerrar* is not a labile verb, the intransitive version must also take the RM in periphrastic causative constructions. The example in (100) shows that *cerrarse* can also be inserted in periphrastic causatives with a non-agentive cause. That is, *cerrarse* behaves like intransitive *subir* in this respect, the only difference being that *subir* is unmarked.

(98) (Sánchez López 2002)

- a. La inflación subió los precios.
the inflation raise.PST.3SG the prices
'Inflation raised prices.'
- b. Los precios subieron.
the prices rise.PST.3PL
'Prices rose.'

- c. La inflación hizo subir los precios.
the inflation make.PST.3SG raise.INF the prices
'Inflation made prices raise.'

(99) (Sánchez López 2002)

- a. El viento cerró la ventana.
the wind close.PST3SG the window
'The wind closed the window.'
- b. La ventana se cerró.
the window REFL.3 close.PST3SG
'The window closed.'
- c. *El viento hizo cerrar la ventana.
the wind make.PST3SG close.INF the window
'The wind made the window close.'

- (100) El viento hizo cerrarse la puerta de mi cuarto, se
the wind make.PST.3SG close.INF-REFL.3 the door of my room REFL.3
rompió la cerradura y no puedo entrar.
break.PST.3SG the lock and not can.PRS.1SG go.in.INF
'The wind made the room's door close, the lock broke and now I can't go in.'
(February 25th 2012, Buenos Aires, Argentina. Tweet.)

In my opinion, the main problem with these accounts that use the distinction between internally and externally caused events to explain the marking patterns of anticausatives is that they are too categorical and hence the predictions they make are too strong. As shown above, several verbs (although a minority in the "rest of the territory") might undergo the causative-anticausative alternation and be either never or seldom marked. In this sense, Kailuweit's (2011) approach seems more appropriate. Noting that labile causative-anticausative verbs are to be found in all Romance languages (as well as German), and that they are not verbs with the same meanings across languages, he adopts a less strict approach, considering the contrast between external and internal causation to be one criterion among several others: "*If the difference between internal and external causation is coded, external causation will be coded with the pronominal construction and not with the bare intransitive construction*" (Kailuweit 2011: 110, emphasis added). In this sense, the typological generalisation that "a factor favoring the anticausative expression type is the probability of an outside force bringing about the event. Conversely, the causative expression type is favored if the event is quite likely to happen even if no outside force is present" (Haspelmath 1993: 103) has the potential to explain the behaviour of the two types of anticausatives (marked vs. unmarked).

Vivanco (2016) applies Haspelmath's approach to Spanish in a quantitative study. She establishes a relationship between the proportion of causative uses of a

verb and its behaviour regarding anticausative marking. In a sample of 30 verbs, she observes that verbs with more than 50% of causative uses belong to class A, while verbs with less than 30% of causative uses tend to belong to class B. Class C anticausatives are mostly found in the gap between these two extremes.⁴ Only four of our typically unmarked verbs are found in her sample: *cambiar*, *engordar*, *hervir* and *mejorar*. *Cambiar*, *engordar* and *hervir* show low percentages of causative uses (10.68%, 30.61%, 15.19%, respectively) and are classified in class B by Vivanco. On the other hand, *mejorar* shows a higher degree of causative uses (44.05%) and is classified as a class C anticausative by Vivanco. Our data clearly suggest that this verb is very resistant to the presence of the RM (cf. Figure 14) and casts some doubt on the suitability of a three-fold classification that does not take into account the large differences within class C.

In order to give a picture that better reflects these differences on the probability of marking, I replicated (and expanded) Vivanco's (2016) study using 83 of the most frequently anticausative verbs in our sample (see Appendix 5). A random sample ($n = 150$) of each of these verbs was extracted from the corpus *European Spanish Web 2011* (eseuTenTen11); the probability of these verbs showing a transitive use was calculated, as was the probability of them being marked in intransitive uses.⁵ This is a more data-driven approach, since I do not classify the verbs in class A, B or C introspectively. Figure 16 plots each of these verbs according to their probability of showing the RM in anticausative uses (y-axis) and of its probability of having transitive uses (x-axis). The distribution of the points shows no clear correlation between these two parameters, and that "class A" verbs (i.e. a verb that always takes the RM) show a full range of probability of being used in a transitive construction.

4. This distribution points to the fact that reflexive marking is the default strategy for anticausative verbs in Standard Spanish (Vivanco 2016), which fits my findings for most peninsular varieties, but not for north-western ones.

5. I have chosen eseuTenTen for two main reasons. First, it is a extremely large web corpus, which offers a sample of several styles of language ranging from blogs to newspapers. Second, its search engine offers an easy way of exporting the data, which was not true for most reference corpora of Spanish at the time of the study. I extracted all examples (or up to 100,000) of 83 verbs of our sample, 100,000 being the limit established by the corpus, and randomly selected 150 tokens of each, which I later annotated manually. In calculating the probability of transitive uses, examples in the passive reflexive, in causative constructions, in purely reflexive constructions, and with omitted objects were considered transitive: examples in the impersonal passive or with prepositional objects were considered intransitive. For the probability of the RM, only anticausative uses were considered. Ambiguous examples, auxiliary uses, and instances where the form was not a verb were disregarded. A summary of the dataset can be found in Appendix 5.

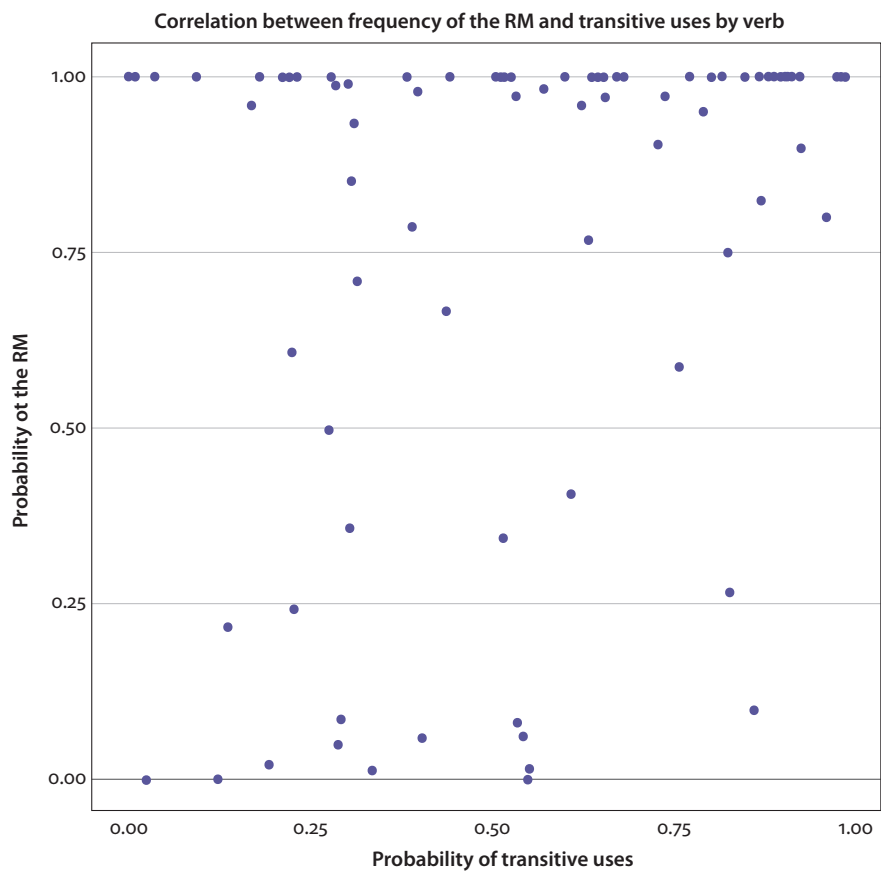


Figure 16. (Lack of) correlation between the relative frequency of the RM and the probability of appearing in a transitive context in corpus eseüfTenTen11

However, this lack of correlation might have several explanations. On the one hand, it might be that this relationship is found only in a subset of anticausatives, as Kailuweit (2011) has proposed. On the other hand, it might be that this operationalisation of “the probability of an outside force bringing about the event” (Haspelmath 1993: 103) is problematic. While Haspelmath proposes that this probability is directly linked to the proportion of causative uses vs. anticausative uses, causes can also be codified in anticausative uses (via prepositional adjuncts or datives, as said above). Finally, the correlation may exist but be mediated by a number of other factors, such as those that will be discussed in the remainder of this chapter. This seems to be the case, as discussed in Section 4.7.

4.4 The aspectual properties of the predicate

Recent proposals explain the variable distribution of the RM in anticausatives on the basis of aspectual motivations. For instance, Kailuweit (2012) maintains that, if marking codes an aspectual contrast, then the reflexive verb will be telic and the bare intransitive will be atelic.⁶ A more detailed proposal has been put forward by Vivanco (2016), for whom the presence or absence of the RM in class C anticausatives is crucially determined by the event structure of the predicate. She identifies two classes of class C anticausatives depending on how their aspectual properties affect the distribution of the RM: durative change-of-state verbs and inherently telic accomplishments. Although I will not discuss her proposal in detail here, I will examine the specific construction that she deems to be the strongest argument in favour of attributing an aspectual function to the RM.

Durative change-of-state verbs such as *envejecer* and *enfriar* consist of a chain of accomplishments, each one followed by a (comparative) result state. According to Vivanco, the RM is sensitive to the presence of a final result state, which is the reason why even class A verbs such as *secarse* reject the presence of the RM “when subordinated to a verb that selects atelic predicates, since in that case the last accomplishment is cancelled” (Vivanco 2016: 312):

- (101) Poner la ropa a secar(*se)
 put.INF the clothes to dry.INF-REFL.3
 ‘Put clothes to dry.’

The contrast illustrated in (101) is quite robust and absolutely confirmed in my data. My sample attests 171 examples with durative change-of-state verbs with either an open or closed scale in constructions similar to (101), that is, “*poner* or a similar verb + *a* + infinitive”, all of which were unmarked.⁷ The same goes for the structure “*tener* ‘to have’ + gerundio”, which is semantically very similar, again, all examples in our sample (8/8) were unmarked. However, I do not believe that attributing the absence of the RM in these examples to the event structure selected by the predicate

6. The split between marked and unmarked anticausatives has been explained by using the distinction between telic and atelic events in other Romance languages, see Schäfer (2008) and Cennamo & Jezek (2011) for a critical view.

7. I only considered durative infinitives in this count – achievements were disregarded. The inflected verbs documented in this construction mostly encode the idea that the subject of the anticausative either is at a location or is carried to a location: *botar* (only Gal.) ‘throw’, *colgar* ‘hang’, *dejar* ‘leave’, *echar* ‘throw’, *estar* ‘be’, *estirar* ‘stretch’, *llevar* ‘carry’, *meter* ‘put into’, *poner* ‘put’, *quedar* ‘stay’, *sacar* ‘get out’, *tender* ‘hang’, *traer* ‘bring’. I also considered cases with *dejar* where the preposition *a* was missing.

is correct. First of all, I disagree with Vivanco's (2016) characterisation of these structures, shown in (102), i.e. as rejecting a final result state. While I agree that this structure does not focus on the final result state but on the process, I believe that examples such as (103) clearly imply that the final result state is reached. That is, even if the final result state is not in focus, neither is it rejected.

(102) [poner a → [L1 → E+SECO, L2 → E+SECO, L3 → E+SECO ... || L4 → E+SECO]]

- (103) a. [la lana] Pues la lavamos y después ya se
 well ACC.3SG wash.PRS.1PL and afterwards already REFL.3
 pone a secar al sol.
 put.PRS.3SG to dry.INF at.the sun
 'Well, we wash [the wool] and afterwards we put it to dry in the sun.'
 (Lucillo, León, COSER 2644)
- b. La grasa lo poníamos a derretir, a calentar, a punto de
 the fat ACC.3SG put.PST.1PL to melt.INF to heat.INF to point of
 hervir y lo echábamos sobre la sosa.
 boil.INF and ACC.3SG throw.PST.1PL over the soda
 'We put the fat to melt, to heat up, until it almost boiled, and we poured it
 over the caustic soda.' (Cigales, Valladolid, COSER 4407)

Second, the arguably synonymic structure "*poner* or a similar verb + *a que* + inflected verb in the subjunctive" (see (104)) shows the opposite marking pattern. That is, anticausatives in this construction are massively marked – 35 out of 42 examples (83.3%).⁸ The different marking pattern between the two constructions is rather extreme, as Table 27 shows, despite being close synonyms. Note that since no examples of the inflected verb construction were found in north-western varieties, this table only shows data from the rest of the territory.

- (104) a. Y se ponía a que se secara ahí.
 and REFL.3 put.PST.3SG to that REFL.3 dry.PST.SBJV.3SG there
 'And it was put there so it would dry there.'
 (Povedilla, Albacete, COSER 222)
- b. Luego eso lo dejaban a que se enfriara,
 later that ACC.3SG leave.PST.3PL to that REFL.3 cool.PST.SBJV.3SG
 porque entonces no había frigorífico.
 because then no have.PST.3SG fridge
 'Afterwards they left it to cool, because there were no fridges then.'
 (Arjona, Jaén, COSER 2301)

8. This alternation of infinitive and subjunctive *que* clauses appears in several other domains of the grammar of Spanish (NGLE 2009: 26.11).

Table 27. Frequency of the RM in the “*poner* (or a similar verb) *a* + infinitive”/“*tener* + gerund” construction, depending on the inflection of the verb in the COSER data (rest of the territory)

	+ non-finite clause (infinitive/ gerund)	+ finite <i>que</i> clause (subjunctive/ conditional)*
Unmarked (RM)	141 (100%)	7 (16.7)
Marked (No RM)	0 (0%)	35 (83.3%)
Total	141 (100%)	42 (100%)

Fisher’s exact test odds ratio estimate (1000 resamples):** Min.: Inf, 1st Qu.: Inf, Median: Inf, Mean: Inf, 3rd Qu.: Inf, Max.: Inf

* Some varieties of peninsular Spanish are known to use the conditional in contexts where other varieties typically use the imperfect subjunctive (see Pato 2003, 2012 for a more detailed account).

** Because the observations of our total sample are not independent (i.e. many locations offer more than one example) and a generalised linear mixed model is not reliable (due to low figures in some cells), I have resampled the data 1,000 times (hierarchical bootstrap) and applied a Fisher’s exact test to each of them. The measures given here correspond to the distribution of odds ratio estimates over these 1,000 samples. Values greater than one indicate a higher probability of positive cases (i.e. either unmarked non-finite verbs or marked finite verbs) than of negative cases (i.e. either marked non-finite verbs or unmarked finite verbs).

Moreover, in the similar structure with the preposition *para* instead of *a*, the presence of the RM is possible, as shown in (105). That is, the rejection of the RM in this structure is both related to the use of the infinitive and the use of the preposition *a* with a final value in the construction discussed by Vivanco (2016).

- (105) a. Y el año pasao colgué uno para secarse y no
and the year last hang.PST.1SG one for dry.INF-REFL.3 and no
me gusta, porque se queda muy seco.
DAT.1SG please.PRS.3SG because REFL.3 become.PRS.3SG very dry
‘And last year I hung one up for it to dry and I don’t like that, because it
gets very dry.’ (Quintana de los Prados, Burgos, COSER 939)
- b. Pimientos, pero pimientos secos, de esos que se cuelgan
Peppers but peppers dry.PL of those that REFL.3 hang.PRS.3PL
a secar, no los verdes
to dry.INF no the green.PL
‘Peppers, but dried peppers, those that are hung up for drying, not the
green ones.’ (Navalmoral de la Mata, Cáceres, COSER 1015)

Third, other structures that focus on the process and not on the final result state (such as the progressive periphrasis “*estar/ir* + gerund”) do not reject the clitic (see (106a, b)). The same can be said for structures that focus on the beginning of the process (see (106c)). These three examples show activity readings with marked durative change-of-state verbs:

- (106) a. Lo dejas que se fría mucho, a lo mejor
 ACC.3SG leave.PRS.2SG that REFL.3 fry.PRS.3SG much to the better
se está friendo dos o tres horas, y se pone
 REFL.3 be.PRS.3SG frying two or three hours and REFL.3 put.PRS.3SG
 muy tostaíto.
 very toasted
 ‘You let it fry well, it might be frying two or three hours and it gets very
 toasted.’ (Manzanares el Real, Madrid, COSER 2910)
- b. [Las ovejas] cuando se van haciendo mayores, las
 when REFL.3 go.PRS.3PL making older ACC.3PL
 van sacrificando
 go.PRS.3PL slaughtering
 ‘When [the sheep] are getting older, they start slaughtering them.’
 (Etxauri, Navarra, COSER 3209)
- c. Cuando el pan empezaba a abrirse a la media hora o así
 when the bread start.PST.3SG to open-REFL.3 to the half hour or so
 ‘When the bread started opening up, half an hour later or so.’
 (Castrillo de Valdelomar, Cantabria, COSER 1205)

Finally, the absence of the RM is also possible in aspectual periphrases that select a durative verb and focus on reaching the final state, such as “*tardar en* + infinitive” (‘to take time to + infinitive’, see (106a)) or “*terminar de* + infinitive” (‘to finish + gerund’, see (106b)). This contradicts Vivanco’s (2016: 234) claim that unmarked anticausatives can have both activity and achievement readings, whereas they reject accomplishment readings, which must take the RM.

- (106) a. El jamón se tie- | tarda mucho en secar
 the ham REFL.3 have.PRS.3SG take.long.PRS.3SG much in dry.INF
 también.
 also
 ‘And ham is..., it also takes a long time to dry’
 (Menagaray, Álava, COSER 109)
- b. [el vino] primero se fermentaba en el lagar y así de
 first REFL.3 ferment.PST.3SG in the press and so of
 terminar de fermentar...
 finish.INF of ferment.INF
 ‘[Wine] was first fermented in the press and as soon as it finished ferment-
 ing...’ (Banastón, Huesca, COSER 2206)

To sum up, my data quantitatively support the claim that the RM is rejected in the construction “*poner... + a* + infinitive”, but data from synonymous structures and other aspectual periphrasis suggest that this is not necessarily related to the aspectual

event structure of the predicate. This specific structure shows atypical behaviour, acting like a fixed expression. This can be seen by comparing all instances of anticausative verbs in the infinitive or gerund. I have grouped them in five different classes: (1) instances in unequivocally atelic periphrases (such as “*estar/ir* + gerund” ‘to be V+ing’, “*empezar a* + infinitive” ‘to begin to V’); (2) instances in unequivocally telic periphrases and constructions (such as “*costar* + infinitive” ‘to be difficult that V’, “*después de* + infinitive” ‘after V+ing’, “*llegar a* + infinitive” ‘completely V’, “*parar/terminar de* + infinitive” ‘stop V+ing’, “*ponerse en punto de* + infinitive” ‘to be about to V’, “*sin* + infinitive” ‘without V+ing’, “*tardar en* + infinitive” ‘to take long to V’, “*volver a* + infinitive” ‘to V again’); (3) the “*poner... + a* + infinitive” or the “*tener* + gerund” construction, (4) contexts very similar to the “*poner... + a* + infinitive” construction, but with preposition *para* or with the *a* clause in a coordinated (not subordinated) sentence to the *poner*-clause;⁹ and (5) instances in contexts that are not aspectually restricted.¹⁰ The frequency of the RM in these different constructions (see Figure 17) shows very clearly the special status of the “*poner... + a* + infinitive” construction. It is the only context that categorically rejects the presence of the RM and it differs strongly from other contexts, be they telic or atelic, where the presence of the RM is almost always the norm. Focussing on this construction leads Vivanco (2016) to claim that the distribution of the RM responds categorically to aspectual contrasts, but Figure 17 shows that this is not so. From this, it is clear that this test cannot be used to argue for a distribution of the RM with anticausatives based on aspectual differences, since the “*poner... + a* + infinitive” structure is a fixed construction, one of whose requirements is to take unmarked anticausatives.

Is the RM’s distribution affected probabilistically by the verbal aspect? A number of authors claim that this is indeed the case, based on introspective judgements. However, the data in Figure 17, which show rather similar frequencies of the RM for telic, atelic and not aspectually restricted periphrases, do not support these intuitions. The fact that many authors have argued for a role of verbal aspect in the distribution of the RM is nevertheless relevant evidence, since it is clear that native speakers perceive some aspectual differences between at least some marked and unmarked examples. Anticausatives (and conversive) verbs as a class seem to be associated, at least prototypically, with some event types (such as changes of state), which are also associated with some aspectual features, such as telicity. However,

9. An example of an infinitive in a coordinated *a* + infinitive clause:

(i) Y se colgaban y a secar.
and REFL.3 hang.PST.3PL and to dry.INF
‘And they were hung so that they would dry.’ (Mahide, Zamora, COSER 4617)

10. I have excluded causative periphrases, since there were very few of them.

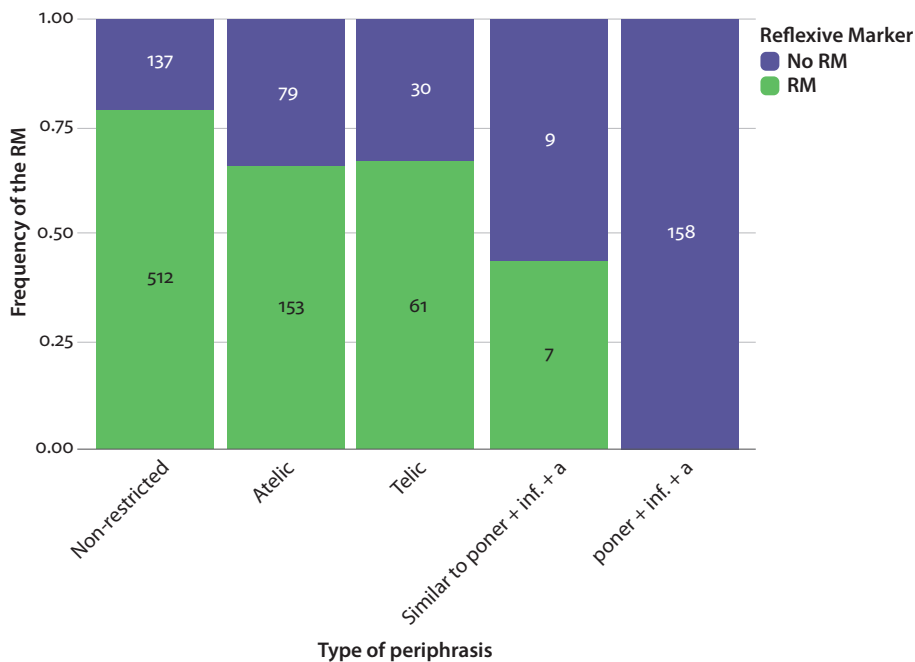


Figure 17. Frequency of the RM in anticausative infinitives or gerunds by the aspectual characteristics of the periphrasis (rest of the territory)
All verbs, regardless of their aspectual characteristics, are included in this count.

reducing the distribution of the RM to a single aspectual characteristic is probably too simplistic. Again, proposing tendencies that apply to only a subset of the data (as Kailuweit (2012) does) seems a more fruitful approach.

4.5 The event structure of the predicate: The acceptability of datives

Among the syntactic differences that Vivanco (2016) notes between marked and labile anticausatives, she mentions that only the former accept datives with an unintentional cause or an affected reading. Because she provides the most detailed discussion of this topic, I follow her argumentation closely in this section. In her words, “there are no affected datives in the absence of *se*, unless affectedness can be inferred from a possessive relationship” (Vivanco 2016: 277, my translation). For Vivanco (2016), the distribution of unintentional cause and affected datives goes hand in hand with the distribution of the RM because both categories are connected to the same event structure, that is, they all require a dynamic subevent

followed by a final result state.¹¹ When there is no result state (i.e. in labile anticausatives), affected datives cannot appear because they are located between those two subevents and unintentional causers are not felicitous because “the semantics of causation requires an end result state, it requires that the event culminates in a state” (Vivanco 2016: 319, my translation).

More specifically, regarding class C achievements, Vivanco claims that the presence of the RM triggers a contrast in the duration of the event: unmarked achievements are not durative at all, while the presence of the RM adds some duration to the event. This would explain the incompatibility of dative causers or affected datives with unmarked achievements – since their achievement subevent is simultaneous to their final result state, these datives have no room left in the syntactic structure. Therefore, she claims that the dative in example (107a) cannot be interpreted either as an unintentional causer or as an affected participant, and that it can only a possessor interpretation. The presence of the RM, however, creates a contrast between both subevents, leaving room for affected datives, as can be seen in (107b).

(108) (Vivanco 2016)

- a. A Yuridia le ha despertado el niño
to Yuridia DAT.3SG has wake.up.PTCP the child
‘The child of Yuridia woke up.’ [✓ possessor, *causer, *affected]
- b. A Romeo se le despertó Julieta
to Romeo REFL.3 DAT.3SG wake.up.PST.3SG Juliet
‘Juliet woke up [to Romeo].’ [*possessor, ✓ affected]

Nonetheless, the COSER interviews attest one example of an unmarked achievement with an affected dative. The fact that there is only one example must not be seen as proof that these are rare, but as a consequence of the fact that there seem to be very few class C achievements. Actually, Vivanco (2016) only identifies two verbs in this class: *despertar* ‘to wake up’ and *reventar* ‘to burst, to blow out’. The single COSER example shows the unmarked *reventar* with an affected dative (see (109a)). As example (109b) shows, this combination is also possible in Standard Spanish. That is, the contrast identified by Vivanco (2016) is not borne out by at least one of the two verbs in this class.¹²

11. The different event structures deemed to be associated to marked and unmarked anticausatives have been also discussed in detail in Vivanco (2021).

12. I find this contrast hard to accept even introspectively. On the one hand, it is often very hard to keep the possessive and the affected readings apart, since affectedness often comes from (sometimes temporal) possession. On the other hand, example (108a) is hard to interpret for me and seems to be something I would never say in my own idiolect.

- (109) a. Tú echabas el líquido ahí, y lo dejabas
 you throw.PST.2SG the liquid there and ACC.3SG leave.PST.2SG
 abierto, y luego te reventaba todo aquello, y to
 open and afterwards DAT.2SG explode.PST.3SG all that and all
 la mierda que tenía, echaba.
 the shit that have.PST.3SG throw.PST.3SG
 ‘You poured the liquid in there and left it open and afterwards it will burst
 [on you] and all the shit it had will come out.’
 (Aulesti, Vizcaya, COSER 4501)
- b. A los once o doce años le reventó un barreno en
 to the eleven or twelve years DAT.3SG explode.PST.3SG a drill in
 la mano.
 the hand
 ‘When he was eleven or twelve a drill exploded in his hand.’
 (Camilo José Cela, 1988, *Cristo versus Arizona*, España, CREA)

Regarding deadjectival class C verbs, which can have both an atelic process reading and a telic achievement reading, Vivanco (2016) gives a different explanation. As discussed in Section 4.4, she considers that the unmarked version does not focalise the final result state. Dative unintentional causers are then rejected because a final result state is a semantic requirement for causation, according to this author. Affected datives are acceptable as long as they are also possessive datives, as in (110a), where the subject of the anticausative is an inalienable possession of the dative.

However, I have argued above, I do not believe that this aspectual contrast between marked and unmarked anticausatives is so clear-cut. It seems to me that the uneven distribution of dative causers with these can be explained in simpler terms. As also stated above (Section 4.3), anticausatives depict events that can be produced by a cause, since it is a requirement for a verb to enter into anticausative alternation to be able to have a cause subject. However, dative causers do not encode just any kind of cause, but specifically human ones (Schäfer 2008; Fábregas 2021). Anticausatives that cannot have a human cause cannot take dative causers and that is why Luisa in (110a) cannot be interpreted as the cause of the event. However, (110b, c) show that dative causers are possible both with marked and unmarked class C deadjectival verbs, as long as they are semantically interpretable.

- (110) a. A Luisa le ha envejecido la piel
 to Luisa DAT.3SG have.3SG age.PTCP the skin
 ‘Luisa’s skin has aged.’ [√ possessor-affected, *causer] (Vivanco 2016)
- b. [el guiso] Hay que ir vigilando para que no se
 have.PRS.3SG that go.INF watching for that no REFL.3
 nos queme y se nos ennegrezca
 DAT.1PL burn.SBJV.3SG and REFL.3 DAT.1PL blacken.SBJV.3SG
 ‘One has to watch [the stew] so that it does not get burnt [to us] and blacken
 [to us].’ (eseuftenTen11, gastronomía.laverdad.es)

- c. ¿Puedo prepararles algún estimulador casero
 can.PRS.1SG prepare.INF-DAT.3PL one stimulating.agent homemade
 para que me engorden los cogollos?
 for that DAT.1SG fatten.SBJV.3PL the cos.lettuce
 'Can I prepare some homemade stimulating agent so the my cos lettuce
 grows larger?' (eseuTenTen11, cannabismagazine.es)

Let us briefly review the COSER quantitative data regarding the datives that appear with anticausative verbs. There were only 419 examples of anticausatives with a dative participant in the COSER interviews, from which only 36 were found in north-western varieties. The first thing to be noted is that classifying real speech datives into affected, causer or possessive is not easy, since these categories are often intertwined. It is often the case that causers are also possessors (see (110b, c)), and it also might be the case that both affectedness, possession and (unintentional) causality can all be interpreted in a given example (see (111)). I classified the datives attested in the COSER data into four different types: experiencer datives (typically with verbs such as *olvidar* 'to forget' or *pesar* 'to feel sorrow', see (112a)), goal datives (with movement predicates, such as *meterse* 'to get in', see (112b)), possessive datives (where the dative plays no role in the initiation of the event but has a possessive relationship with the subject of the event, see (112c)), and causer/affected datives (where the dative either played a role in the initiation of the event or its only relationship with the event is that it was affected by it). It must be noted that dative causers also tend to have some kind of possessive relationship with the subject (if only a temporary one). In the literature, it is common to focus on unintentional causers, such as those in (112d) (Schäfer 2008; Vivanco 2016). However, there are many cases where the dative is interpreted as an "intentional" causer, in the sense that the participant carries out some preparatory actions to trigger the change of state (compare the clearly unintentional dative (112d) to (112e)). I see no reason to differentiate between them, since it seems that the importance of unintentional causers in both Schäfer's (2008) and Vivanco's (2016) argumentation lies in the fact that they are causers, but not in their lack of intentionality. Finally, it is worth pointing out that all these datives (even goal datives) are somehow affected. That is, establishing a clear-cut distinction between possessive datives and causer/affected datives is almost impossible when we look at naturally occurring data.

- (111) ¿Ha dejao la [NP] algo puesto ahí en la
 have.PRS.3SG leave.PTCP the [PN] something put.PTCP there in the
 lumbre que te se queme?
 fire that DAT.2SG REFL.3 burn.PRS.3SG
 'Did the [PN] leave something over there on the fire which can get burn?'
 (Tablones, Granada, COSER 1834)

- (112) a. Oy, me se olvidan las cosas.
oops DAT.1SG REFL.3 forget.PRS.3PL the things
'Oops I keep forgetting stuff.' (Almadén de la Plata, Sevilla, COSER 3806)
- b. No se le metía el aceite por dentro.
not REFL.3 DAT.3SG put.in.PST.3SG the oil for inside
'The oil wouldn't get inside it.' (Bacares, Almería, COSER 404)
- c. No se me curaba esa herida.
not REFL.3 DAT.1SG heal.PST.3SG that injury
'That injury that I had wouldn't heal.'
(Aguaviva de la Vega, Soria, COSER 3923)
- d. Se nos ha roto el molinillo.
REFL.3 DAT.1PL have.PRS.3SG break.PTCP the grinder
'Our grinder broke.' (Valencia del Ventoso, Badajoz, COSER 728)
- e. Le echo una poquita sosa y ya me
DAT.3SG throw.PRS.1SG a few caustic.soda and already DAT.1SG
se cuaja.
REFL.3 set.PRS.3SG
'I add a little bit of caustic soda and then in sets [to me].'
(Espera, Cádiz, COSER 1107)

At any rate, the COSER interviews show a few clear examples of involuntary dative causers with unmarked predicates, located in Galicia, as illustrated in (113). In what follows, I investigate whether the presence of a dative has an effect on the presence of the RM in both varieties under study.

- (113) [el molino] Ahora está medio abandonado, porque... me
now be.PRS.3SG half abandon.PTCP because DAT.1SG
rompió abajo.
break.PST.3SG down
'It is now half abandoned, because it broke at the bottom.'
(Zas, La Coruña, COSER 2403)

Figure 18 shows the distribution of the presence of the RM by dative type in our two regions of interest. Due to the extremely low figures in north-western varieties, it is hard to assess the effect of this parameter on the distribution of the RM. In the rest of the territory, however, it seems that datives of all types very frequently co-occur with the RM. This is not surprising, since in this area anticausatives are very frequently marked with the RM.

However, the comparison of anticausatives that take a dative and those that do not (Figure 19) clearly shows that the presence of the dative has an effect on the presence of the RM in the rest of the territory, since it is considerably more likely that an anticausative verb with a dative is marked (95.6%) than when there is no dative (79.3%). In north-western varieties no effect is detected, but we should bear in mind that there are only a few examples with datives.

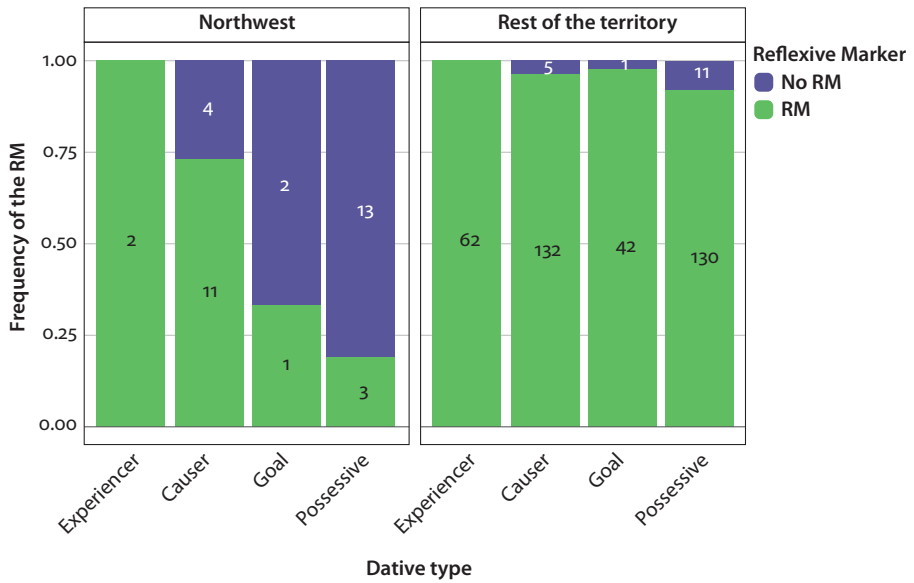


Figure 18. Frequency of the RM by kind of dative

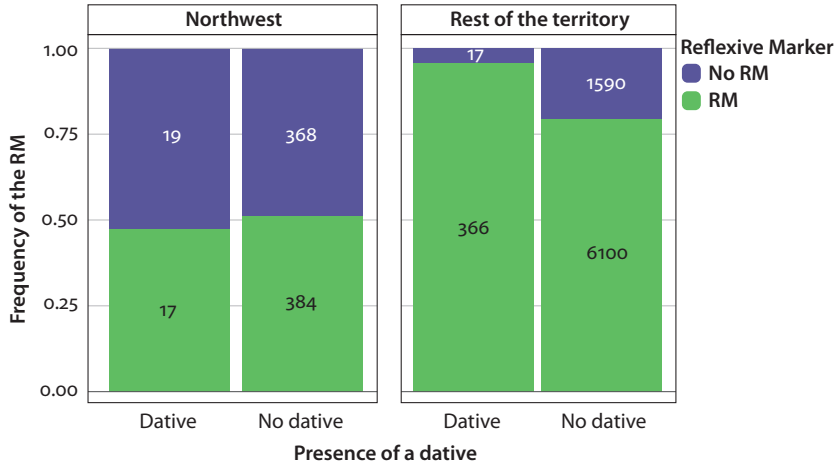


Figure 19. Frequency of the RM by presence of a dative participant

In order to reduce the dependency of the data, in Figure 20 I plot the correlation between the frequency of the RM and the frequency of a dative *by verb*. No correlation is apparent in north-western varieties – again, note that only a few verbs take datives at all. In the rest of the territory no linear relationship can be found, although there is a trend in which verbs which take datives more often seem to take the RM frequently, while the reverse is not necessarily true (i.e. many verbs that rarely take datives are also take the RM frequently).

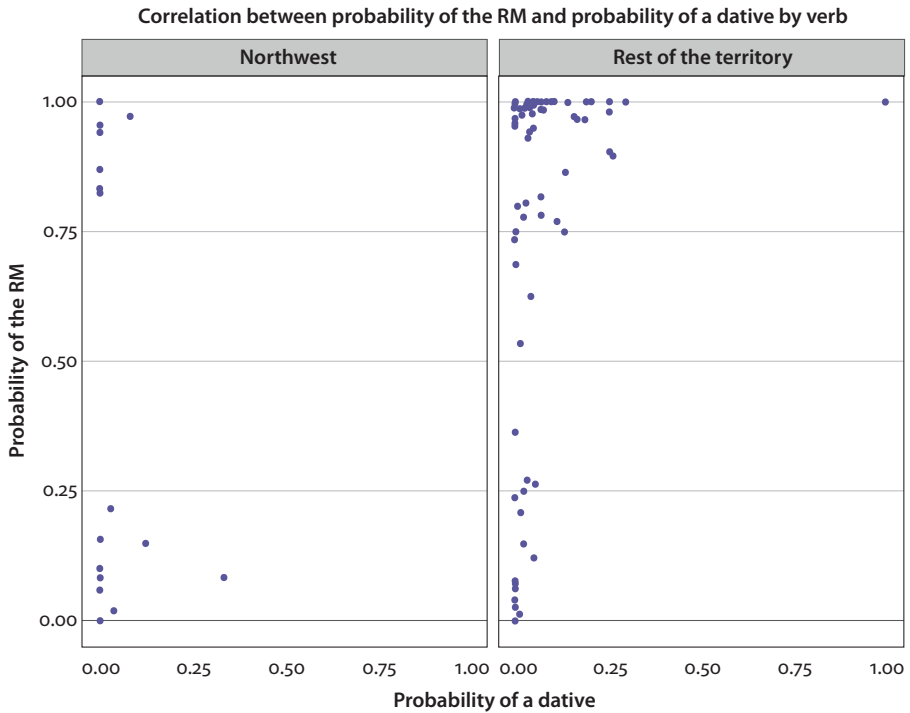


Figure 20. Correlation between the probability of the RM and the probability a dative by verb and geographical area

To sum up, it indeed seems that the presence of an affected dative (since they all seem to be affected) is a significant factor in the presence of the RM. However, I do not believe this to be an unequivocal indicator of the aspectual function of the clitic, since it is clearly related to the semantics of anticausative constructions. As noted above (Section 4.3), some authors have proposed that anticausative verbs are more prone to be marked when they are typically interpreted as having an external cause. The behaviour of the datives might be seen as indirect proof of this relationship between the RM and the presence of an external cause, since datives often encode such an external cause. The assimilation of other datives to dative causers

is understandable, since, we have said, dative causers are normally understood as being affected by the event and also typically show a possession relationship with the verb subject. These semantic connections provide a potential source of analogy that could have triggered a higher frequency of marking if a dative is present.

4.6 Animacy of the subject

I now turn to the effect of the animacy of the subject in the marking patterns of anticausative verbs. Monge (1955) points out that the most important landmark in the evolution of the reflexive pronoun is the possibility of taking inanimate subjects. This possibility was very rare in Early Latin but had already become frequent in Late Latin. Obviously, the original reflexive function can only appear with animate subjects. The extension to inanimate subjects, which has been claimed to have a strong element of personification in the early stages (Richter 1909; Monge 1955), would have allowed the reflexive pronoun to become an intransitivising device, in what Monge (1955: 14) describes as “the grammaticalization process of a meaningful element” (my translation). Accordingly, “the use of the pronominal form with middle and intransitive value is explained by the weakening of the reflexive meaning when the subject is animate and by the effect of analogy and personification if the subject is inanimate” (Monge 1955: 16, my translation). A very similar argument is put forward by Waltereit (2000): “A corollary of the grammaticalization process is that reflexive markers can combine with non-human subjects, because affectedness – as the new, grammaticalized meaning of the reflexive marker – is not restricted to humans” (Waltereit 2000: 266). Nevertheless, studies with a synchronic orientation have not focussed specifically on the animacy of the subject as a possible factor that affects the productivity of the RM, although several authors have made a distinction in their typologies of reflexive constructions depending on the animacy of the subject, as discussed in Section 1.2.3.

As we know, in grammaticalization theory “the source meaning uniquely determines the grammaticization path that the gram will travel in its semantic development” (Bybee, Perkins & Pagliuca 1994: 12). Accordingly, a natural hypothesis is that the presence of the RM in anticausative verbs might be influenced by the animacy of the subject, being higher with animate subjects than with inanimate ones. Moreover, the expectation is that such an effect is greater in north-western varieties, where the lower productivity of the RM suggests that this marker is at an earlier stage of development.

At this point, the strong association that other constructions with a change of diathesis showed with animate subjects in the COSER interviews is worth noting: (1) only ten out of 839 (1.2%) examples of de-objective verbs show inanimate

subjects, (2) only two out of 455 (0.4%) examples of conversive verbs show inanimate subjects, (3) only 11 out of 325 (3.4%) examples of antipassive verbs show inanimate subjects. These data hint at a strong association between the presence of RM and animate subjects across its valency-changing function.

In the COSER data, the animacy of the subject seems to affect the presence of the RM to a modest extent in the rest of the territory and to a greater extent in north-western varieties. As Figure 21 shows, animate subjects are more favourable to the presence of the RM than inanimate subjects in both areas.

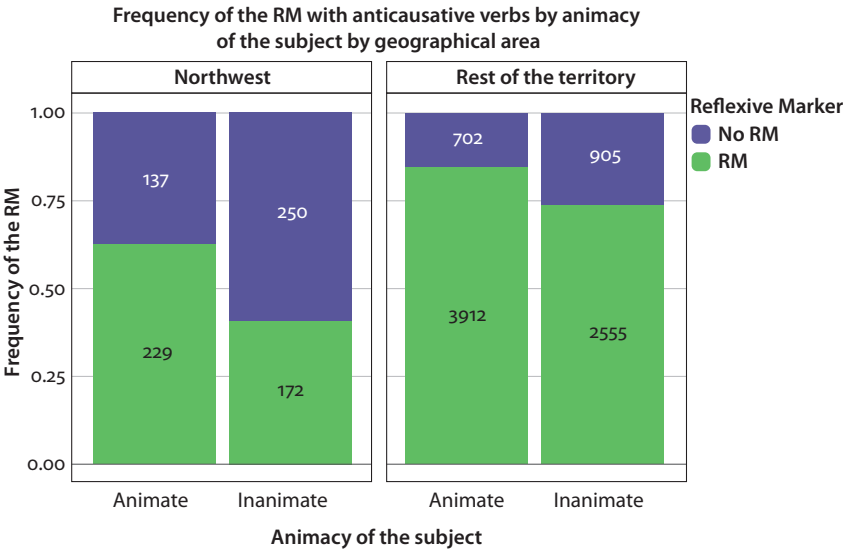


Figure 21. Frequency of the RM with anticausative verbs by animacy of the subject in the rest of the study area

In order to reduce the dependency of the data, the correlation between presence of the RM and probability of the subject being animate *by verb* and geographical area (only verbs with ten or more occurrences are considered here) can be checked. Interestingly, no correlation is apparent from Figure 22. However, this does not necessarily mean that the relationship between animacy of the subject and the presence of the RM does not exist, only that it is not verb-dependent.

This interpretation finds some support if we aggregate first the examples by animacy of the subject and, afterwards, by the verb. In this case we cannot study the correlation between two probabilities, but we can observe the distribution of the four resulting groups, as shown in Figure 23. Again, only verbs with ten or more occurrences are considered, which now reduces our sample considerably:

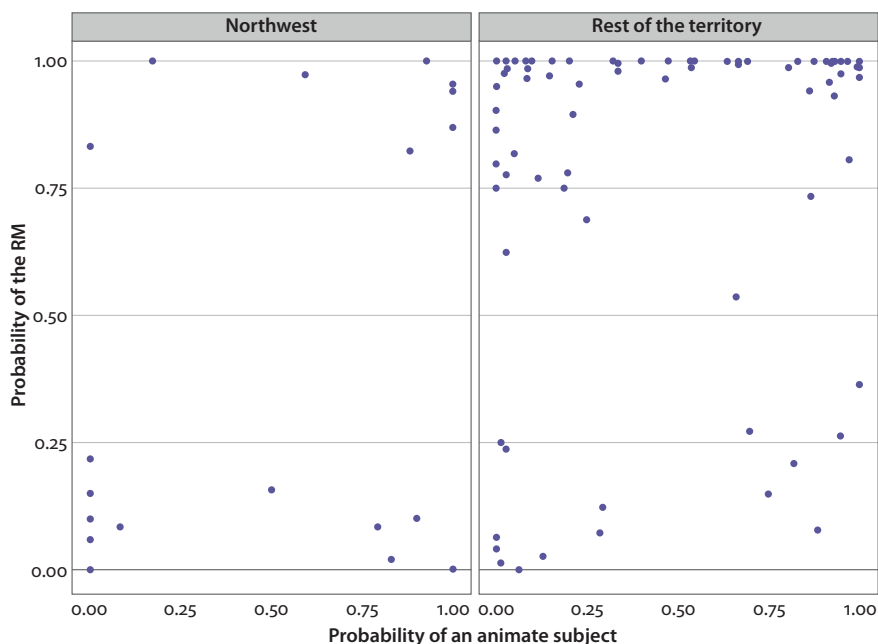


Figure 22. Correlation between the probability of the taking RM and animate subjects by verb and geographical area

12 verbs were found in each group (Animate/Inanimate) in the northwest, and 58 and 50 were found in the rest of the territory. In north-western varieties, Figure 23 shows an extreme difference between the medians in both groups, as well as a more cohesive marking pattern in anticausatives with inanimate subjects. The opposite is true for the rest of the territory. This is consistent with anticausatives with animate subjects being older than anticausatives with inanimate subjects, and with north-western varieties being at an earlier stage of the development than the rest of the territory: while in the rest of the territory marking of anticausative verbs with animate examples is almost complete (with some outliers), and verbs with inanimate subjects are catching up, in the northwest anticausative verbs with animate subjects have not reached complete productivity and anticausative verbs with inanimate subjects tend to strongly resist the presence of the RM.

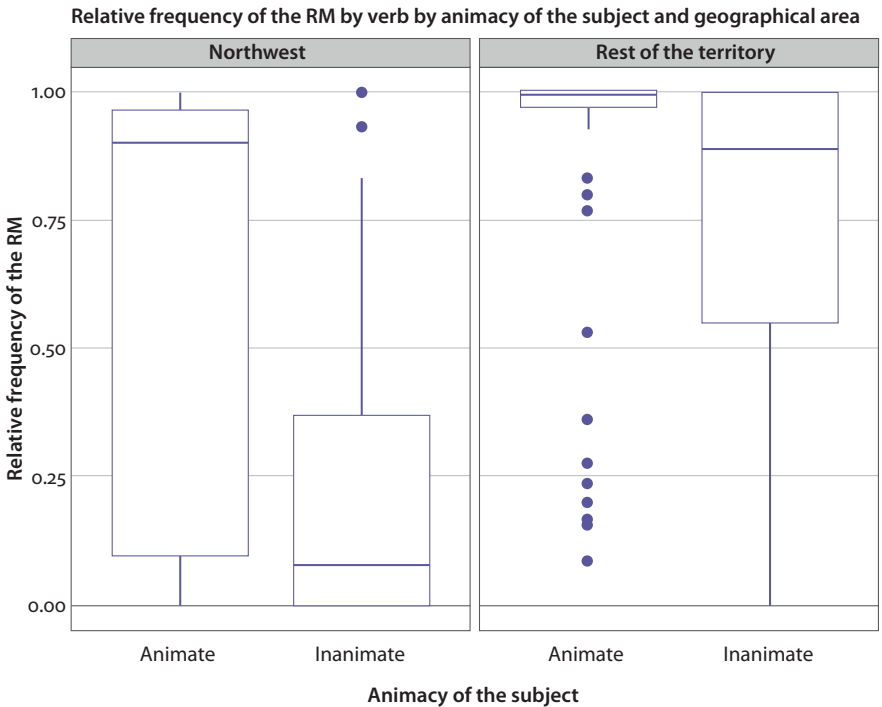


Figure 23. Boxplots for the relative frequency of the RM with anticausative verbs by animacy of the subject and area in the COSER data

Animate subjects (Northwest): Min. = 0.0000; 1st Qu. = 0.09649; Median = 0.90145; Mean = 0.59980; 3rd Qu. = 0.96591; Max. = 1.00000
 Inanimate subjects (Northwest): Min. = 0.0000; 1st Qu. = 0.00000; Median = 0.07941; Mean = 0.27431; 3rd Qu. = 0.37050; Max. = 1.00000
 Animate subjects (Rest of the territory): Min.: 0.08696, 1st Qu.: 0.97012, Median: 1.00000, Mean: 0.87032, 3rd Qu.: 1.00000, Max.:1.00000
 Inanimate subjects (Rest of the territory): Min.: 0.0000, 1st Qu.: 0.5521, Median: 0.8864, Mean: 0.7040, 3rd Qu.: 0.9991, Max.: 1.0000

The effect of the animacy of the subject, of course, is not an absolute determiner of the presence of the RM, and there are a number of verbs that even seem to show the opposite contrast. Vivanco notes that Spanish anticausative verbs allow for both animate and inanimate subjects regardless of their marking, but also offers the contrast in (114), which seemingly contradicts the tendency that animate subjects prefer marked anticausatives. In these examples, however, the animacy of the subject collides with the likelihood of an external cause, since *envejecer la madera* ‘to make wood older’ can easily be conceived of as a causative event (see (115)), while *envejecer a Luisa* ‘to make Luisa older’ is only acceptable in a fantasy world.

(114) (Vivanco 2016)

- a. La madera se ha envejecido con los años
 the wood REFL.3 have.PRS.3SG age.PTCP with the years
 'The wood has aged with the years.'
- b. *Luisa se ha envejecido con los años
 Luisa REFL.3 have.PRS.3SG age.PTCP with the years
 'Luisa has aged with the years.'

(115) Para envejecer la madera y darle un toque más oscuro
 for age.INF the wood and give.INF-DAT.3 a touch more dark
 se suele quemar superficialmente con un soplete
 REFL.3 tends.3SG burn.INF superficially with a blowtorch
 'In order to make the wood look older and give it a darker look, one normally
 burns it superficially with a blowtorch.'
 (esTenTen, <http://www.agrada.es/ayuda/envejecer-mueble-nuevo>)

4.7 Generalised linear mixed model

So far, I have analysed each factor independently. However, a joint quantitative analysis that takes all relevant parameters into account can be useful as a means of comparing the effect size of each parameter and to generate a more fine-grained analysis where we control for each factor. In what follows, only the 83 verbs that were analysed in Section 4.3 (i.e. those that were examined in corpus *eseuTenTen11*) are considered. Moreover, all examples that appear in any kind of periphrases or construction analysed in Section 4.4 are disregarded, since there are only a few of them. Table 28 shows the results of a generalised linear model that predicts the presence of the RM depending on three categorical fixed factors (geographical

Table 28. Generalised linear mixed model. Presence of the RM depending on geographical area [reference level: Rest of the territory], subject animacy [reference level: Inanimate], presence of a dative [reference level: No dative], probability of transitive uses in corpus *eseuTenTen11* (fixed factors) and the place (random factor)

Effect	Term	Estimate	Std. error	Statistic	Conf. low	Conf. high
fixed	(Intercept)	-0.478	0.125	-3.816	-0.724	-0.233
fixed	Area: Northwest	-1.433	0.25	-5.739	-1.922	-0.943
fixed	Subject animacy: Animate	0.96	0.093	10.378	0.779	1.141
fixed	Presence of a dative: Yes	0.933	0.248	3.768	0.448	1.418
fixed	Probability of being transitive	3.816	0.184	20.762	3.456	4.176
random	sd (Intercept)	0.701	NA	NA	NA	NA

(COSERID)

area, subject animacy, and presence of a dative) and one quantitative fixed factor (probability of transitive uses in corpus *eseuTenTen11*), plus one random factor (the specific town of each interview).

The model allows us to better understand the parameters that affect the presence of the RM in anticausative verbs. Interestingly, the factor that favours the presence of the RM most strongly is the probability that the verb appears in transitive contexts in the *eseuTenTen11* data. This seems surprising, since within the *eseuTenTen11* data there was no clear correlation between the probability of an event being encoded in a transitive construction and the probability of the same verb being marked in an intransitive construction. Controlling for the effect of other variables, however, it seems that the distinction between events with an external cause (those that appear more frequently in transitive contexts) and events with an internal cause is relevant to understanding the presence of the RM. Second, the geographical variety is the second strongest variable that affects the presence of the RM – north-western varieties disfavour the presence of the RM, with an effect that is about three times smaller than the probability of appearing in a transitive context. Finally, both the presence of the dative and the animacy of the subject have a similar effect on the presence of the RM, which is about four times weaker than the probability of appearing in a transitive context.

4.8 Conclusions

In this chapter I have analysed more closely the behaviour of the RM within the anticausative clause. Specifically, I have looked at the internal composition of the class in terms of the presence of the RM, and analysed quantitatively the effect of five factors in the marking patterns of anticausatives, namely, the distinction between externally and internally caused events, the aspectual properties of the predicate (as indicated by the presence of a verbal periphrasis), the presence of a dative, the animacy of the subject, and the geographical variety.

First, my data clearly show that some verbs are more resistant to the RM than others, as suggested in other studies. The fact that even the most resistant verbs show at least sporadic marking in the rest of the territory suggests that the pressure from the rest of the class is rather strong.¹³

13. Enrique Arias & Bouzouita (2013) noted that the anticausative construction was “completely developed” (my translation) as early as the thirteenth century, since it was already more frequent than the direct reflexive by that time. However, a comparison of their data (the productivity of the anticausative construction with the RM increases considerably from that period to the modern language) with the data presented in this chapter suggests that this was not completely the case. An analysis of their data in the terms proposed here would probably be very enlightening.

Second, the frequency of causative uses of these verbs, used to operationalise the likelihood of an external cause, seems to be a strong predictor of the presence of the RM when controlling for other factors. That is, while my close approximation of Vivanco's (2016) study with a larger sample did not show a clear correlation (Section 4.3), this parameter has the strongest effect in the generalised linear mixed model (Section 4.7).

Third, my data empirically support the claim that the construction "*poner... + a + infinitive*" rejects marked anticausatives, as proposed by Vivanco (2016). However, I have argued against this being proof of the link between a specific event structure and the presence of the RM, since the comparison with constructions with similar aspectual characteristics suggests that the construction "*poner... + a + infinitive*" is a fixed construction that shows a peculiarly strong resistance to the presence of the RM.

Fourth, I have shown that the presence of any kind of dative favours the presence of the RM, although this effect seems to be stronger in the rest of the territory than in north-western varieties. I have argued that this is not necessarily related to the event structure of the predicate, as claimed by Vivanco (2016), but that it might have to do with the fact that the dative often encodes an external cause and external causes favour the presence of the RM.

Fifth, as regards the animacy of the subject, whose role has been hinted at in several studies (Fernández Ramírez 1986; Cartagena 1972; Martín Zorraquino 1979; Cano 1981; Benavides 2010), I have argued that the quantitative evidence supports the hypothesis of this factor being relevant to the development of the RM. The effect of the animacy of the subject can confirm and update Monge's claim that the spread of the RM to constructions with inanimate subjects is a late development. In many of the contexts in which the RM appears (conversive, de-objective, anti-passive, transitive verbs and, of course, purely reflexive constructions), inanimate subjects are rather rare, which means that the RM is strongly associated *de facto* with animate subjects. In categories where inanimate subjects occur frequently, like anticausatives, the RM is more frequent with animate subjects, confirming that this association had an impact on the evolution of the RM. The COSER data reflect very clearly the fact that anticausatives with inanimate subjects are a more recent "conquest" of the RM.

Finally, the dialectal perspective also offers some rather interesting insights. On the one hand, north-western varieties disfavour the presence of the RM in general. On the other hand, not all factors seem to have the same effect in all varieties. This has implications for the diachronic reconstruction of the history of the RM. The COSER data show effects of the animacy of the subject in the two areas under study. This makes sense if this parameter is so strongly determined by the source construction. The data on datives, however, suggest that this parameter is not relevant

in the northwest, as opposed to the relevance of its effect in the rest of the territory. In turn, this suggests that the association between the presence of a dative and the RM came into play later, and indeed is not (yet?) found in the northwest.

Reflexive intransitive verbs

5.1 Introduction

In this chapter I focus on intransitive reflexive verbs that show no change of valency. As described in Section 1.2.2, there is broad consensus in the current literature on attributing an aspectual function to the RM in this category. I have previously argued against the idea of the RM having an aspectual function when it attaches to verbs that show no change of valency (De Benito Moreno 2015; De Benito Moreno 2021). On these lines, I have shown that the sharp aspectual contrasts cited in the literature are in fact not found when an exhaustively analysis of corpus data is undertaken. On the contrary, such analysis actually shows that the distribution of the RM in these verbs is both much broader and less systematic.

In this chapter I offer an alternative account of the role of the RM in intransitive verbs, based on the diachronically-oriented hypothesis that the attachment of the RM to intransitive verbs results from the analogical extension of the RM from older functions, namely, its reflexive, reciprocal and change-of-valency functions. For such a hypothesis I follow Fischer's (2008) and De Smet's (2012) use of analogy (see also Section 1.5.2). My claim that these uses of the RM are analogical, however, does not entail that this particle is meaningless.¹ On the contrary, I believe that the meaning of the RM in these constructions is crucially determined by its older functions, which determine the contexts in which it is most likely to appear. That is, the systematic attachment of the RM to reflexive, reciprocal and diathesis reducing constructions has caused the association of the RM with some meanings (having to do with the semantics of the verb, the animacy of the subject, the existence of an affected external cause or an involved participant, see Chapter 4) that are pervasive with these constructions, and which in itself has caused the extension of the RM to contexts with similar meanings but where no change of diathesis or operation takes place.

In this chapter I will support this hypothesis with a quantitative analysis of the data on the association between the RM with intransitive verbs and such meanings. It is important to note, however, that an account of the RM in these terms does

1. Maldonado (1989) rejects any interpretation of the RM on analogical terms, considering that it entails that the RM would have no meaning.

not predict that the RM would have one exact and well-defined function. On the contrary, it explains its broad and apparently unsystematic distribution. Sections 5.2 and 5.3 focus on the analogical effect of the semantic class of the verb and the animacy of the subject, respectively. Sections 5.4 to 5.9 analyse in more depth the behaviour of some frequent reflexive intransitive verbs, in order to show how the specific meaning of the marked variants can be related to previous values of the RM.

5.2 Analogical effects of the semantic class of the verb

In total, 69 different intransitive verbs exhibit the RM at least once in the COSER data.² In the following section, I will consider the semantic classes of these verbs, in order to see whether the extension of the RM to these verbs, which cannot be attributed to a syntactic function of the RM (since verb valency is not altered), can be analysed in terms of analogy with the semantic classes considered in Chapter 3.

Some 49 of these verbs (that is, 71%) fit at least one of the semantic categories considered by Kemmer (1993). Figure 24 provides an overview of the distribution of these verbs across semantic categories, showing the notable significance of translational motion and spontaneous events, which account for 19 (27.5%) and 14 (20.3%) of all the verbs respectively. These were two of the main semantic classes with which reflexive verbs with a change of diathesis are associated (see Section 3.4), suggesting that this association (i.e. analogy) could effectively have played a role in the extension of the RM to those verbs. Another 23.1% are either naturally reciprocal events (5 verbs), body processes (4), emotion events (4), non-translational motion (2) or grooming verbs (1). Only 29% of the verbs (20 in total) are not easily classifiable in any of these categories. This situation is reminiscent of Mous's judgement on non-reversible verbs in Cushitic languages: "The [middle] deponents fall into the same subsets of meaning that Kemmer has set up in her typology of middle markers. Apparently verbs with a lexical meaning in the middle semantic domain tend to vacuously add a middle affix" (Mous 2001: 77). Whether this affixation is really vacuous, or to what extent it is so, will be the topic of the remainder of this book.

The intransitive verbs that show the RM at least once are provided in Table 29, classified according to their semantic category. Of course, it should be noted that the fact that some of these verbs do not easily fit in Kemmer's semantic classes does

2. In fact, a total of only 65 are documented, but a few of them are polysemic and are thus classified in several semantic categories (see Table 29). Note that, as mentioned in Section 4.2.1, *subir* 'to go up', *bajar* 'to go down', *pasar* 'to pass' and *pasear* 'to go for a walk' are also considered here, despite showing a transitive-intransitive alternation.

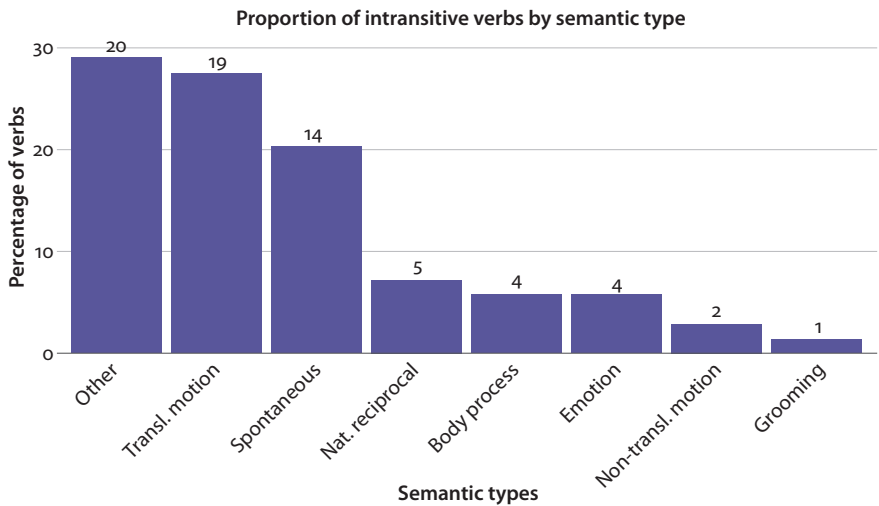


Figure 24. Semantic categories of reflexive intransitive verbs in the COSER data

not rule out analogy as a reason for the RM to appear. Analogy may be triggered by a few specific verbs with related meanings (for example, *vivir* is not a spontaneous change of state, but is related to *morir*, which is a spontaneous change of state that may take the RM). Specific analogical processes like that (which may be behind the presence of the RM in every case, not only in those belonging to the “Other” semantic class) are, however, much harder to trace and would demand a more fine-grained study. A brief note on these is given in Section 5.9. At any rate, the quantitative data in this section and its comparison with the data in Section 3.4 are clear evidence of the role of analogy (through paradigmatic semantic similarity) triggered by verbal semantics in the presence of the RM with intransitive verbs.

Table 29. Reflexive intransitive verbs by semantic category in the COSER data

Semantic type	Verbs
Other	<i>abusar</i> ‘to abuse’, <i>cenar</i> ‘to have dinner’, <i>dejar</i> ‘to leave’, <i>desertar</i> ‘to desert’, <i>esperar</i> ‘to wait’, <i>estar</i> ‘to be’, <i>existir</i> ‘to exist’, <i>jugar</i> ‘to play’, <i>mocear</i> ‘to be young’, <i>pasar</i> ‘to happen’, <i>pesar</i> ‘to weigh’, <i>quedar</i> ‘to stay, to remain’, <i>reposar</i> ‘to rest’, <i>retumbar</i> ‘to resound’, <i>seguir</i> ‘to continue’, <i>servir</i> ‘to be useful’, <i>tardar</i> ‘to take time’, <i>trabajar</i> ‘to work’, <i>valer</i> ‘to be worth’, <i>vivir</i> ‘to live’
Translational motion	<i>acudir</i> ‘to go’, <i>andar</i> ‘to walk’, <i>bajar</i> ‘to go down’, <i>correr</i> ‘to run’, <i>descender</i> ‘to go down’, <i>entrar</i> ‘to go in’, <i>escapar</i> ‘to escape’, <i>inmigrar</i> ‘to immigrate’, <i>ir</i> ‘to go’, <i>llegar</i> ‘to arrive’, <i>marchar</i> ‘to leave’, <i>montar</i> ‘to climb in’, <i>pasar</i> ‘to pass’, <i>pasear</i> ‘to walk’, <i>salir</i> ‘to go out’, <i>subir</i> ‘to go up’, <i>venir</i> ‘to come’, <i>volar</i> ‘to fly’, <i>volver</i> ‘to go back’

(continued)

Table 29. (continued)

Semantic type	Verbs
Spontaneous	<i>arder</i> ‘to burn’, <i>caer</i> ‘to fall down’, (<i>en</i>) <i>canecer</i> ‘to rot’, <i>crecer</i> ‘to grow’, <i>dormir</i> ‘to sleep’, <i>esbarar</i> ‘to slip’, <i>escullar</i> ‘to drain’, <i>morir</i> ‘to die’, <i>mostear</i> ‘to turn into grape juice/must’, <i>nacer</i> ‘to be born’, <i>pasar</i> ‘to happen’, <i>resbalar</i> ‘to slip’, <i>salir</i> ‘to go out’
Naturally reciprocal	<i>conformar</i> ‘to be in agreement’, <i>discutir</i> ‘to discuss’, <i>luchar</i> ‘to fight’, <i>pelear</i> ‘to fight’, <i>reñir</i> ‘to quarrel’
Body processes	<i>cagar</i> ‘to have a shit’, <i>mear</i> ‘to piss’, <i>orinar</i> ‘to urinate’, <i>pisar</i> ‘to urinate’
Emotion	<i>gustar</i> ‘to like’, <i>reír</i> ‘to laugh’, <i>sonreír</i> ‘to smile’, <i>desear</i> ‘to wish’
Non-translational motion	<i>brincar</i> ‘to jump’, <i>saltar</i> ‘to jump’
Grooming	<i>ir</i> ‘to wear’

A very strong argument in favour of an analogical origin of the extension of the RM to intransitive verbs comes from marked intransitive naturally reciprocal verbs. Five such verbs are attested in my data: *conformar*, *discutir*, *luchar*, *pelear*, *reñir*. All these verbs are activities and hence cannot be easily subsumed under the aspectual functions that have been proposed for the RM, that typically connect the presence of the RM with telicity (Zagona 1996; De Miguel & Fernández Lagunilla 2000; Sanz & Laka 2002, among others). However, their semantic relationship with one of the most conspicuous functions of the RM, namely, marking reciprocal constructions, is beyond doubt and suggests that the RM attached to these intransitive verbs is analogically related to reciprocal constructions.

The frequency of the RM with these verbs is far from the systematic appearance of the RM in reciprocal constructions (see De Benito Moreno 2015). From 81 examples of naturally reciprocal verbs attested in the COSER interviews, only 36 (44.4%) examples were marked.

It is also worth noting that the frequency of markedness is verb-specific to a very high degree. This is shown in Figure 25 below, where the presence of the RM is depicted together with the frequency of each verb. The two most frequent verbs, which have very similar meanings, show almost a mirror-like picture regarding markedness: *discutir* is almost always unmarked, while *pelear* clearly prefers the presence of the RM. *Reñir*, also a synonym, looks more similar to *discutir*, while *luchar* is somewhere in the middle (but note that there were only a small number of examples overall for these verbs, even fewer in the case of *conformar*).

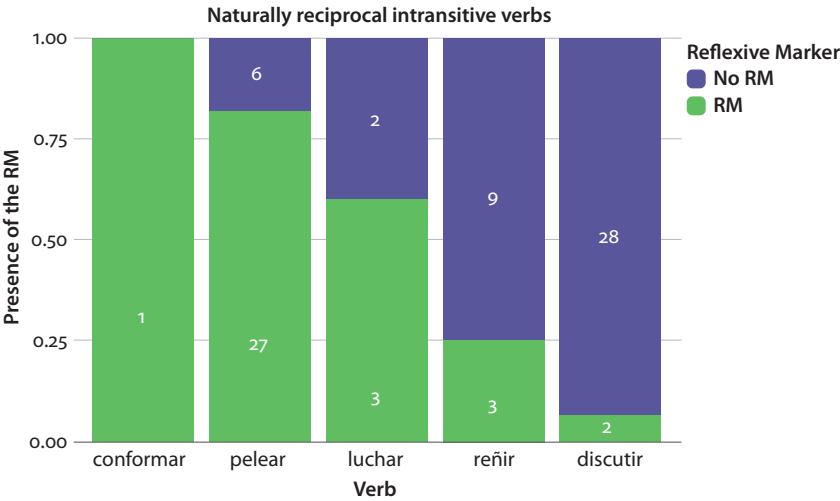
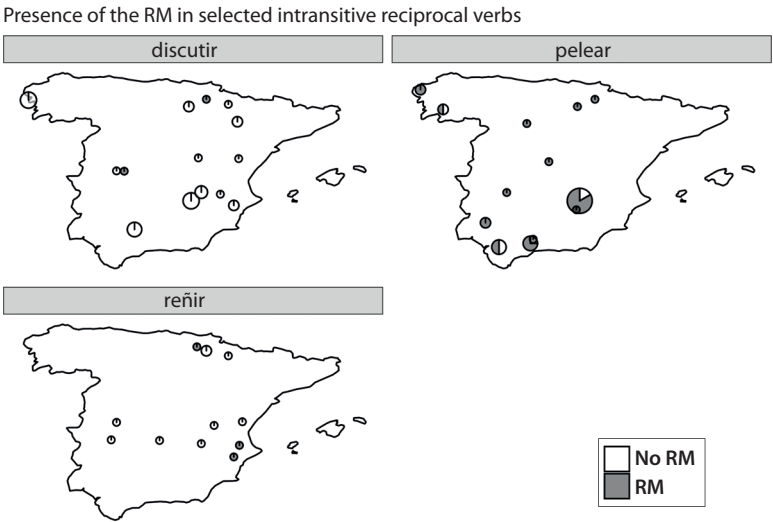


Figure 25. Frequency of the RM with symmetric intransitive verbs in the COSER data

The scarcity of data makes it difficult to know whether the presence of the RM in these cases is spatially distributed, although Map 10 suggests that it is not. For the three most frequent and variable intransitive reciprocal verbs (that is, *discutirse*, *reñirse* and *pelear*) it seems that the infrequent variant appears sporadically, following no clear geographical pattern.



Map 10. Frequency of the RM in symmetric intransitive reciprocals in the COSER data

5.3 Analogical effects of the animacy of the subject

The data in this section (and the remainder of the chapter) are restricted to the 25 intransitive verbs that I collected exhaustively, which were also considered in Section 3.2.6, with the addition of *subir(se)*, *bajar(se)* and *pasar(se)*, which were collected from the whole corpus.³ In Chapters 3 and 4 these verbs were considered anticausatives, but due to their abnormal behaviour as such they will be considered here as well as intransitives (as is typically done in the literature).

The animacy of the subject was one of the factors found to have an effect on the frequency of the RM with anticausative verbs (see Chapter 4). Regarding intransitive verbs, Figure 26 shows that the animacy of the subject seems to have no effect in the northwest, where intransitive verbs are marked extremely rarely, regardless of the characteristics of the subject. In the rest of the territory, however, intransitive verbs show more reflexive marking when the subject is animate (see Figure 26). It is also important to note that animate subjects are far more frequent than inanimate ones in these verbs – they represent 77.4% of the subjects (1,668 out of 2,156) in the northwest and 73.8% in the rest of the territory (9,769 out of 13,233). This seems to indicate that the RM tends to be attached to intransitive verbs that normally take animate subjects.⁴

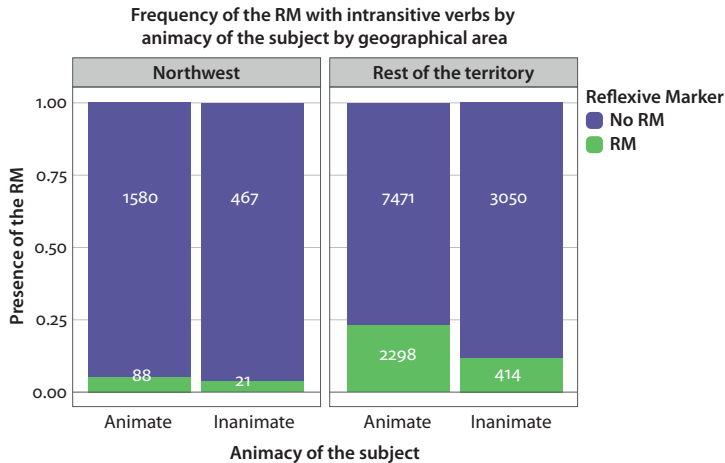


Figure 26. Frequency of the RM with intransitive verbs by animacy of the subject in the northwest

3. Metaphorical uses of all these verbs were excluded.

4. Variationist accounts of intransitive (and transitive) reflexive verbs have also taken animacy into account (Aaron 2003, 2004; Aaron & Torres Caucullos 2005; Torres Caucullos & Schwenter 2008; Rivas 2011; De la Mora 2011). Overall, our data on the association between the RM and animate subjects in middle constructions contradict García’s (1975: 4) claim that it is “particularly frequent with inanimate subjects”

If we plot each verb with ten or more occurrences as a dot placed on the x-axis according to the frequency with which it takes the RM, we clearly see that, in the rest of the territory, verbs with animate subjects show a higher degree of reflexive markedness (see Figure 27). The reduced sample in the north-western area allows us to see that intransitive verbs are very rarely marked in this area, with a few exceptions. Comparing this figure with Figure 28, which shows the same information for anticausative verbs, we see the same pattern for these verbs (as already discussed in Section 4.6) – more verbs with animate subjects show more frequent marking than verbs with inanimate subjects. A clear difference is also apparent, however: marking is far more frequent with anticausative verbs than with intransitive ones. This fits the hypothesis that the marking of intransitive verbs is a later development caused

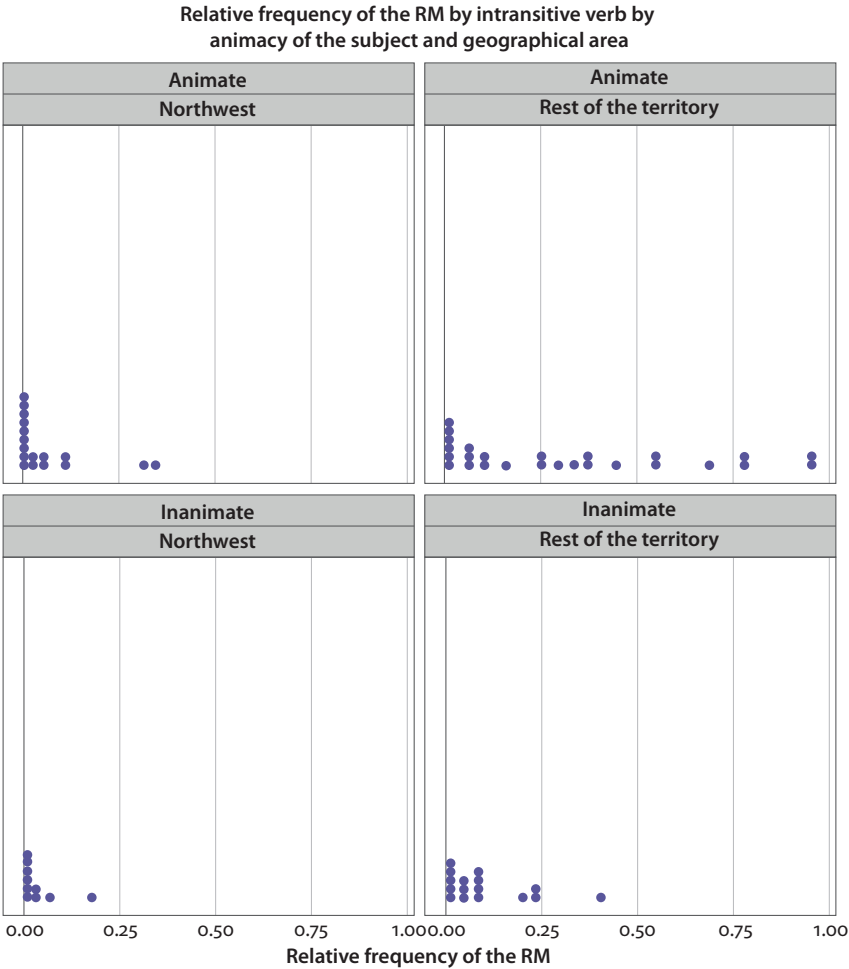


Figure 27. Dotplots for the relative frequency of the RM with intransitive verbs by animacy of the subject and area in the COSER data

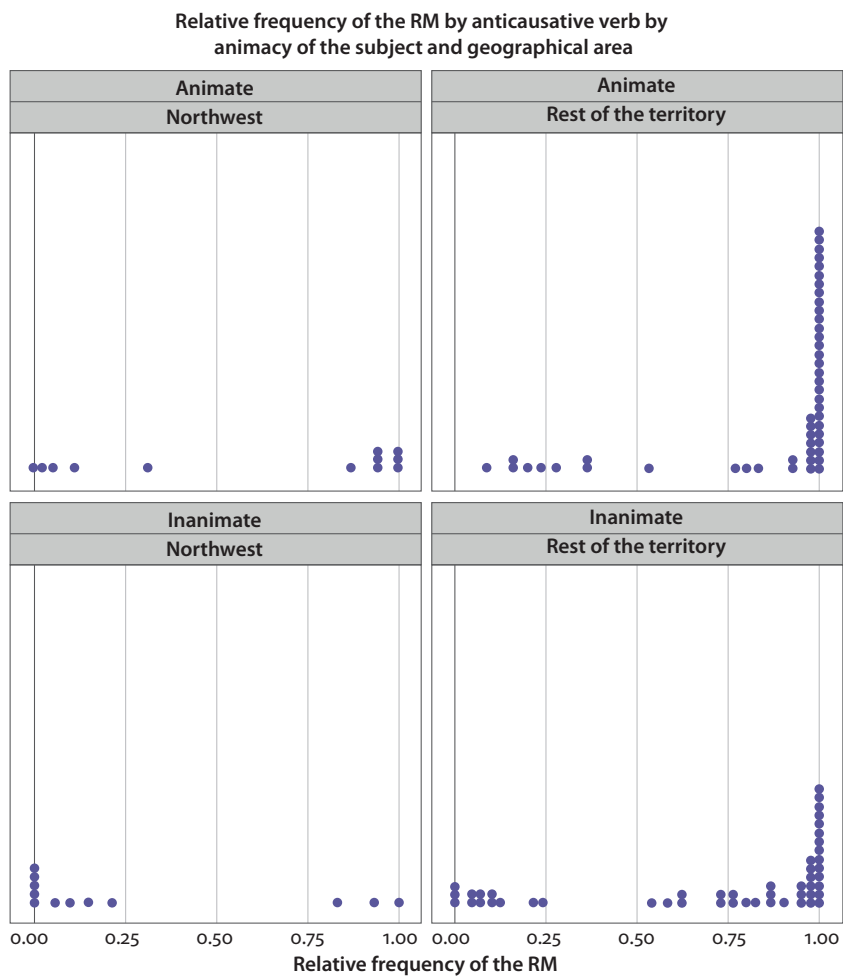


Figure 28. Dotplots for the relative frequency of the RM with anticausative verbs by animacy of the subject and area in the COSER data

by analogy with verbs with a change of diathesis. That is, the similarities between both classes regarding the effect of the animacy of the subject are seen here as the product of syntagmatic analogy, while the differences in the frequency of the RM are due to the fact that one construction is older than the other.

5.4 Intransitive reflexive verbs of translational motion

In this section I will deal with some of the most frequent intransitive movement verbs found in my data, namely, *ir(se)* ‘to go’, *marchar(se)* ‘to leave’, *salir(se)* ‘to go out’, *escapar(se)* ‘to escape’, *venir(se)* ‘to come’, *subir(se)* ‘to go up’, *bajar(se)* ‘to go down’, *montar(se)* ‘to climb in’ and *pasar(se)* ‘to pass’.⁵ Many of these have been considered in studies on the so-called aspectual *se* (De Miguel & Fernández Lagunilla 2000; Sánchez López 2002), where contrasts such as (116) and (117) have been used in support of the aspectual function of the RM, which would be related to the focalisation of a result state (De Miguel & Fernández Lagunilla 2000). Besides the aspectual contrast, the verbs in these examples show a difference in meaning. Sánchez López has noted that marked movement verbs might show “a sense of abandonment of the place of origin” (Sánchez López 2002: 118, my translation). Three of these verbs, *salir(se)*, *subir(se)*, *bajar(se)*, have also been studied in detail from a variationist perspective in Mexican Spanish, providing a useful point of comparison (Aaron 2003, 2004; Aaron & Torres Cacoullós 2005; Torres Cacoullós & Schwenter 2008).

- (116) *(Me) voy de aquí para no volver.
 REFL.1SG go.PRS.1SG from here for not go.back.INF
 ‘I’m leaving not to come back.’ (Sánchez López 2002: 118)

- (117) (Sánchez López 2002: 120)
- a. Se subió a un árbol y allí estuvo subido
 REFL.3 climb.PST.3SG to a tree and there be.PST.3SG climb.PTCP
 más de dos días.
 more of two days
 ‘He climbed a tree and was up there more than two days.’
 - b. ?Subió a una montaña y allí estuvo subido
 climb.PST.3SG to a mountain and there be.PST.3SG climb.PTCP
 más de dos días.
 more of two days
 ‘He climbed a mountain and was up there more than two days.’

In the following paragraphs I will briefly explore the semantic differences that can be found between the marked and unmarked variants of these verbs. First, however, it is worth noting that these verbs show very different proportions of marked variants, as seen in Figure 29.⁶

5. Because there are few metaphorical uses of these verbs, I have excluded them from this section.

6. A similarly uneven distribution is found by Aaron (2003) for Mexican Spanish. These contrasts were already found in old Spanish: Martín Zorraquino (1998) found that *venirse* ‘to come’ was far less frequent than *irse* ‘to go/leave’.

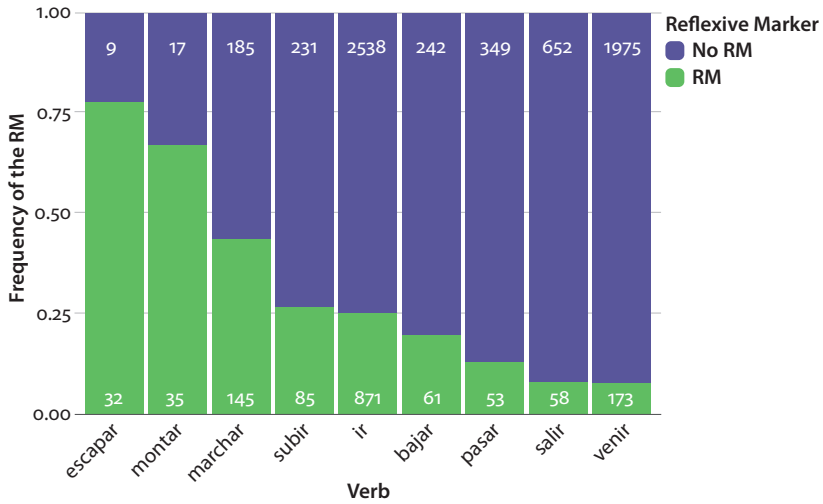


Figure 29. Frequency of the RM with frequent translational motion intransitive verbs⁷

Figure 29 shows that verbs with more specific meanings (such as *marchar(se)* ‘to leave’, *escapar(se)* ‘to escape, to get away’ and *montar(se)* ‘to get into, to climb to’) show far higher frequencies of the RM than verbs with broader meanings (such as *ir(se)* ‘to go’, *salir(se)* ‘to get out’ and *subir(se)* ‘to go up, to climb’). This is especially interesting given that the reflexive forms of the latter (i.e. *irse*, *salirse* and *subirse*) seem to be associated with meanings similar to *marchar(se)*, *escapar(se)* and *montar(se)*. This has often been noted in the literature for a verb like *irse*, which is typically translated as ‘to leave’, as opposed to *ir* ‘to go’, but can also be argued for *salirse* and *subirse*.

For instance, a close examination of the data shows that *salirse* ‘to get out’ tends to appear in contexts where the subject is exiting a place against a physical or social obstacle (see also Maldonado (1999) and Aaron (2003) for a similar but broader description), a meaning clearly related to *escapar(se)* ‘to escape, to get away’. Many of these examples refer to someone leaving the army, school, political service, a religious order, or even normality (see (118)).⁸ Most examples of *salir(se)* in such

7. In my data from rural speech, *salirse* is rather more frequent (9%) than what Aaron & Torres Cacoulios (2005) found for educated speech in Spain (3%, 13/486) and much closer to what they found for Mexican Spanish (12%, 66/557).

8. Aaron & Torres Cacoulios (2005) briefly analyse the use of *salirse* in Peninsular Spanish and note that the presence of the RM is associated with examples meaning “[to exit] from a physically delimited space or against a physical obstacle” (Aaron & Torres Cacoulios 2005: 622), but they do not find examples of abstract delimited spaces or obstacles. The COSER data clearly show that such an abstract meaning is also possible (and common) in Peninsular Spanish. Nevertheless,

contexts show reflexive marking (34/50, 68%), while marked examples in other contexts are extremely rare (24/660, 3.6%).

- (118) Se salió de cura y se fue a Madrí.
 REFL.3 go.out.PST.3SG of priest and REFL.3 go.PST.3SG to Madrid
 'He ceased being a priest and went to Madrid.'
 (Alcalá de la Selva, Teruel, COSER 4102)

Regarding *subir(se)* 'to go up, to climb', the marked form *subirse* is preferred when the action depicted indicates getting/climbing into something (sometimes with a collateral change of posture), as opposed to instances simply depicting going somewhere that is higher (but on the same continuous path as the source of the movement), where *subir* is preferred (see (119)). The marked form is found in 79.5% (58/73) of the total cases where *subir(se)* appears in "climbing" contexts, while *subir(se)* in other contexts is seldom marked: 11.1% (27/243). Interestingly, *bajar(se)*, which depicts the same kind of movement but in the opposite direction, shows a similar pattern: from 16 examples of the "climbing down" interpretation, 68.8% (11 cases, although note the low total figures) show the RM, as opposed to the 17.4% (50/287) of marking found in other contexts. This parallels Torres Cacoullos & Schwenter's (2008) findings that the most common meaning of *subirse* and *bajarse* is 'to go into / out of a vehicle'. That is, the meaning typically associated with *subirse* is close to the meaning of *montar(se)* 'to get into, to climb to'.

- (119) ¿Pa qué van a subirse al árbol?
 for what go.PRS.3PL to climb.INF-REFL.3 to.the tree
 'What are they going to climb a tree for?' (Luzuriaga, Álava, COSER 0107)

Of course, these figures show that the association between the RM and these specific meanings or contexts is not perfect or systematic, but rather a statistical tendency. The same happens with *ir(se)* 'to go'. The marked variant *irse* is rather frequent with goal complements (26%, 314/1209 examples of *ir(se)* with goal complements with a 'to' or *para* 'for'). In these cases, no "sense of abandonment" is conveyed and it is difficult to argue that the specific phase of leaving a place is focussed on, as shown in (120). That is, while the RM is required (no unmarked examples were found) when *ir(se)* takes the meaning 'to leave', *irse* very often shows other meanings. This same variability is found in other dialects, as shown by the analyses of *salir(se)*, *subir(se)* and *bajar(se)* in Mexican Spanish (Aaron 2004; Aaron & Torres Cacoullos 2005; Torres Cacoullos & Schwenter 2008).

the other three meanings of *salirse* that they find in Mexican Spanish, namely 'leaving abruptly', 'leaving permanently' and 'leaving in order to have a good time' (Aaron & Torres Cacoullos 2005: 610), seem to be rare in Peninsular Spanish.

- (120) [en las bodas] Pero ya no, ya la gente se va
 but already no already the people REFL.3 GO.PRS.3SG
 al restaurán.
 to.the restaurant
 [at weddings] 'But not now, now people go to a restaurant and that's all.'
 (Talaván, Cáceres, COSER 1020)

This lack of sharp semantic contrasts between marked and unmarked variants argues against a clearly defined syntactic function of the RM. The data show, nevertheless, that the RM attaches more frequently to some contexts than others. Following my hypothesis that the distribution of the RM is driven by paradigmatic and syntagmatic analogy with older functions of the RM, I argue in what follows that the contexts where the RM appears more frequently are connected to the behaviour of this marker in its older functions.

Regarding the characteristics of the subject, my data show that, for some of these verbs, the presence of the RM is affected by the animacy of the subject. Recall that this parameter was also found to be relevant to the productivity of the RM with anticausatives (see Section 4.6). For instance, as a translational motion intransitive, *montar(se)* only admits animate subjects. This suggests that the type of movement depicted by *montar(se)* (and preferred by *subirse*, as noted above) is more agentive than the type of movement preferred by *subir*, which admits both animate and inanimate subjects.⁹ It is no surprise, then, that the presence of the RM is higher for both *subir* and *bajar* (which depicts the same movement in the opposite direction) when they have animate subjects. This is shown in Figure 30: while the role of animacy is clear for the rest of the territory, there are very few data for north-western varieties, preventing us from drawing any conclusions.

This fits well with the proposals of a number of authors that the presence of the RM is related to the role of the subject and not to the aspectual characteristics of the predicate (Maldonado 1988; Di Tullio 2012). These authors have mentioned the notion of agentivity or volitionality as crucial in understanding the role of the RM with these verbs. Of course, animate subjects are more typically agentive than inanimate agents. The association of the RM with animacy, due to its original reflexive meaning (see Section 4.6 above), could have produced a further association between the RM and agentivity.

Concerning verbs related to *marchar(se)* (i.e. *ir(se)* and *venir(se)*, the latter depicting the same movement but with opposite deixis), my data actually suggest

9. Actually, Maldonado (1999) considers that in *Adrián se subió a la mesa* (as opposed to *Adrián subió la escalera*) there is an energy increase, which implies more speed and dynamism.

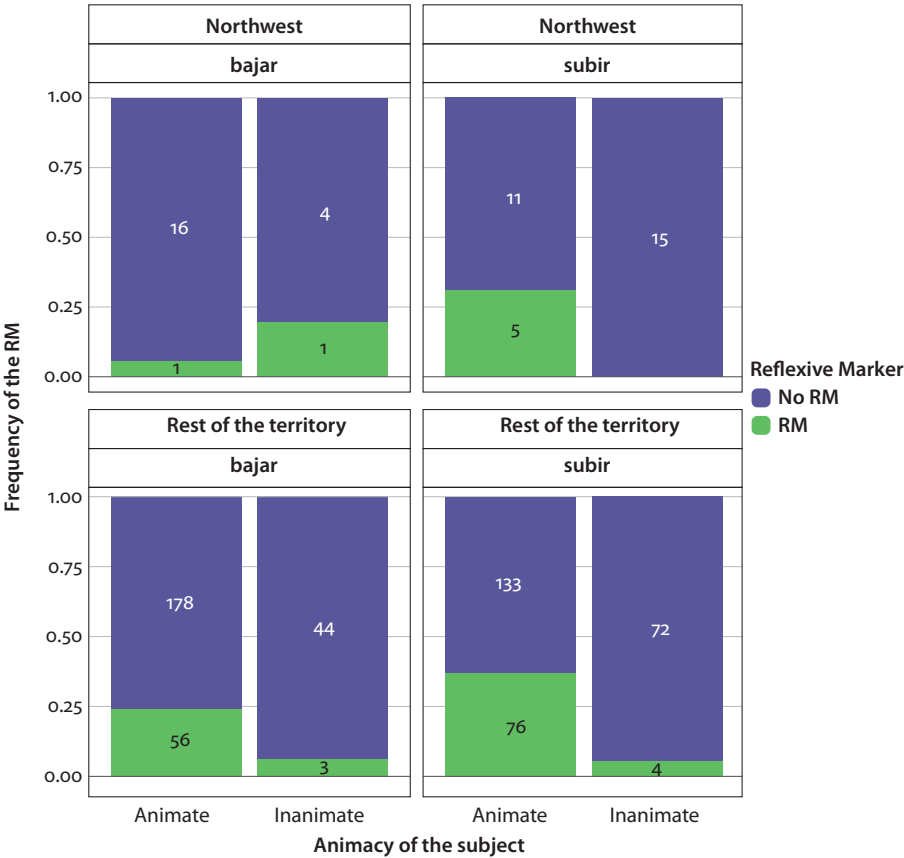


Figure 30. Frequency of the RM by animacy of the subject with *bajar(se)* and *subir(se)* by geographical area

that it is not exactly animacy, but agentivity which is playing a role in the presence of the RM. On the one hand, animacy of the subject seems to be a relevant factor for *marchar(se)* and *venir(se)*, but not for *ir(se)*, as shown in Figure 31. Again, note that the verb that most frequently takes the RM, *marchar(se)*, strongly prefers animate subjects overall. On the other hand, agentivity seems to be a relevant factor for *ir(se)* and *venir(se)* (the data are insufficient for *marchar(se)*), as shown in Figure 32. This figure depicts the frequency of the RM depending on whether the verb is in the imperative or not (and includes only examples with animate subjects, which are the only ones that can appear in the imperative). Imperatives are used as a form of grammatical proof to identify agentive verbs and are thus

unequivocally agentive.¹⁰ Again, the association of the RM and agentivity could have been produced as a side effect of its association with animacy, but might also be the product of a direct analogy with autocausative verbs, that is, anticausative verbs with animate and agentive subjects (see Section 3.2.1).

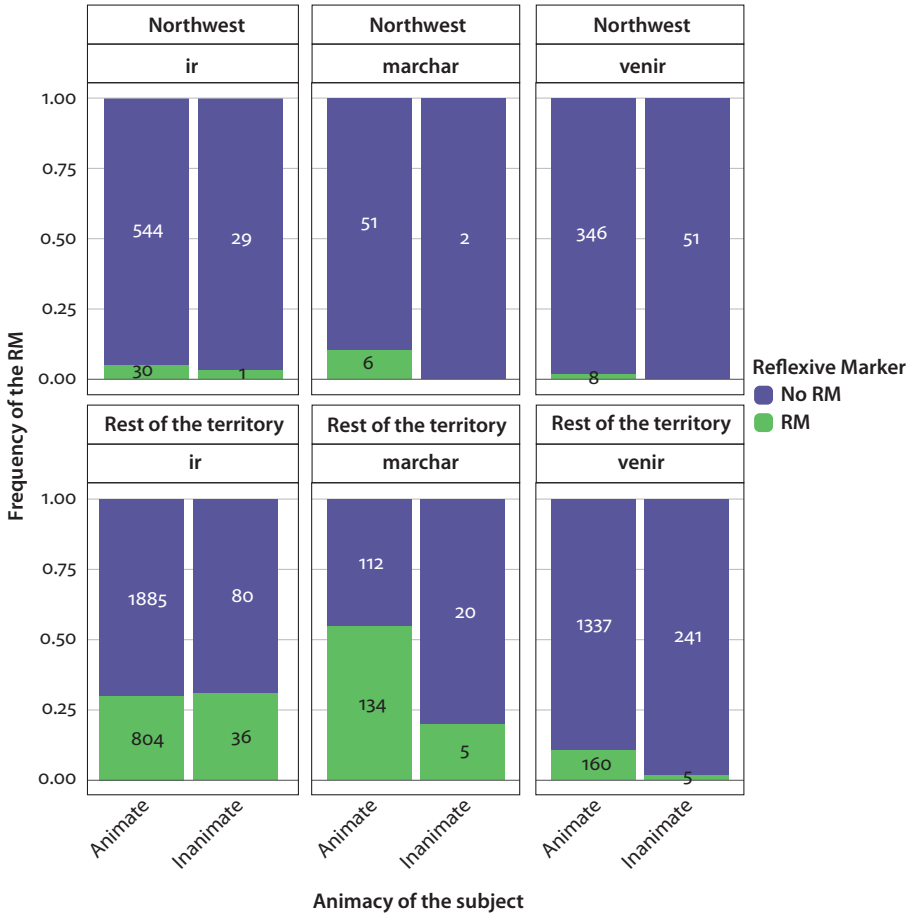


Figure 31. Frequency of the RM by animacy of the subject with *ir(se)*, *marchar(se)* and *venir(se)* by geographical area

10. Imperatives are also diaphasically marked, since they are rarely used in formal contexts. Although my data are not suited to the analysis of the diaphasic conditions that favour the presence of the RM, we should recall that they are very likely to play a role (Cartagena 1972). Thus, the effect of the imperative could be related to such diaphasic conditions instead of (or together with) agentivity.

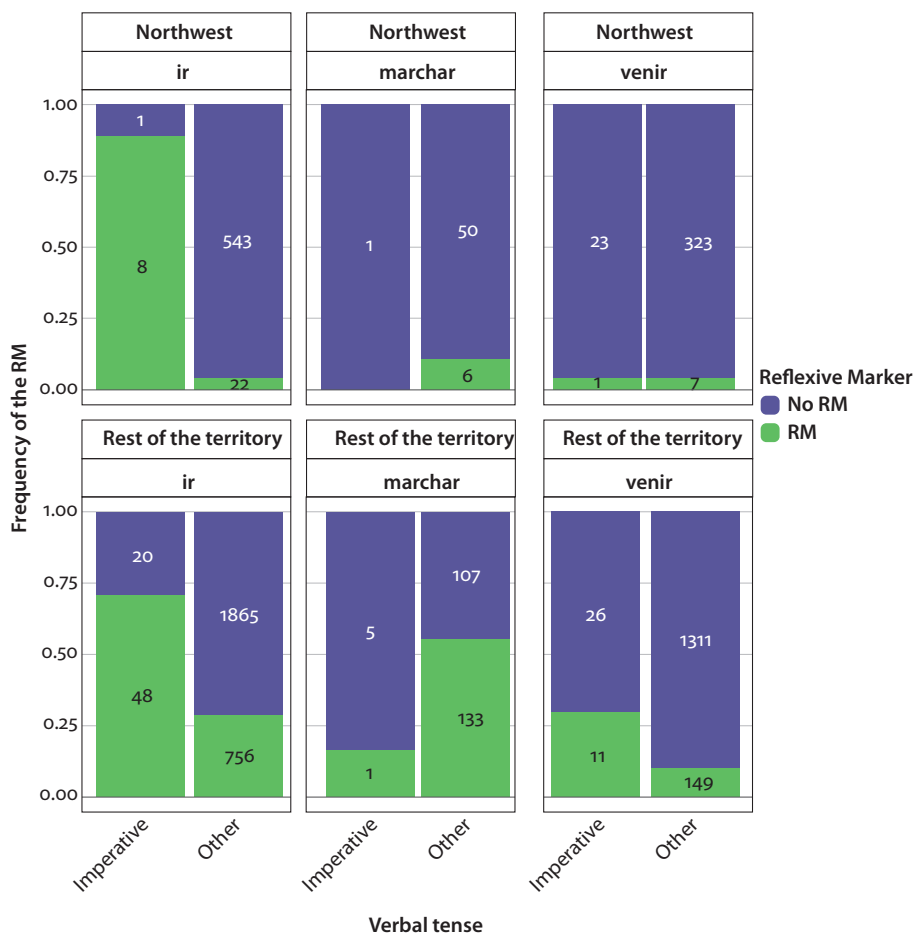


Figure 32. Frequency of the RM by verbal tense with *ir(se)*, *marchar(se)* and *venir(se)* by geographical area

However, not all the verbs considered here show an effect of the animacy of the subject – *salir(se)*, *escapar(se)* and *pasar(se)* show no trace of such an effect. I believe, nevertheless, that a connection between older functions of the RM and the distribution of the RM with these verbs can be found in the meanings they tend to encode. As observed above, *salirse* seems to be most frequent in contexts where the source of the movement is a place that ought not to be abandoned. This meaning is inherent to *escapar*, which shows a far higher frequency of the RM (see Figure 29 above). *Pasar(se)* shows a similar situation, since most marked examples imply the crossing of a limit that should not have been crossed, as shown in (121). The marked form appears in 72.7% (32/44) of the cases in such contexts, while in other contexts *pasar(se)* is only marked in 5.9% of the cases (21/358).

- (121) Bueno, nos estamos pasando a otra cosa...
 good REFL.1PL be.PRS.1PL passing to other thing
 ‘Well, we’re moving to something else now...’ (Ausejo, La Rioja, COSER 2501)

Hence, these verbs show a probabilistic association between the RM and situations where the subject is contravening the expectations of an external participant. I believe this to be associated with a common implicature that arises with anticausative verbs, as was discussed in Section 4.5, namely, the existence of an external argument that causes (directly or indirectly) the event, even against this participant’s will, as with unintentional causers (see (122)). If this is so, we would expect a high association of the RM with affected source datives.¹¹

- (122) Se me ha roto un vaso
 REFL.3 DAT.1SG have.PRS.3SG break.PTCP a glass
 ‘I broke a glass unintentionally’

The parallel is clear in the case of *escapar(se)*, where the typical role of the dative is to encode the affected external argument against whose will the subject is escaping, as shown in (123). From (12) examples found in my sample, only one is unmarked, but since *escapar* is almost always marked, this is no evidence of an association between the RM and the dative.

- (123) El soldao que se le escapaba tenía que cumplir
 the soldier that REFL.3 DAT.3SG escape.PST3SG have.PST3SG that keep.INF
 la condena.
 the sentence
 ‘The soldier that got away from him had to serve a sentence.’
 (Higuera, Albacete, COSER 0211)

My data are too scant to check also the effect of the different types of datives in the distribution of the RM with *pasar(se)* ($N = 9$) or *salir(se)* ($N = 18$). Aaron & Torres Cacoullos (2005) analyse the behaviour of *salir(se)* in Mexican Spanish and find an association between datives and the presence of the RM, since 66% of the time that a dative appears they find the marked form *salirse* (10/15, but note the low figures). They do not discuss the interpretation of these datives specifically, but take it to be a measure of affectedness. They find similar figures for Old Spanish, proposing that a high level of affectedness has always favoured the form *salirse* over the course of the history of the language. It should be noted, however, that this affectedness does not refer to the subject of *salirse*, but to an external participant, providing a plausible link with anticausative constructions.

11. For the pair *subir(se)* / *bajar(se)*, Torres Cacoullos & Schwenter (2008) note that most marked instances with inanimate subjects co-occur with a dative, although they do not discuss the interpretation of these datives.

The behaviour of *ir(se)* can shed some light on the connection between datives, affectedness and translational motion. The examples of *ir(se)* with a dative found in the COSER show a perfect split between *ir*, which takes datives that encode the goal of a movement (see (124)), and *irse*, which takes datives that encode either the (affected) source of a movement (see (125a)), the (affected) cause of a physical movement (see (125b)), or an affected possessor (see (125c)).

- (124) [sosa cáustica] Hombre, te va pa una mano o pa un
 man DAT.2SG go.PRS.3SG for a hand or for an eye
 ojo, hay que tener mucho cuidado.
 have.PRS.3SG that have.INF much care
 [caustic soda] 'Well, it can splash to your hand or your eye, one must be careful.'
 (Mahíde, Zamora, COSER 4617)
- (125) a. Decía la gente que se le iba la leche, no
 say.PST.3SG the people that REFL.3 DAT.3SG go.PST3SG the milk not
se va nada, se va la que no quiere
 REFL.3 go.PRS.3SG nothing REFL.3 go.PRS.3SG the that not want.PRS.3SG
 dar teta.
 give.INF tit
 'People said the milk was gone, milk does not go, it goes if you don't want
 to breastfeed.'
 (Zufre, Huelva, COSER 2122)
- b. [el cántaro] Así se me iba de lao, justo.
 so REFL.3 DAT.1SG go.PST3SG of side exactly
 '[The jug] slipped to the side, like that, exactly.'
 (Tramaced, Huesca, COSER 2229)
- c. [su nieto] Que me se ha ido, a jugar a otro
 that DAT.1SG REFL.3 have.PRS.2SG go.PTCP to play.INF to other
 pueblo.
 town
 '[his grandson] He's left, he plays (football) in a different town now.'
 (San Mateo de Gállego, Zaragoza, COSER 4719)

That is, *irse* is associated with datives that encode the source of the movement. This is reminiscent of the well-established fact that *ir(se)* must take the RM when a source complement with the preposition *de* is present (see (116)), this also in Galician (Bull 1952; Babcock 1970; Martín Zorraquino 1979, 1993 Cidrás Escáneo 1991; Morimoto 2008):

- (126) Foi*(se) da oficina ás dúas.
 go.PST.3SG-REFL.3 of.the office to.the two
 'He left the office at two.'
 (Cidrás Escáneo 1991: 117)

This is confirmed in the COSER data, as long as the preposition *de* ‘of’ is not com-
mutable by *desde* ‘from’ (see also De Benito Moreno 2015). Interestingly, both *salirse*
and *pasarse* seem also to show a stronger association with source complements with
de, as can be seen in Figure 33,¹² despite the low figures in some categories: marked
salirse and *pasarse* reach their highest frequency (17% and 42% respectively) with
source complements with *de*, despite the higher frequency of the unmarked forms
in all contexts. Aaron & Torres Cacoullos (2005) report a similar association be-
tween *salirse* and source complements with *de* as early as Old Spanish. Note that
Figure 33 also shows that the marked forms of these verbs are virtually non-existent
in the northwest.

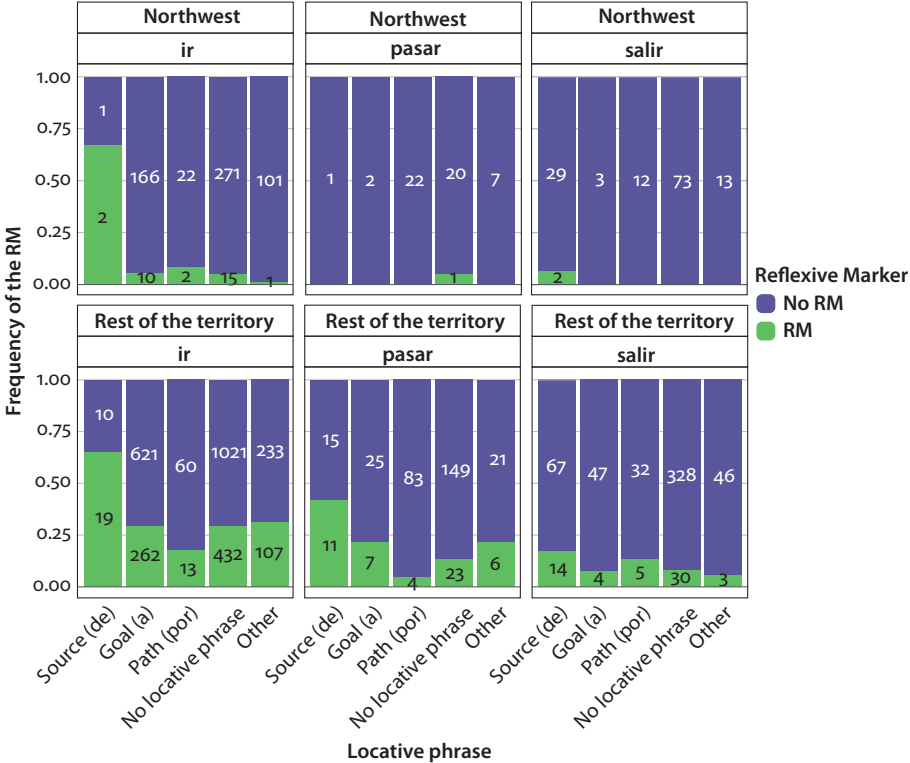


Figure 33. Frequency of the RM with *ir(se)*, *salir(se)* and *pasar(se)* by locative phrase by geographical area

12. Figure 33 considers only the four most frequent contexts: the presence of a goal complement with preposition *a*, the presence of a source complement with the preposition *de*, the presence of a path complement with the preposition *por* and the absence of locative complements.

It is worth recalling that *ir(se)* seems to be one of the oldest reflexive intransitive verbs (Martín Zorraquino 1993; Bogard 2006). In this sense, I believe that *salirse* and *pasarse* show an incipient situation which could lead to the creation of a sharp grammatical contrast: due to the association of the RM with undue movements, it appears more frequently in the discourse with *de* source complements and this could end up in a strong association between these two characteristics, as is the case with *irse*. That is to say, I propose that different translational motion reflexive intransitive verbs can be seen as representing different diachronic stages. The older they are, the sharper the contrast between the unmarked and the marked variant. I fully agree with Aaron & Torres Cacoullós (2005), who propose that the current function of *se* in *salirse* (which they describe as “the subjectification of [a] counter-expectation marker”) “appears to proceed via the absorption of contextual meaning in the frequently occurring construction with *de* ‘from’” (Aaron & Torres Cacoullós 2005: 609). I believe that the same idea holds for other verbs, such as *pasar(se)* or *ir(se)*. This, however, does not mean that there is a saturation point in the possibilities of the RM: as *irse* clearly shows, there are some contexts where a sharp grammatical contrast is found, while in other contexts the marked and unmarked forms might appear. In some cases, this process might end in the conversion of the intransitive verb into a non-reversible reflexive verb, that is, a verb that always shows a RM: *escapar(se)* seems to be on its way to completing this process.

Due to their different semantics, other verbs in this class show an association between the marked form and another preposition. For instance, in the rest of the territory, *subirse* is more frequent than *subir* when the goal complement is a prepositional phrase with *en*, despite the fact that the most frequent goal preposition with *subir(se)* is *a* (see Figure 34). This is undoubtedly due to the characteristics of the movement most frequently depicted by *subirse*, as can be seen by the fact that *montar(se)* – which always depicts this kind of movement – clearly prefers the goal preposition *en*, taking *a* very seldom. Note that Figure 34 again shows the resistance of north-western varieties to the presence of the RM (and that *montar* was completely absent in this area). Unfortunately, although Torres Cacoullós & Schwenter (2008) also report a strong association between the presence of the RM and the co-occurrence of some prepositions (*a*, *en*, *de*), they do not consider any differences between them, so we do not know if the differences found in the COSER data replicate their findings.

To conclude, I have shown several similarities between translational motion intransitive reflexive verbs and anticausative verbs. I have argued that the distribution of the RM in translational motion intransitive reflexive verbs is linked both to a higher degree of responsibility of animate subjects and to the existence of an external cause or affected participant, paralleling its distribution with anticausative verbs. Importantly, I have noted that not all verbs in this category show the same behaviour, and have proposed that this is due to the fact that the presence of the RM

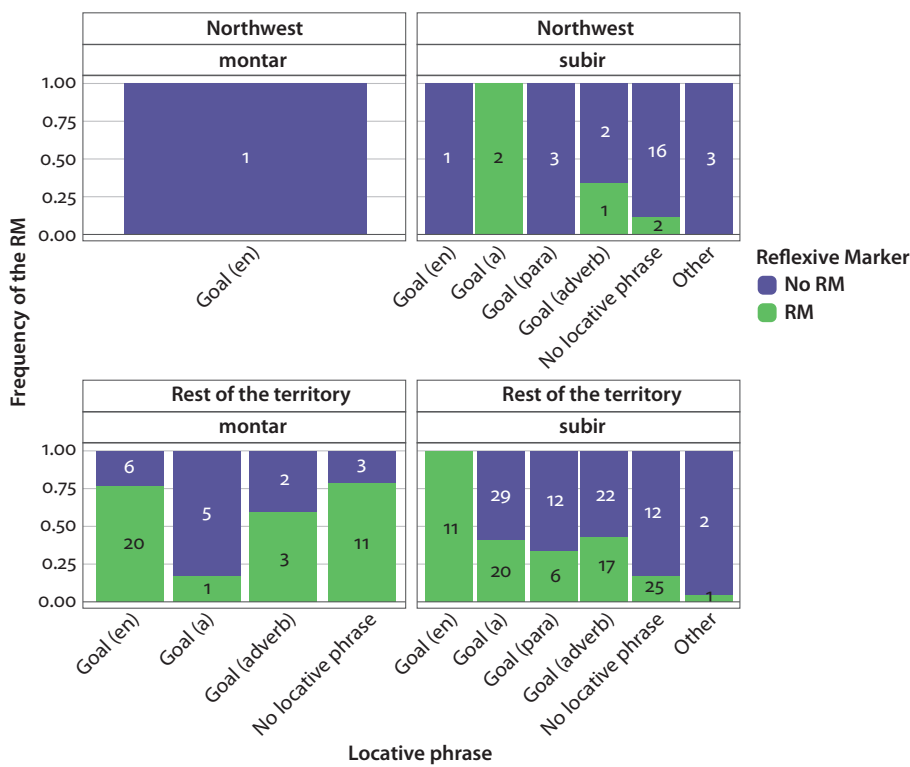


Figure 34. Frequency of the RM with *montar(se)* and *subir(se)* by locative phrase by geographical area

has arisen as the product of analogies with different verbs and because the reflexive verbs considered here are not at the same developmental stage, with some being older than others. A generalised linear mixed model based on the data of the rest of the territory (Table 30) confirms that both being in the imperative and having an animate subject favour the presence of the RM (the former having a stronger effect), as well as the fact that there is notable variability by verb.

Moreover, I have proposed that the strict grammatical contrasts that we find nowadays with some verbs (such as *irse*) are the result of the absorption of contextual meaning, that is, the obligatorification of the RM in the contexts where it most frequently appeared at the beginning of the process. In this sense, the association of *irse* with a specific preposition is not due to the original function of the RM as an aspectual marker, but a consequence of the fact that the contexts which trigger the presence of the RM (i.e. responsible subjects or external causes / affected participants) often present a *de* source complement.

Table 30. Generalised linear mixed model. Presence of the RM in the rest of the territory depending on subject animacy [reference level: Inanimate] and verbal tense (imperative vs. other) [reference level: Other] (fixed factors) and the place and verb lexeme (random factors)

Effect	Term	Estimate	Std. error	Statistic	Conf. low	Conf. high
fixed	(Intercept)	-1.336	0.575	-2.322	-2.463	-0.208
fixed	Subject animacy: Animate	0.654	0.134	4.871	0.391	0.917
fixed	Verbal tense: Imperative	1.159	0.184	6.293	0.798	1.519
random (COSERID)	sd (Intercept)	0.777	NA	NA	NA	NA
random (Verb)	sd (Intercept)	1.648	NA	NA	NA	NA

5.5 *Morir(se)* and *caer(se)*

In this section I focus on two verbs which have been considered instances of the so-called aspectual *se*, namely, *morir(se)* ‘to die’ and *caer(se)* ‘to fall (down)’. These two verbs seem to be connected somehow, inasmuch as the presence of the RM seems to be related to the notion of (lack of) intentionality (see also García Fernández 2015). As regards *morir(se)*, Maldonado (1988: 161) notes that the marked form “implies a high level of accidentality, i.e. the speaker’s expectations play an important role in the construal, as they constitute the agonistic forces overwhelmed by the undesirable happening”. Other authors have noted that *morirse* never represents a violent death, while *morir* may (Cuervo 1954 [1907]: § 103, Merino Hernández 2019, but see deMello 1997, whose study suggests that it might not be so in every Spanish variety). A contrast like the one shown in (127), however, can be explained both on the grounds of “accidentality” and in terms of aspectual factors. Thus, de Miguel & Fernández Lagunilla (2000) consider that, since the RM highlights the phase where the change of state occurs, only secondary predicates referred to the state of the subject at that precise moment are acceptable.

(127) (De Miguel & Fernández Lagunilla 2000: 35)

- a. Juan (**se*) murió fusilado.
John REFL.3 die.PST.3SG shoot.PTCP
‘John died of a shot.’
- b. Juan *se* murió tranquilo.
John REFL.3 die.PST.3SG calm
‘John died calmly.’

Similar observations have been made for *caer(se)*. Maldonado (1988) models the difference between *caer* and *caerse* on the distinction between absolute and energetic construals, proposing that *caerse* appears in unintentional and unexpected

events (energetic construals), while *caer* appears in situations where the event of falling is expected (absolute construals), producing contrasts such as the one in (128). While this generalisation is intuitively true, it feels too strong. The COSER data show examples where *caerse* is used even if an intentional action is performed in order to cause the event of falling (see (129)).

(128) (Maldonado 1988: 154)¹³

- a. Juan (*se) cayó al agua con toda elegancia.
 Juan REFL.3 fall.PST.3SG to.the water with all elegance
 ‘Juan fell into the water with elegance.’
- b. Juan *(se) cayó al agua vestido.
 Juan REFL.3 fall.PST.3SG to.the water dress.PTCP
 ‘Juan fell into the water dressed.’

(129) Entonces se escurrían los panales con la mano. Y debajo
 then REFL.3 drain.PST.3PL the honeycomb with the hand and below
se caía la miel.
 REFL.3 fall.PST.3SG the honey
 ‘Then the honeycombs were drained by hand. And the honey fell down.’
 (Fuente del Pino, Murcia, COSER 3107)

Other authors have associated the presence of *se* with aspectual notions. For instance, Sanz & Laka (2002) claim that *caer* depicts an activity, while *caerse* depicts an accomplishment, on the basis of examples such as (130). However, this does not seem to hold across the board, as (129) above shows, in which *caerse* is an activity.

(130) Vio una estrella cayendo(*se).
 see.PST.3SG a star falling-REFL.3
 ‘S/he saw a star falling.’ (Sanz & Laka 2002: 327)

A more fine-grained distinction is proposed by Teomiro (2013), who explains the contrast in (131a) by proposing that *caer* denotes a change of location, while *caerse* denotes a change of state (see also García Fernández 2015).¹⁴ However, once again this seems to be too strong a claim, since it is unclear whether examples such as (131b), with an explicit locative complement, convey less of a change of location than (131a).

13. I would have not starred any of those examples as ungrammatical, although I do agree with the general pattern.

14. Despite the fact that *caer(se)* has been considered a verb of movement (Sánchez López 2002) or a change of location (Teomiro 2013), the variation in the presence of the RM has been associated with different types of complements of manner (De Miguel & Fernández Lagunilla 2000) instead of locative complements, as has been the case for other verbs of movement, such as *irse* or *venirse* (see Section 5.4).

- (131) a. Juan (**se*) *cayó* cinco metros.
 John REFL.3 fall.PST.3SG five meters
 'John fell five meters.' (Teomiro 2013)
- b. [*las manzanas, de los árboles*] *Se caen* al suelo y
 REFL.3 fall.PRS.3PL to.the floor and
allí se pudren.
 there REFL.3 rot.PRS.3PL
 '[apples, from the trees] They fall to the floor and they rot there.'
 (Quintana de los Prados, Burgos, COSER 0939)

The COSER data offer statistical support for the association of the RM with the notion of accidentality. For *morir(se)* I have operationalised it by distinguishing violent deaths from other kind of deaths. When the death is not violent (or there was no hint that it was so), 68.5% (267/390) of the examples are marked with the RM, whereas with violent deaths *morir(se)* is marked half the time (50%, 17/34). Such a high degree of markedness seems to contradict the claims described above, although it is worth noting that from the 17 marked examples, 11 refer to animals, whose deaths are of course violent, but are most likely not conceived of as such in a rural society, where the slaughter of animals is a necessary element. Regarding *caer(se)*,¹⁵ I have made a distinction between expected or planned falls and accidental or unwanted falls. In the latter case, 72.8% (123/169) of examples of *caer(se)* are marked, whereas when the fall is expected *caer(se)* very rarely takes the RM (8.4%, 19/227), showing a very sharp contrast between the two situations.¹⁶

It is clear, then, that the RM is associated with the meaning of accidentality to a large extent. I believe that this association is related to the meanings that anti-causative constructions typically take, especially the fact that such constructions reject external agents, that is, agents that are not the subject (see Section 3.2.1). This explains why *morirse* is rarely used when the subject was a human killed by someone else, and why *caerse* is typically used when the event of falling is the product of an involuntary event or an indirect cause. The reason why these two verbs show similar behaviours is probably the fact that they depict events that can frequently be caused by an agent. In fact, they both have a frequent lexical causative pair, *matar* and *tirar*. Moreover, *caer* can be used as causative in some western varieties of Peninsular Spanish (De Benito Moreno 2015; Jiménez-Fernández & Tubino 2014,

15. Instances of *caer* which do not mean 'to fall down' but show other meanings, such as *caer bien* 'to get along' or *caer dormida* 'to fall asleep', were excluded.

16. Ten examples were considered to be ambiguous.

Lara Bermejo 2020).¹⁷ In addition, *muerto*, the past participle of *morir*, can be used as the past participle of *matar* ‘to kill’ in the literary language (NGLE 2009: 4.12j). In what follows I analyse quantitatively the influence of the parameters that have been shown to be relevant for the distribution of the RM with anticausative verbs, i.e. the animacy of the subject and the presence of a dative.

First of all, it is worth noting that the frequency of the RM differs greatly between *morir(se)* and *caer(se)*. While *morirse* was even more frequent than *morir* in my data,¹⁸ with 67% (284/424) examples of *morir(se)* showing the RM, *caer* was far more frequent than *caerse* – only 31.3% (147/406) of the cases of *caer(se)* showed the RM. However, this difference might be attributable to the animacy of the subject. On the one hand, of course, *morir* is clearly a verb that typically takes animate subjects (see Figure 35).¹⁹ With *caer(se)*, on the other hand, the animacy of the subject is a highly relevant parameter in explaining the distribution of the RM, as shown in Figure 35. *Caerse* is far more frequent with animate subjects, reaching a very similar distribution to *morir(se)*, while it is dispreferred with inanimate subjects.²⁰ Figure 35 also illustrates quite clearly that marking is dispreferred for both verbs in north-western varieties, even if the data are scant.

The presence of a dative seems to have a crucial effect on the presence of the RM. Both DeMello’s (1997) and Merino Hernández (2019) confirm the well-established contrast according to which only the marked form *morirse* is possible when a dative is present ($N = 12$ and $N = 35$, respectively). DeMello attributes this situation to the

17. The data in De Benito Moreno (2015: 681–683) suggest that the presence of such alternation does not increase the frequency of the RM, as one might have expected, since anticausative verbs are typically marked by the RM (see also Section 4.3). The villages that showed the causative alternation (transitive *caer* vs. intransitive *caer(se)*) showed no increase in the presence of the RM in *caer(se)*. That is, when *caer(se)* has a causative reading, it behaves like *subir(se)*, *bajar(se)* and *pasar(se)*, thus fitting better in the group of intransitives.

18. Merino Hernández (2019) finds very similar proportions in Mexican Spanish. DeMello (1997) found frequencies between 20% (La Paz) and 66% (Madrid) of *morirse* in his study on *morir(se)* in the educated speech (*habla culta*) of several Spanish-speaking cities.

19. There were only 3 examples of *morir(se)* with inanimate subjects, one referring to the disappearance of an old habit, a second one referring to a plant, and the final one explicable in terms of metonymy (the subject is *panales* ‘honeycombs’ but the death event refers to the *abejas* ‘bees’).

20. This contrasts with Maldonado’s claim (1988: 154) that “[t]he contrast between reflexive and non-reflexive formations does not depend on whether the participant is animate or not”. Of course, such a claim is intended to highlight the fact that both animate and inanimate subjects can appear with both forms, *caer* and *caerse*. However, only by studying large amounts of data from a quantitative point of view can we discover more subtle tendencies, which, as shown in Figure 35, indicate that animacy of the subject is actually a very relevant factor.

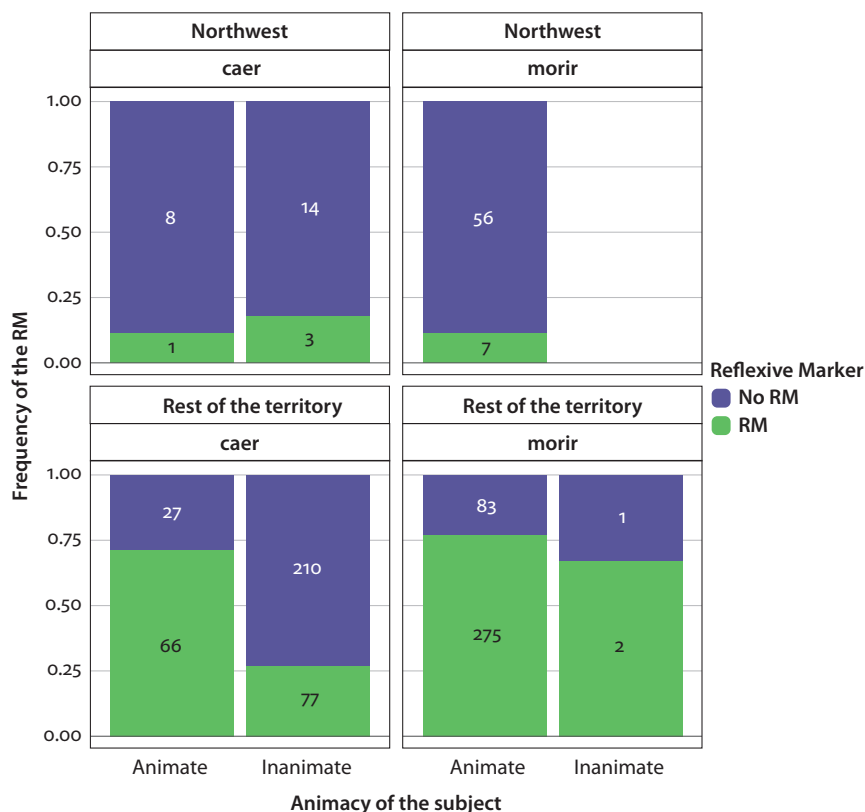


Figure 35. Frequency of the RM by animacy of the subject and geographical area in *caer(se)* and *morir(se)*

notion of ‘separation’ that can be seen in these examples, which he connects with Bull’s (1952) claims about *salirse*. The situation is rather similar in our data, with one caveat: the effect of the dative is not the same across varieties. The presence of the dative has no effect in north-western varieties, while it triggers the presence of the RM in the rest of the territory (see Figure 36) – the only unmarked example was found in León, an area that shows many linguistic similarities to the northwest. Regarding *caer(se)*, Figure 36 suggests that the presence of a dative favours the presence of the RM, although the absence of the RM is also allowed in the rest of the territory. Again, no traces of such an effect are found in the northwest.

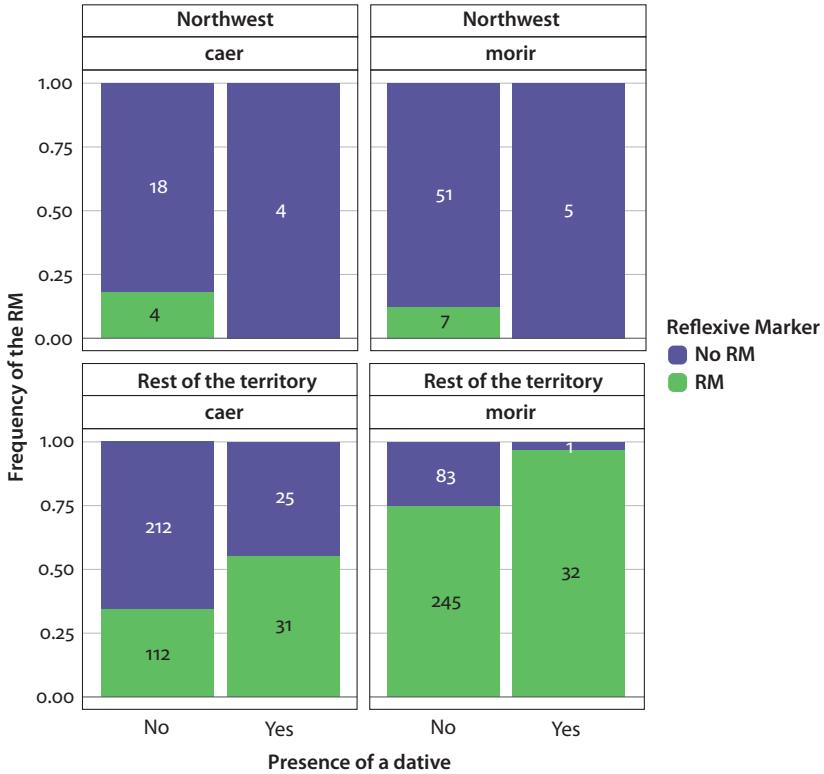


Figure 36. Frequency of the RM by presence of a dative and geographical area with *caer(se)* and *morir(se)*

Although scant, a closer look at the data (see Figure 37) reveals that in most of the unmarked cases of *caer* the dative is a (goal) locative (see (132)). On the contrary, *caerse* is more frequent with datives that express a source or a possessive (see (133a)), two values that are often found simultaneously, or a cause (see (133b)). That is, the behaviour of the RM with *caer(se)* and *morir(se)* regarding both the animacy of the subject and the presence of a dative is reminiscent of the behaviour of anti-causative verbs. I propose that these similarities are due to paradigmatic analogy.

- (132) Y, claro, le caía agua al... aquello.
and clear DAT.3SG fall.PST.3SG water to.the that
'And, of course, water fell over that.' (Barrax, Albacete, COSER 0204)

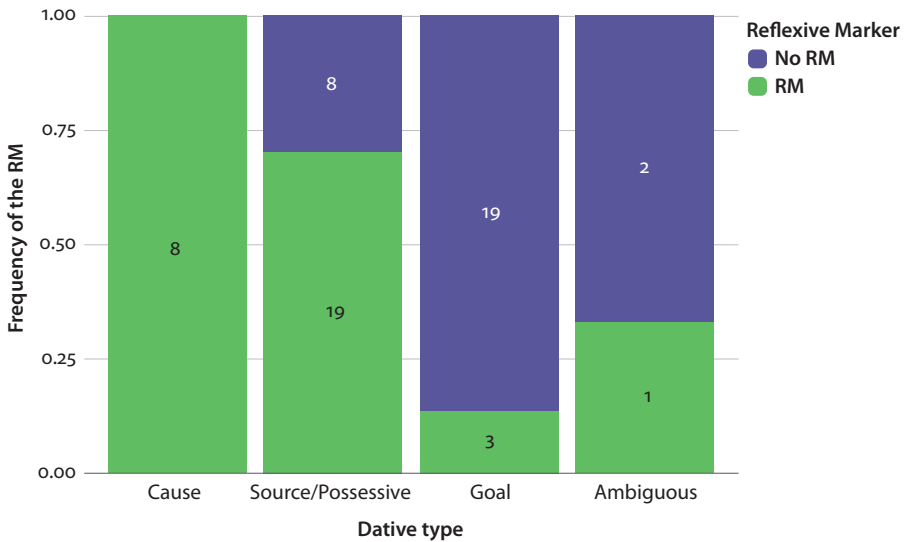


Figure 37. Frequency of the RM by dative type with *caer(se)*

- (133) a. Y uno..., se le cayó un botón a la camiseta.
 and one REFL.3 DAT.3SG fall.PST.3SG a button to the t-shirt
 ‘And one... a button fell off his t-shirt.’ (Liétor, Albacete, COSER 0214)
- b. Un papelillo te se ha caído.
 a paper DAT.2SG REFL.3 have.PRS.3SG fall.PTCP
 ‘You’ve inadvertently dropped a paper.’
 (Antequera, Málaga, COSER 3001)

A generalised linear mixed model based on the data from the rest of the territory (see Table 31) confirms that both the presence of a dative and having an animate subject favour the presence of the RM (having the former a stronger effect) and that both verbs behave very similarly.

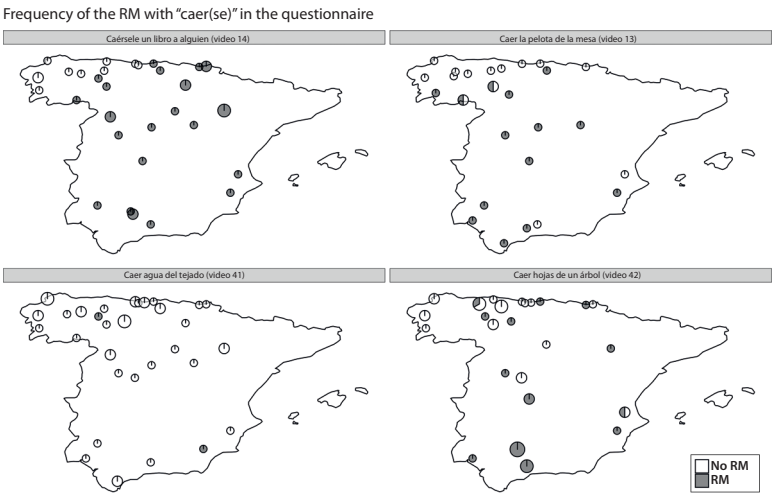
The data of the questionnaire allow us to investigate a little further the differences between *caer* and *caerse*. First of all, thanks to video 14, *caér(se)le un libro a alguien* ‘to fall (a book) that was being carried by someone’, the questionnaire confirms that the presence of a dative coding the involuntary cause (and source of the movement) typically triggers the presence of the RM, with the exception of

Table 31. Generalised linear mixed model. Presence of the RM in the rest of the territory depending on subject animacy [reference level: Inanimate], presence of a dative [reference level: No dative] and verb lexeme (*caer* vs. *morir*) [reference level: *Caer*] (fixed factors) and the place (random factor)

Effect	Term	Estimate	Std. error	Statistic	Conf. low	Conf. high
fixed	(Intercept)	−1.423	0.204	−6.96	−1.824	−1.022
fixed	Subject animacy: Animate	2.471	0.337	7.34	1.811	3.13
fixed	Presence of a dative: Yes	1.732	0.353	4.903	1.04	2.424
fixed	Verb: <i>morir</i>	0.278	0.304	0.913	−0.319	0.874
random (COSERID)	sd (Intercept)	0.961	NA	NA	NA	NA

north-western varieties, which still prefer *caer* in this case (see Map 11). This dialectal difference parallels the one found in Section 4.5 for anticausative verbs, once more supporting the hypothesis of the diachronic connection between both uses.

There are three more instances from the questionnaire which show a contrast that cannot be attributed either to animacy or to accidentality or expectedness. Videos 13 (*caer la pelota de la mesa* ‘to fall (the ball) out of the table’), 42 (*caer hojas, frutos de un árbol* ‘to fall (leaves, fruits) from a tree’) and 41 (*caer agua del tejado* ‘to fall (water) from the roof’) were designed to elicit examples of *caer* in different contexts. Maldonado (1988) considers the example *La pelota se cayó de la mesa* and argues that the RM is expected in this case because “the ball is expected to remain on the table and its falling constitutes an unexpected action, an event that



Map 11. Frequency of the RM with *caer(se)* in the questionnaire

contradicts a normal expectation” (Maldonado 1988: 153). Map 11 confirms that the presence of the RM is the norm in such an example in most of the area under study, although north-western varieties clearly prefer *caer* in this context and the absence of the RM is also possible outside the area.

However, note that a very unexpected event, like water falling off the roof on a sunny day, is consistently conveyed by *caer* in the whole study area (except for a few exceptions), as Map 11 shows. On the other hand, an expected event such as leaves falling off a tree, is more often conveyed by *caerse* than by *caer* (again leaving aside north-western varieties), as shown by Map 11.

That is, we find a very consistent contrast that is borne out for most speakers (except for north-western speakers) but we cannot explain it on the grounds of expectedness. Actually, I believe that this contrast is related to the notion of movement from a delimited closed space or against an obstacle, as discussed regarding *salir(se)* (see Section 5.4). In the cases of both *Caer la pelota de la mesa* (video 13) and *Caer hojas, frutos de un árbol* (video 42) the subject falls out of a delimited closed space (the table and the tree). In the case of *Caer agua del tejado* (video 41), however, the roof is not a delimited closed space, since it is actually designed to make water fall. This can ultimately be connected (as with *salir(se)*) to the distinction between external and internal causation and the notion of affected external participants that were discussed in Sections 4.3 and 4.5 and for anticausatives. *Caer la pelota de la mesa* is an event that very likely lacks an agent, but needs an external cause and typically affects (negatively) an external human participant (whoever wanted the ball on the table). *Caer hojas, frutos de un árbol* does not affect any external human participant, but typically needs an external cause (such as the wind, which was mimicked in the video). Finally, *caer agua del tejado* is the event that can be most easily conceived of as spontaneous, since the inclination of the roof enables the downward movement, requiring no external cause.

5.6 Corporal (physical and psychological) processes

In this section I will focus on some verbs that depict corporal (physical and psychological) processes, such as *dormir(se)* ‘to sleep’, *reír(se)* ‘to laugh’, *cagar(se)* ‘to have a shit’, *mear(se)* ‘to piss’. While *dormir(se)* and *reír(se)* are often mentioned in works on aspectual *se* or reflexive intransitive verbs (Sánchez López 2002; De Miguel & Fernández Lagunilla 2000; Martín Zorraquino 1979; Maldonado 1999), *cagar(se)* and *mear(se)* have never been considered in the literature, as far as I know. As regards *reír(se)*, Martín Zorraquino (1979) notes that the presence of *se* is required when a source phrase with *de* is present (see (134)) – a pattern clearly reminiscent of *irse* ‘to go’, which also requires the RM when a source phrase is present.

- (134) a. Juan (se) rio mucho.
 John REFL.3 laugh.PST.3SG much
 'John laughed a lot.'
- b. Juan *(se) rio de alguien.
 John REFL.3 laugh.PST.3SG of someone
 'John laughed at someone.'

Regarding *dormir(se)*, many authors have noted that *dormirse* seems to be specialised in the punctual event of 'falling asleep', while *dormir* has the more general meaning of 'sleeping'. While this has been connected to the aspectual function of the RM by some authors (Sánchez López 2002; De Miguel & Fernández Lagunilla 2000), certain others have noted that this distribution of the RM might be due to other factors. For instance, González Vergara (2006) notes that *dormir* actually has a transitive variant (*dormir* means 'to put someone to sleep') and that *dormirse* can hence be interpreted as an anticausative which would mean 'to fall asleep', while intransitive *dormir* would then mean 'to sleep'.

Intuitively, there is also a semantic difference between the marked and unmarked pairs of *cagar(se)* and *mear(se)*, which is quite clear when the verb is in the progressive (see (135)). There is a difference between the unmarked forms (*cagar*, *mear*), which in the gerund only depict an activity that is currently taking place, and the marked forms (*cagarse*, *mearse*), which in the gerund can also depict the urgent need to carry out such an activity. This difference seems to be borne out in all imperfective tenses.

- (135) a. Estoy {cagando /meando}
 be.PRS.1SG shitting pissing
 'I'm taking a {shit / piss}.'
- b. Me estoy {cagando /meando}
 REFL.1SG be.PRS.1SG shitting pissing
 'I'm about to take a {shit / piss}' – 'I'm taking a {shit / piss}'.

Interestingly, this contrast is paralleled by other verbs, such as *dormir(se)* and *caer(se)*, which in the unmarked form clearly imply that the action is taking place, while in the marked form they might also mean that the action is about to take place (see (136)).

- (136) a. Estoy {durmiendo /cayendo}
 be.PRS.1SG sleeping falling
 'I'm {sleeping / falling}.'
- b. Me estoy {durmiendo /cayendo}
 REFL.1SG be.PRS.1SG sleeping falling
 'I'm about to {fall asleep / fall down}' – 'I'm {falling asleep / falling down}'.

This association between the RM and the incipency of a corporal process seems to be related to lack of control – physical processes such as *cagar(se)*, *mear(se)* and *dormir(se)* cannot always be controlled by the subject and might happen against the subject's will. The incipency reading refers precisely to the phase of the event when the subject feels they are about to (spontaneously, without being able to control it) suffer the corporal process.²¹

This lack of control associated with *dormir(se)*, *cagar(se)* and *mear(se)* is reminiscent of a class of verbs that is systematically marked by the RM, namely, conservative emotion verbs (see Section 3.2.4). Emotion events, inasmuch as they also depict corporal (psychological) processes, might also occur against the will of their subjects and show a similar behaviour in the imperfective tenses, actually preferring the incipient reading:

- (137) Me estoy asustando
 REFL.1SG be.PRS.1SG getting.scared
 'I'm about to get scared' – 'I'm getting scared.'

Lack of control of the subject is also found in a number of anticausative verbs with human subjects, that differ from the proper reflexive reading in the volitionality of the subject: while the proper reflexive is volitional, the anticausative is not (since it rejects the presence of an agent):

- (137) Me he cortado con el cuchillo
 REFL.1SG have.PRS.1SG cut.PTCP with the knife
 Reflexive: 'I've cut myself with the knife.' ~ Anticausative: 'I've got cut with the knife.'

Moreover, a further connection between *cagar(se)* and *mear(se)* and previous functions of the RM can be identified. The marked forms *cagarse* and *mearse* appear in contexts that are clearly reminiscent of semantically reflexive actions, since they are carried out physically *on* the subject's body. This is illustrated in (138). The marked variant (138a) does not require the object of the spatial adverb to be specified, since

21. The possibility of a verb taking an incipency reading in the progressive periphrasis has been linked to the lexical aspect of the verb. According to the NGLÉ (2009) this reading is only possible with states, while De Miguel & Lagunilla (2000: 18–19) note that this possibility identifies inceptive achievements (*logros ingresivos*), such as *marearse* 'to get dizzy'. *Mear(se)* and *cagar(se)* do not seem to belong in either of these two categories, since they are durative processes. If we were to conceive of them as events with an initial culmination, they would contradict Miguel & Lagunilla's (2000) claim that only inceptive achievements in which the culmination point is followed by a state (like *caer(se)*) admit the RM, since *mear(se)* and *cagar(se)* do not have a final state but rather a process (as with *floreecer* 'to flower', which is unmarked).

it is interpreted by default as the body of the subject (i.e. as reflexive), contrary to what happens with the unmarked variant (138b).

- (138) a. Se {meó /cagó} encima
REFL.3 piss.PST.3SG shit.PST.3SG above
'S/he {pissed/shat} himself/herself.'
b. {Meó /cagó} encima ??(de la mesa)
piss.PST.3SG shit.PST.3SG above of the table
'S/he {pissed/shat} on (the table).'

Once again, the corpus data show that this split in meaning is far from perfect. First of all, these verbs were not very frequent in the COSER interviews, which makes it difficult to perform a quantitative analysis. As Figure 38 shows, the reflexive variant *dormirse* is rather infrequent and both *cagarse* and *mearse* are less frequent than their corresponding unmarked variants.²² The opposite situation is found with *reír(se)*, which is almost always marked in the COSER data.²³

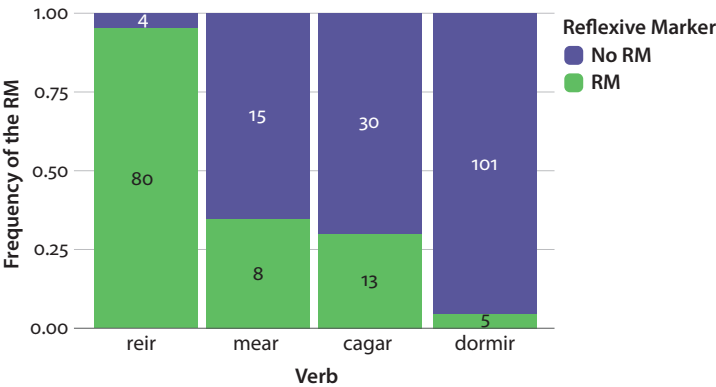


Figure 38. Frequency of the RM with intransitive verbs depicting corporal processes

22. *Dormir(se)* was collected exhaustively only in subcorpus E, while the figures for *cagar(se)*, *mear(se)* and *reír(se)* correspond to the whole corpus.

23. Maldonado (1999) also acknowledges this greater frequency of *reírse* as opposed to *reír*, which he connects to the fact that *reír* typically implies a sudden change (as opposed to *sonreír*, for instance).

Since five examples of *dormirse* are not enough to offer any generalisations, I also collected the occurrences of *dormirse* (but not of *dormir*) in subcorpus NE, obtaining 17 examples in total. These data allow us to see that the RM in *dormir* is not only specialised in marking inceptive meaning. Although most (12 out of 17, 70.6%) of the examples of *dormirse* clearly mean ‘to fall asleep’, such as (139), there are also examples of variation between the marked and the unmarked form, with and without the inceptive meaning (see (140)). That is, the differences between *dormir* and *dormirse* discussed above show a statistical tendency, but they are not completely systematic.

- (139) Pero dice ella que no se **dormía** hasta que no
 but say.PRS.3SG she that not REFL.3 sleep.PST.3SG until that not
 llegábamos.
 arrive.PST.1PL
 ‘But she says that she wouldn’t fall asleep until we would’ve arrived.’
 (Navalmoral de la Mata, Cáceres, COSER 1015)
- (140) a. [la anestesia no le hacía efecto] “[NP], ¿cuándo te vas a
 NP when REFL.2SG go.PRS.2SG to
dormir?” Y digo: “Pero si no tengo sueño, ¿por
 sleep.INF and say.PRS.1SG but if not have.PRS.1SG sleepiness for
 qué voy a dormir?”
 what go.PRS.1SG to sleep.INF
 [anaesthesia was not working] “[NP], when are you falling asleep?”, and
 so. And I said: “But I’m not sleepy, why would I fall asleep?”
 (Tablones, Granada, COSER 1834)
- b. Los hermanos sí, solían dormir [juntos]. Pero hermana
 the brothers yes used.to.PST.3PL sleep.INF together but sister
 y hermano, no. Cada uno se dormía en su cama, no
 and brother no each one REFL.3 sleep.PST.3SG in their bed not
 los echaban juntos.
 ACC.3PL throw.PST.3PL together
 ‘Brothers used to sleep together, yes. But sister and brother didn’t. Each
 slept in their bed, they weren’t put together.’
 (Fuentes Claras, Teruel, COSER 4117)

As regards *cagar(se)* and *mear(se)*, the COSER data show that uncontrolled contexts clearly trigger the marked variant (6/6 in the case of *cagar(se)* and 8/9 in the case of *mear(se)*). Regarding controlled contexts, the unmarked form is preferred in the case of *cagar(se)* (30/37, 81%) and the only one documented in the case of *mear(se)* (14/14). All in all this suggests that uncontrolled contexts are typically conveyed by the marked form, but that the RM could be spreading to controlled contexts, such as in (141).

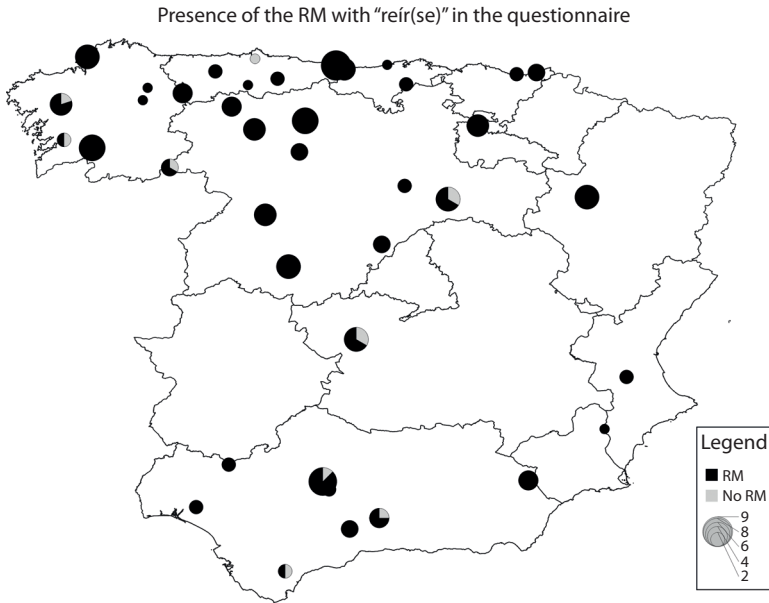
- (141) [la mosca del jamón] Que está cómoda pa poderse cagar ella
 that is comfortable for can.INF-REFL.3 shit.INF she
 allí..., donde está cómoda.
 there where be.PRS.3SG comfortable
 [the ham fly] ‘Where it feels at home to be able to shit there..., wherever it feels
 at home.’ (Ausejo, La Rioja, COSER 2501)

Unsurprisingly, all these verbs took animate subjects in my data. Thus, the opposition between animate and inanimate subjects is not very informative in this case. The same cannot be said for the presence of a dative. Although 13 examples might not allow for definitive statistical analysis, *cagar(se)* shows some interesting tendencies when a dative is present in the structure. When the dative is an affected participant or a possessor (see (142)) both *cagar* (4/7, 57%) and *cagarse* (3/7, 43%) are possible, whereas when the dative is a locative (as in (143)) ($N = 6$), *cagar* was the only attested form. This suggests that the affectedness of an external participant may favour the presence of the RM, as with anticausative verbs (see Section 4.3).

- (142) Que una mosca no fuera y te se cagara por allí.
 that a fly not go.SBJV.3SG and DAT.2SG REFL.3 shit.SBJV.3SG for there
 ‘So that a fly wouldn’t come and shit on your ham.’
 (Enguera, Valencia, COSER 4310)
- (143) [el jamón] Es en otro sitio donde le ha cagao la
 be.PRS.3SG in other place where DAT.3SG have.PRS.3SG shit.PTCP the
 mosca.
 fly
 [the ham] ‘The fly shit on some other part of the ham.’
 (Leza, Álava, COSER 0106)

I would like to mention a rather interesting construction that exploits the combination of the incipient reading of imperfective tenses with these reflexive verbs and the use of “reflexive” *encima* in order to highlight subject affectedness here. Examples such as (144), which were not found in the COSER data but exist in colloquial language, convey the extreme need to carry out an action. Even if those actions are not corporal processes, the construction forces this interpretation by adding the semantic content of ‘feeling the need to do something’.²⁴ Moreover, it admits

24. The examples were provided to me by Twitter users after I had asked them whether they felt a difference between *Estoy cagando* and *Me estoy cagando*. The tweet read “Sin más contexto, ‘Estoy cagando’ es un acción real, pero ‘Me estoy cagando’ es un sentimiento (de que la acción real es incipiente)” ‘With no more context, *Estoy cagando* is a real action, but *Me estoy cagando* is a feeling (that the actual action is incipient)’ and allowed for two different answers (using the survey functionality of Twitter): a) “Sí” ‘Yes’ b) “No, tía, no te ralles” ‘No, girl, don’t be crazy’ and



Map 12. Frequency of the RM with *reír(se)* in the questionnaire data

5.7 *Quedar(se)*, *estar(se)* and *esperar(se)*

In this section I focus on the verbs *quedar(se)* ‘to stay, to become’, *estar(se)* ‘to be’ (both of them analysed only in subcorpus E) and *esperar(se)* ‘to wait’.²⁵ *Quedar(se)* can have different meanings, the two most important being a locative reading (‘to stay, to remain’) and an attributive reading (‘to become’).²⁶ Since both readings show semantic connections with *estar(se)* – although only the locative reading is connected to *esperar(se)* –, I will analyse both readings here. I will, however, always maintain this distinction.

25. Nishida (1994) treats *esperarse* as a transitive verb in examples such as *esperarse un momento*. I disagree with this interpretation, however, and consider *un momento* as a measure phrase that does not share many of the properties of direct objects (such as being pronominalised with an accusative pronoun). For a discussion of why measure phrases are not direct objects, see Rodríguez Molina (2010: 390–412), whom I follow here.

26. Less frequent meanings include (1) ‘to acquire, come into possession of’, where it typically takes a complement with *con*, (2) ‘to meet’, and (3) ‘to agree’, where it typically requires a prepositional object with *en*. Although limited, my data suggest that meaning 1 (possessive) is typically expressed by *quedarse* ($N = 23$), while the unmarked form *quedar* is preferred for meanings 2 (‘to meet’) ($N = 8$), and 3 (‘to agree’) ($N = 6$). No variation was found in any of these contexts.

In what follows I describe how all these verbs show connected behaviours, mostly related to the notion of agentivity. That is, in these verbs the RM seems to be associated with contexts where the subject is more agentive. These verbs, then, differ greatly from the ones analysed in the previous section (where the RM was associated with lack of control of the subject). On the contrary, they are reminiscent of the notion of the higher responsibility of the subject found for human anticausatives (see the discussion of autocausatives in Section 3.2.1).

Before discussing the behaviour of these verbs regarding agentivity, however, two observations must be made. The first refers to the geographical distribution of the RM with these verbs. No examples of *estarse* or *esperarse* were found in north-western varieties (as shown in Map 13), which is why I will only use the data of the rest of the territory when discussing these two verbs.

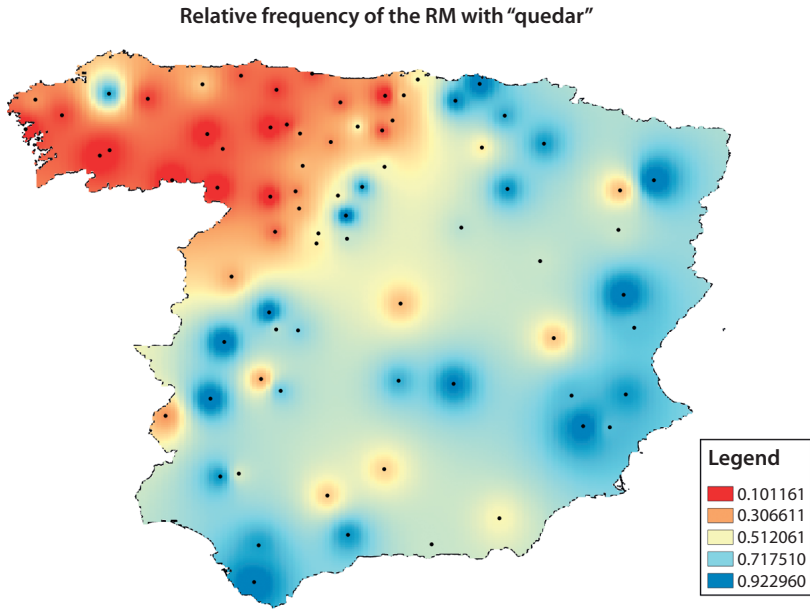
Frequency of the RM with “*estar(se)*” and “*esperar(se)*”



Map 13. Frequency of the RM with *estar(se)* and *esperar(se)* in the COSER data²⁷

Quedarse, although possible in north-western varieties, is considerably less frequent in this area, as shown by Map 14. Accordingly, I will consider these two areas separately when discussing *quedar(se)* in what follows. This geographical pattern clearly suggests a link between the intransitivising function of the RM and the presence of the RM with these verbs, since the former was also less productive in north-western varieties.

27. It should be noted that *estarse* was extremely rare across the territory (see below), which is probably why the map shows some other large areas with no cases of *estarse* at all.



Map 14. Relative frequency of the RM with *quedar* in the COSER data

The second observation refers to the notable differences that these verbs show regarding the frequency of the RM. As can be seen in Figure 39, *quedarse* – in both readings – is just as frequent as *quedar* in the rest of the territory. Marked *esperarse* also shows high frequency – although lower than *quedarse*. However, *estarse* is rather rare as compared to the extremely high frequency of the unmarked form *estar*.²⁸ In the northwest, both readings of *quedar(se)* also show similar frequencies, although far lower than those found in the rest of the territory.

What are the differences between the marked and the unmarked forms? A closer look at locative *quedar(se)* offers an interesting contrast. With human subjects, *quedar* takes on a meaning similar to ‘remain’, while *quedarse* more often conveys a meaning similar to ‘stay’. The main semantic difference between the two is the absence or presence of control of the subject, respectively – that is, the latter is more agentive. In contexts of lack of control (with human subjects, the only ones that can show control), the marked form is used only in 36% of cases (21/58) in the rest of the territory, and in 6% (1/17) of cases in the northwest, the unmarked

²⁸. As mentioned in Section 3.2.6, I collected all the examples of *quedar* in both subcorpus E and in the interviews that showed examples of causative usages of *quedar*. Figures of *estar(se)* in Figure 39 refer only to subcorpus E, from which I collected all the instances of this verb. As regards *esperar(se)*, I collected all its occurrences from the whole corpus.

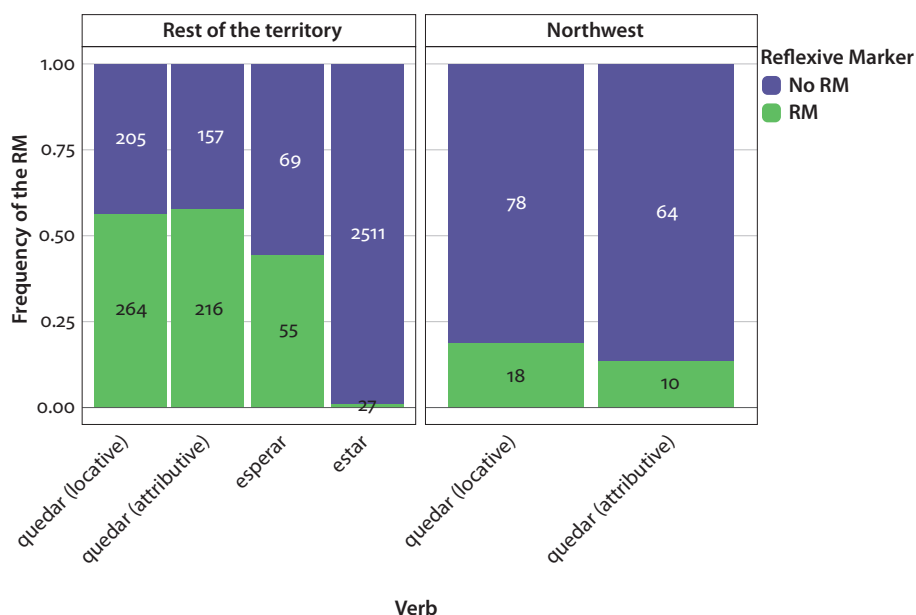


Figure 39. Frequency of the RM with *quedar(se)*, *estar(se)* and *esperar(se)*

form being more frequent (see (145)). On the contrary, when the human subject shows control of the verbal action, the marked form is clearly preferred, used in 96% of cases (171/178) in the rest of the territory, and in 62% (10/16) of cases in the northwest.

- (145) nosotros ya de tantos hermanos quedamos dos, dos na
 we already of so.many siblings remain.PRS.1PL two two nothing
 más
 more

‘Out of so many brothers, only two of us are still alive.’

(Santaella, Córdoba, COSER 1514)

- (146) Los novios nos fuimos y la gente se quedó.
 the grooms REFL.1PL go.PST.1PL and the people REFL.3 stay.PST.3PL
 ‘We (the groom and the bride) left and people stayed.’

(Jarandilla de la Vera, Cáceres, COSER 1012)

Estar(se) and *quedar(se)* are closely semantically related (actually, *stay*, one of the two translations of *quedar(se)* in English, is etymologically related to *estar*) and both admit a locative and an attributive reading. Moreover, the presence of the RM with *estar* has also been associated with agentivity in the literature. According to Sánchez López (2002: 121), the RM with *estar* is a “paradigmatic” case in which

the RM adds a volitional nuance to the verb (see also Morimoto 2008). According to her, this means that the subject of *estarse* is an agent, while the subject of *estar* is an experiencer. Unfortunately, the data for *estarse* are so scant in the COSER interviews that the statistical association of agentivity with the presence of the RM cannot be tested; there are only six total examples of the imperative of *estar(se)*, which is intrinsically agentive, in subcorpus E. In any case, these examples show two unmarked examples (although note that they are negated imperatives), showing that, as always, these contrasts are not absolute:

- (147) [a los hijos] Tú dices: “No estar tan tarde, no estar esto”.
 you.NOM say.PRS.2SG not be.INF so late not be.INF that
 ‘You tell [your children]: don’t be late, don’t be that.’
 (Espera, Cádiz, COSER 1107)

Since the number of examples of *estarse* in subcorpus NE is so low ($N = 27$), I also collected all the instances of marked *estarse* in subcorpus E, garnering a total of 88 tokens of *estarse* from the whole corpus. This allows us to investigate further the contexts where the marked form of *estar* might appear, although it cannot tell us their relative frequencies as compared to the unmarked form. Figure 40 shows the distribution of *estarse* in different syntactic-semantic contexts. Most examples can be categorised into three categories: they were either instances of *estarse* with temporal phrases (see (148)), with adjective attributes (see (149)), or with locative phrases (see (150)).²⁹ Only six examples of *estarse* did not fit in any of these categories (one example showed an interruption, so it was not included in any of these categories). In order to better understand how the behaviour of marked *estarse* differs from the behaviour of unmarked *estar*, I analysed the syntactic context of 150 examples of unmarked *estar*.³⁰ The asymmetry in the behaviour of both forms is clear from Figure 40: while temporal phrases are the most frequent context in the case of *estarse*, they are the least frequent with *estar*, suggesting that they favour the presence of the RM.

- (148) [labor de costura] Me estuve hasta las dos de la mañana y le
 REFL.1SG be.PST.1SG until the two of the morning and ACC.3SG
 llevé hecho.
 bring.PST.1SG make.PTCP
 ‘I stayed until two in the morning and I brought it [the needlework] finished.’
 (Villaverde–Mogina, Burgos, COSER 0959)

29. When both a locative phrase and either a temporal phrase or an attribute were present in the clause, I gave preference to the latter, assuming that the locative reading is the most basic one.

30. These 150 examples were selected randomly from all unmarked examples of *estar* in subcorpus E, using the function *sample_n* in R. One of them also showed an interruption, explaining why the total figures for Figure 40 amount to 149.

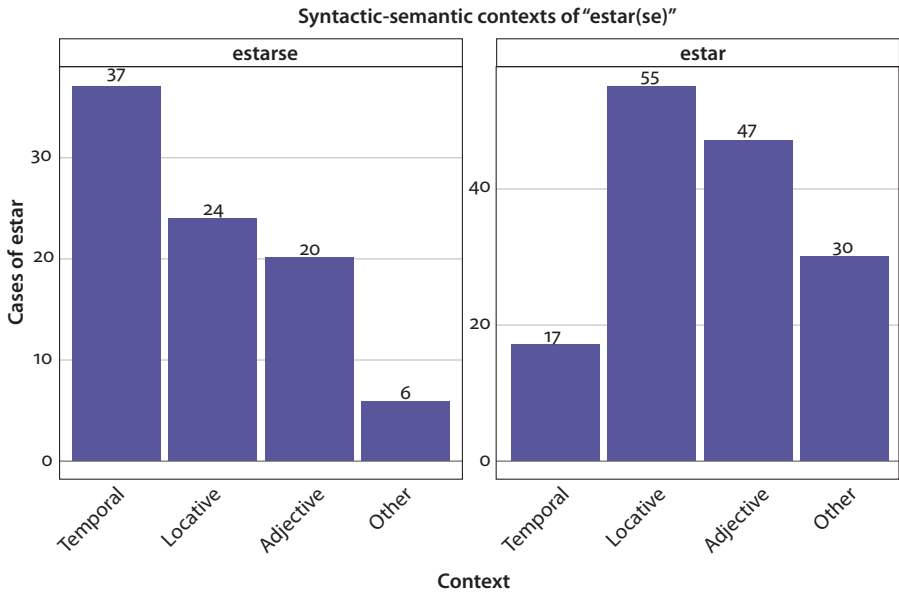


Figure 40. Total occurrences of *estar(se)* by syntactic context in the COSER data

- (149) Oye, no hagas fuerza, o estate quieta.
 hear.IMP.2SG not do.SBJV.2SG strenght or be.IMP.2SG-REFL.2SG still
 ‘Hey, don’t be tense or stay still.’ (Azcona, Navarra, COSER 3203)
- (150) Y si no, puedes estate en tu casa cosiendo.
 and if not can.PRS.2SG be.IMP.2SG-REFL.2SG in your house sewing
 ‘Otherwise you can stay at home sewing.’
 (Porzuna, Ciudad Real, COSER 1417)

The attributive predicates documented with *estarse* in the “*estarse* + adjective” structure provide further proof of the relationship between *estar(se)* and *quedar(se)*, since most of them are related to the notion of ‘stillness’ (see also Morimoto 2008). Actually, the most frequent adjective with *estarse* is *quieto* ‘still’ (45%, 9/20), etymologically related to *quedar*.³¹ Crucially, Bybee & Eddington (2006) identify *quieto* ‘still’ as one of the most frequent adjectives that appears with attributive *quedarse* with human subjects. In the COSER data, the most frequent attributive predicates that appear with *quedarse* with animate subjects (from a total of 60 attributive predicates in 108 different examples) are very similar to those documented by Bybee & Eddington (2006), namely, *solo* ‘alone’, *dormido* ‘asleep’, *parado* ‘still’, *soltero* ‘single,

31. The remaining adjectives are *callado* ‘silent’ (2), *tranquilo* ‘calm’ (1), *sentado* ‘sitting/seated’ (1), *acostado* ‘lying’ (1), *cogido del gancho* ‘grabbed by a hook’ (1), *muerto* ‘dead’ (1), *abrujado* ‘captivated’ (1), *seco* ‘dry’ (1), *cuajado* ‘clotted’ (1), and *cocido* ‘cooked’ (1). Note that many of these are still positions.

not married', and *viudo* 'widower' (see Figure 41).³² The unmarked form *quedar*, however, shows no stronger association with some attributive predicates than with others – it shows 20 different attributive predicates in 39 examples, with only *bien* 'well' appearing more than five times.³³ However, the same is not true for attributive *quedar(se)* with inanimate subjects, where the most frequent predicates are very similar across the marked and the unmarked form:³⁴

A further semantic connection between *estarse* and *quedarse* has to do with the notion of change of state. According to Morimoto (2008, see also De Miguel & Fernández Lagunilla 2000), *estarse* implies a previous change of state, contrary to *estar*. As a copulative verb, *quedar(se)* might convey a change of state and the marked form might have a change-of-state reading where the unmarked form lacks one, as illustrated in (151). Moreover, the notion of change of state is closely linked to anticausatives, as often noted in the literature (Schäfer 2008; Cennamo & Jezek 2011; Vivanco 2016 among others, see also Section 4.3). In all, I believe this to be an indication of the fact that *quedarse* might have been the source of the analogy for *estarse*.

- (151) a. Y el calzado también. Se le quedaba pequeño a uno,
and the shoes also REFL.3 DAT.3SG stay.PST.3SG small to one
pues a ver si a esta le está bueno.
well to see.INF if to this DAT.3SG be.PRS.3SG good
'And also the shoes. If they become too small for one of them, well, let's
see if they fit this one.' (Enguera, Valencia, COSER 4310)
- b. El calzado le quedaba pequeño
the shoes DAT.3SG stay.PST.3SG small
'The shoes were too small for him/her.'

32. I consider here the data for all areas together, in order not to reduce the sample further. Although the adverb *así* 'so' was always amongst the most frequent predicates attested with *quedar(se)*, I do not consider it here, since its lack of semantic content prevents us from making a useful semantic comparison. Figure 41 only shows attributive predicates that appear more than 5 times (and the ones that appear more frequently in the case of *quedar* with animate subjects).

33. The meanings of the other frequent predicates that appear with *quedar* with animate subjects are: *atado* 'tied up', *atontado* 'stunned', *colgado* 'hung / with no plans', *contento* 'happy', *de madre* 'with the function of a mother', and *mal* 'badly'.

34. For *quedar* with inanimate subjects, 69 different predicates in 182 examples were found. For *quedarse* with inanimate subjects, 50 different predicates in 118 examples were attested. The meanings of the most frequent predicates ($N > 5$) are: *blanco* 'white', *bien* 'well', *limpio* 'clean', *duro* 'hard', *seco* 'dry', *cuajado* 'clotted', and *frío* 'cold'.

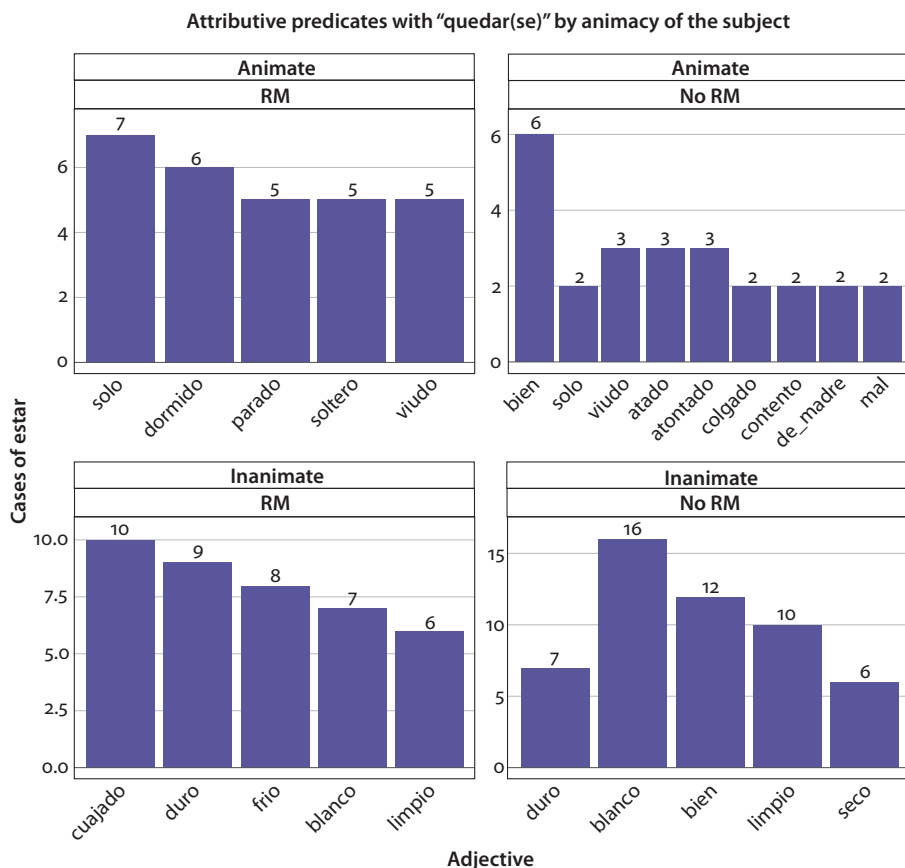


Figure 41. Frequent adjectives that appear with *quedar(se)* with inanimate subjects in the COSER data

As regards *esperar(se)* ‘to wait’, its semantic relationship with the notions of stillness found in both *quedar(se)* and *estarse* is rather clear. Moreover, the fact that *estarse* appears frequently with temporal phrases is another point of connection with *esperar(se)*, since temporal phrases frequently co-occur with this verb. The syntactic context in which *esperar(se)* is used also seems to be a relevant factor in the distribution of the RM. This verb may appear either in absolute contexts – with no complements (see (152)) –, with a prepositional or clausal complement expressing the event waited for (see (153)),³⁵ or with a temporal phrase indicating the amount

35. These sentential complements may or may not be preceded by the preposition *a*. I have considered all examples to be equally intransitive, since the loss of the preposition before a clause is not uncommon in (spoken) Spanish (cf. *queísmo*) and they are not replaceable by an accusative pronoun.

of time one should wait (see (154)). As shown in Figure 42, the absolute construction favours the use of *esperarse* the most. This could be due to the fact that these are the most obviously intransitive contexts and may suggest an influence of the de-objective function of the RM (see Section 3.2.3).

- (152)

Se ponía allí la gente esperando.
REFL.3 put.PST.3SG there the people waiting
'And people stood there, waiting.'
(Antequera, Málaga, COSER 3001)
- (153)

O te esperabas en la esquina a que bajase a
or REFL.2SG wait.PST.2SG in the corner to that go.down.PST.SBJV.3SG to
tu casa para ir al recaó.
your house for go.INF to.the errand
'Or you waited in the corner until she would go down to your house to run
some errands.'
(Porzuna, Ciudad Real, COSER 1417)
- (154)

La que me cortaron, por esperarnos hasta esa hora.
the that me cut.PST.3PL for wait.INF-REFL.1PL until that time
'They were very angry at me, that they had to wait until so late for us.'
(Moraleja de Cuéllar, Segovia, COSER 3706)

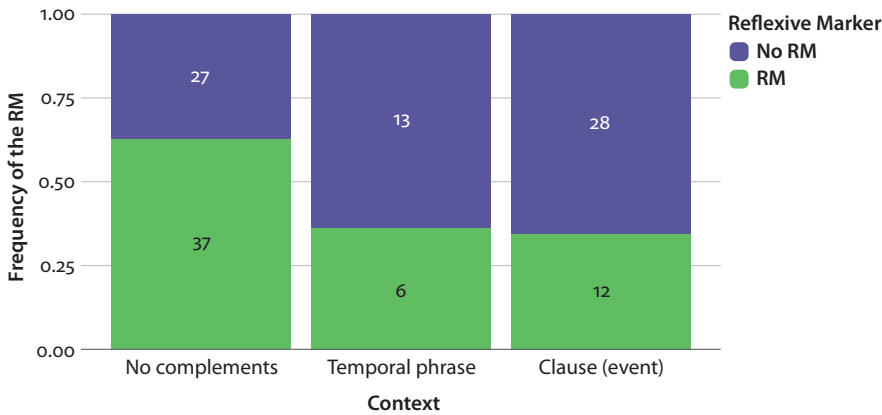


Figure 42. Frequency of the RM with *esperar(se)* by syntactic context in the COSER data

Interestingly, my data show that *esperarse* is more frequent when the verb is used in the imperative (see Table 32) than with other verb tenses. This again suggests a connection between the presence of the RM and higher agentivity or control, similar to the one proposed for *estar(se)* and *quedar(se)* above. Sadly, the relationship between the presence of the RM and the imperative cannot be checked for *quedar(se)*, because no examples of the imperative were found in my corpus.

Table 32. Frequency of the RM with *esperar* by verb tense in the COSER data

	Imperative	Other verb tenses
<i>esperar</i>	26 (40%)	43 (73%)
<i>esperarse</i>	39 (60%)	16 (27%)
Total	65 (100%)	59 (100%)

Fisher's exact test odds ratio estimate (1000 resamples):* Min.: 3.742, 1st Qu.: 5.717, Median: 6.341, Mean: 6.399, 3rd Qu.: 7.031, Max.: 10.823

* Because the observations of our total sample are not independent (i.e. many locations offer more than one example) and a generalised linear mixed model is not reliable (due to low figures in some cells), I have resampled the data 1,000 times (hierarchical bootstrap) and applied a Fisher's exact test to each of them. The measures given here correspond to the distribution of odds ratio estimates over these 1,000 samples. Values greater than one indicate a higher probability of positive cases (i.e. either unmarked non-finite verbs or marked finite verbs) than of negative cases (i.e. either marked non-finite verbs or unmarked finite verbs).

Since *esperar(se)* always takes animate subjects, we cannot check the statistical relevance of the animacy of the subject in the presence of the RM. However, this is a relevant factor determining the frequency of the RM with both readings of *quedar(se)* in both geographical areas, as Figure 43 shows.³⁶ Regarding *estar(se)*,

36. Although Bull (1950) does not mention animacy of the subject as a factor determining the presence of the RM with *quedar*, he states that “[t]he reflexive, clearly, is the sign of responsibility”, which suggests that it should be dispreferred with inanimate subjects, which are typically incapable of being responsible.

In a recent study, Van Gorp (2014) has proposed that pseudocopulative (i.e. attributive) *quedarse* presents a more subjective perspective compared to *quedar*, which favours a more objective perspective. Our data on attributive *quedar(se)* might explain how the association of *quedarse* with a more subjective perspective came about, since we expect animate subjects to favour a subjective perspective, while the opposite is true for inanimate subjects. It should be noted that Van Gorp's (2014) data show a vast majority of animate subjects (93.6% of her 800 examples). This huge difference with our data could be due to the fact that she uses a more restrictive concept of pseudocopulative *quedar(se)*, restricted to change-of-state events (and excluding stative interpretations, for example). The differences in our corpora (written novels vs. spoken interviews with rural speakers) may also have caused this discrepancy.

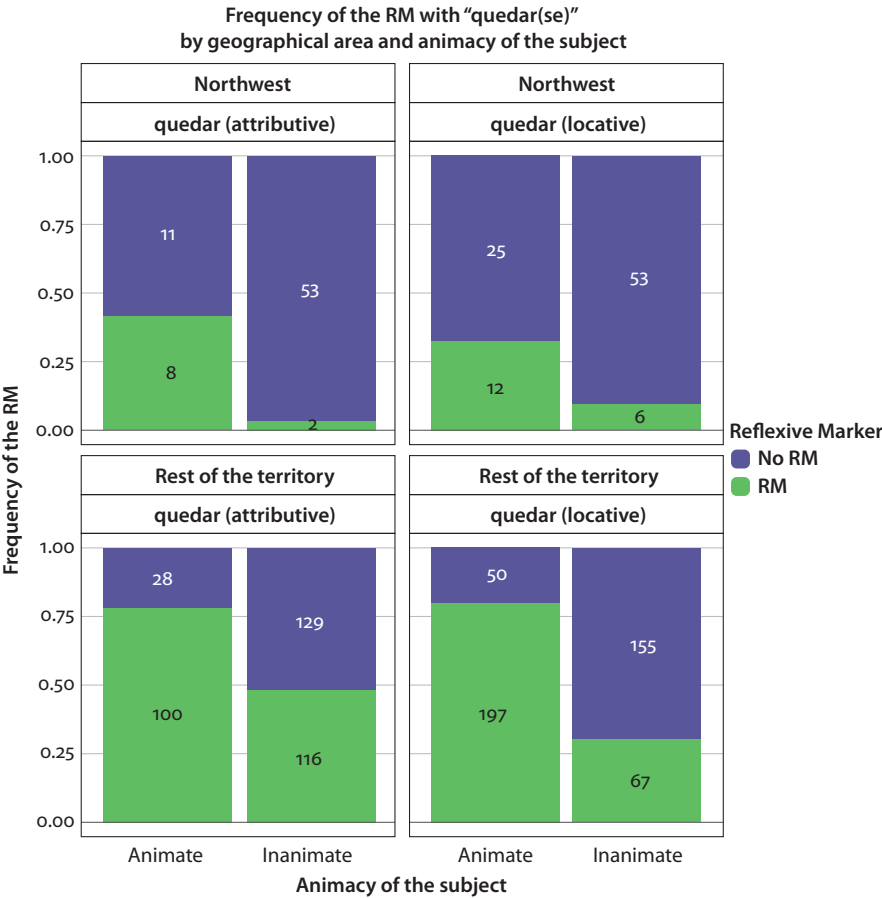


Figure 43. Frequency of the RM with *quedar(se)* by animacy of the subject in the COSER data

the data, although scant, are suggestive (see Table 33), since most of the marked examples occur with animate subjects. However, the frequency of *estarse* is so low that the effect of any parameter is rather small.

Table 33. Frequency of the RM with *estar* by animacy of the subject in the rest of the territory

	Animate	Inanimate
<i>estar</i>	1556 (98.5%)	955 (99.7%)
<i>estarse</i>	24 (1.5%)	3 (0.3%)
Total	1580 (100%)	958 (100%)

A generalised linear mixed model based on the data of the rest of the territory (see Table 34) confirms that having an animate subject favours the presence of the RM and also that differences by verb are quite strong.

Table 34. Generalised linear mixed model. Presence of the RM in the rest of the territory depending on subject animacy [reference level: Inanimate] (fixed factor) and the place and verb lexeme (random factors)

Effect	Term	Estimate	Std. error	Statistic	Conf. low	Conf. high
fixed	(Intercept)	-2.571	1.3	-1.977	-5.119	-0.023
fixed	Subject animacy: Animate	2.046	0.194	10.535	1.666	2.427
random (COSERID)	sd (Intercept)	1.117	NA	NA	NA	NA
random (Verb)	sd (Intercept)	2.57	NA	NA	NA	NA

In conclusion, these three verbs offer further proof of the analogical nature of the RM with intransitive verbs. Their similar meanings have caused the association with the RM in similar contexts, i.e. those of high agentivity, as shown by the examination of the examples and by the association of the RM with the imperative. The initial association with high agentivity was most likely produced through the connection between the RM and agentive subjects in autocausative verbs. Again, it is possible to propose that the productivity of the RM is older with some verbs. In this case it is clear that marked *quedarse* shows a higher productivity than *estar(se)* and *esperar(se)*, not only because its frequency is far higher, but because it is also found in north-western varieties. Thus, the attachment of the RM to *quedar* is likely to be older. The fact that *estarse* is modelled on an analogy with *quedarse* is quite clear due to its association with the adjective *quieto* and the general notion of “stillness”, which is inherent to *esperar(se)*.

5.8 Spontaneous *pasar(se)*

In this section I will briefly discuss the case of the spontaneous reading of intransitive *pasar(se)* ‘to happen, to go by’, which seems to be a paradigmatic case of the aspectual function of the RM (although it has not been considered in studies on the aspectual *se*, as far as I am aware).³⁷ Intuitively, *pasarse*, as opposed to *pasar*, seems to acquire the nuance ‘to be over’, clearly presenting a result state. In what follows I will try to

37. The translational motion readings of *pasar(se)* were considered in 5.4 above.

explain this behaviour by relating the distribution of the RM to spontaneous *pasar* to other meanings that seem to be associated with the RM.

The examples I discuss in this section were considered in Section 5.2 as the spontaneous reading of *pasar(se)*. This use is only found with inanimate subjects, typically with events, periods of time, or feelings/sicknesses, which in itself is surprising, since I have been highlighting the association of the RM with animate subjects. However, it should not be overlooked that spontaneous verbs are also frequently marked with the RM (see Section 3.3.3). The COSER interviews attest a total of 518 examples of spontaneous *pasar(se)*, of which only 42 show the RM (8.1%), that is, spontaneous *pasarse* is rather rare. Only two of these instances of *pasarse* are found in north-western varieties (2/46, 4.3%), so in the remainder of the section I will only address the rest of the territory, where 40 out of 472 (8.5%) occurrences of spontaneous *pasar(se)* show the RM.

When the subject is an event, the data, although scant, suggest that the RM is favoured in contexts meaning that something is over (see (155)). When this semantic nuance is present, the RM is preferred (7/13, 54%); when it is not, the unmarked variant is almost the only possibility (406/408, 99.5%).

- (155) Cuando se pasaron ya las fiestas.
 when REFL.3 pass.PST.3PL already the fest
 ‘When the town festival was over.’ (Cabra del Santo Cristo, Jaén, COSER 2303)

With subjects that express feelings, emotions or sickness, the meaning of *pasar(se)* is typically ‘to disappear’. My sample attests only nine examples with this kind of subjects, all of them marked. Eight mean ‘to disappear’ (see (156a)), while (156b) does not.

- (156) a. A lo mejor si le había dado a alguno mala
 to the best if DAT.3SG have.PST.3SG give.PTCP to someone bad
 gana, el mirar pa que se le pasara y
 wish the look.INF for that REFL.3 DAT.3SG pass.SBJV.3SG and
 darle agua.
 give.INF-DAT.3SG water
 ‘If maybe someone was feeling bad, try to make it go away and give them water.’ (Fuentes Claras, Teruel, COSER 4117)
- b. Venta, que está ardiendo, la vecina me llamó, uy, que
 Venta that is burning the neighbour REFL.1SG call.PST.3SG oh that
se me pasa el escalofrío.
 REFL.3 DAT.1SG pass.PRS.3SG the shiver
 ‘Venta [the house’s name], it is burning, my neighbour had called, oh my god, I get the shivers.’ (Leitza, Navarra, COSER 3214)

When the subject is food, *pasar(se)* means ‘to rot, to decay, to go bad’, a meaning clearly related to the notion of being over: the time of eating the food has passed, is over. In these contexts, *pasar(se)* is seemingly always marked, although I have very few examples ($N = 6$):

- (157) Y en ese tiempo no se pasaba, no había tiempo pa
 and in that time not REFL.3 pass.PST3SG not have.PST.3SG time for
pasarse la leche.
 pass.INF-REFL.3 the milk
 ‘And in that period of time it wouldn’t go bad, there was not enough time so
 that the milk would go bad.’ (Oliván, Huesca, COSER 2222)

When the subject is a period of time, the difference between ‘happen, go by’ and ‘be over’ is blurred. Both the marked and the unmarked variants of *pasar(se)* can appear in this context (see (158)), with similar frequencies: the RM appears 44% of the time (16/36).

- (158) a. Cuando se pasen equis años, se van a quedar
 when REFL.3 pass.PRS.3PL x years REFL.3 go.PRS.3PL to stay.INF
 extranjeros, españoles no, si es que no paren las
 foreigners spaniards no if is that not give.birth.PRS.3PL the
 mujeres.
 women
 ‘When x years go by, there will only be foreigners, no Spaniards, women
 don’t give birth anymore!’ (Povedilla, Albacete, COSER 0222)
- b. De que panan dos días no hay quien coma
 of that pass.PRS.3PL two days not have.PRS.3SG who eat.PRS.3SG
 este pan de ahora.
 this bread of now
 ‘After two days, you can’t eat the bread they bake today.’
 (Lucillo, León, COSER 2644)

Is the frequency of the RM in contexts where a final result is implied so high because this is the meaning with which the RM is associated? Or is it the product of the association of the RM with other characteristics, which typically occur in instances where *pasar* encodes a final limit? In what follows I will argue for the latter option. First of all, let us look at the frequency of the RM depending on the semantic characteristics of the subject. Despite the scarcity of the data, Figure 44 shows that the RM seems to be especially frequent with subjects depicting either feelings or sicknesses, or food. Event subjects clearly seem to be the ones that least favour the presence of the RM, while temporal subjects are somewhere in the middle.

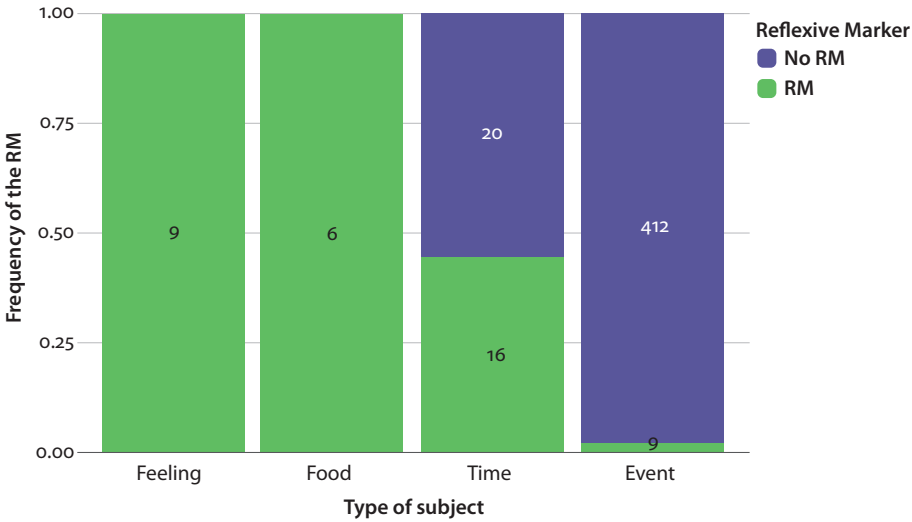


Figure 44. Frequency of the RM with spontaneous *pasar(se)* by the semantics of the subject

This distribution can be connected to two characteristics of *pasar(se)*. First, with the meaning ‘to be over’, it clearly depicts a change of state, as most anticausatives do. As noted above, this meaning is most frequent with feelings and food and least frequent with events. Moreover, a large number of anticausative verbs refer to food going bad, such as *pudrir(se)*, *enranciar(se)* ‘to rot, to go bad’, *agriar(se)* ‘to go sour’, which provide a source of analogy. Second, it is also in these cases that an affected external human participant is most easily interpreted. In fact, from nine examples of feeling subjects, eight show a dative participant encoding an affected human participant.

That is, while spontaneous *pasarse* seems in principle a good example of the aspectual function of the RM, because it seems to focus on the culminative phase of the event, the association of the RM with the meaning ‘to be over’ can be explained as the result of the association of the RM and both the notion of change of state and the affectedness of an external participant.

5.9 More sporadic analogies

There are a number of intransitive reflexive verbs that are documented only infrequently in the COSER interviews, but which can also be explained on the grounds of analogy, be it with verbs with a close semantic meaning or with verbs in similar syntactic constructions.³⁸ Many of these verbs belong in Kemmer's (1993) semantic classes, which can be seen as further evidence of the influence of analogy on the basis of semantic similarity.

First, several intransitive verbs that express spontaneous events are marked with the RM in the COSER interviews. Some of them are rather sporadic and often the source of the analogy is easy to guess: (1) *arderse* 'to burn' is close to the anticausative verb *quemarse* 'to burn'; (2) *(en)canecerse* 'to rot' and *mostearse* 'to turn into grape juice/must' have similar meanings to anticausative *podrir(se)*, *enranciar(se)* 'to rot, to go bad', *agriar(se)* 'to go sour' and other related verbs; (3) *escullarse* 'to drain' is a synonym for anticausative *escurrirse* 'to drain'; (4) *nacerse* 'to sprout' is related to anticausative verbs that encode the life processes of plants, such as *criarse* 'to grow', *marchitarse* 'to wither', *troncharse* 'to snap off'; (5) *crecerse* 'to grow up' and *vivirse* 'to live (somewhere)', both referring to human subjects, might have been influenced by other verbs that convey spontaneous life events, like anticausative *criarse* 'to grow up' or intransitive *morirse* 'to die'.³⁹

Second, some of these verbs can be associated with the uses examined in this chapter. The following belong in Kemmer's classes: (1) *sonreírse* 'to smile', which was documented once in the COSER data and several times in the questionnaire (video 184 was explicitly designed for eliciting it), is undoubtedly related to *reírse* 'to laugh'; (2) *resbalarse*, *esbararse* 'to slip' and *tropezarse* 'to trip' are clearly associated with *caerse* 'to fall down';⁴⁰ (3) *saltar(se)* and *brincar(se)* 'to jump' are motion verbs where the presence of the RM seems to be related to the idea of an event that is not supposed to happen, such as break-ins (see De Benito Moreno 2015: 7.3.1.3) – this idea is connected to the contexts associated with other motion verbs such as *salirse* 'to go out', *escaparse* 'to escape' and *pasarse* 'to pass' (see Section 5.4); (4) *orinarse* and *pisarse* (seemingly a borrowing from the Catalan verb *pixar*) 'to pee' are synonyms for *mearse* 'to take a piss' (see Section 5.6); (5) a large number of translational

38. For the sake of simplicity, I do not discuss all marked examples of intransitive verbs in the COSER data here. All of these were discussed in De Benito Moreno (2015: 7.3.1).

39. De la Mora (2011) also includes examples of marked *crecer(se)* and *nacer(se)* with human subjects.

40. The questionnaire data show that the marked forms are preferred with *resbalar(se)* and *esbarar(se)* and are possible but not frequent with *tropezar(se)* (see De Benito Moreno 2015: 7.3.1.6.4). This depicts the already familiar pattern where verbs with closely related meanings take the RM with different frequencies, suggesting that lexical diffusion is at play.

motion intransitive verbs were documented in the reflexive form just a few times, such as *acudirse* ‘to go’, *andarse* ‘to walk’,⁴¹ *correrse* ‘to run’, *descenderse* ‘to go down’, *inmigrarse* ‘to immigrate’, *volarse* ‘to fly’, *volverse* ‘to go back’,⁴² *llegarse* ‘to arrive’, *entrarse* ‘to go in’. *Llegarse* and *entrarse* are especially noteworthy, since several authors have claimed that they are ungrammatical (De Miguel & Fernández Lagunilla 2000; Sánchez López 2002). While De Miguel & Fernández Lagunilla (2000) acknowledge that both verbs are possible in some varieties, they propose that they have slightly different meanings (*llegarse* being closer to *acercarse* ‘to get closer’, and *entrarse* closer to *meterse* ‘to get in’). This, however, is not true for most of the examples found in the COSER data (see De Benito Moreno 2021, 2015: 7.3.1.2, 7.3.1.2.6).⁴³

Third, the RM appears with a number of intransitive verbs which do not belong in Kemmer’s (1993) semantic classes but that are still reminiscent of other uses of the RM: (1) *tardarse* ‘to take time’ is semantically similar to anticausative *retrasarse* ‘to be late’; (2) *abusarse (del poder)* ‘to abuse one’s power’, *dejarse (de fumar)* ‘to stop smoking’, *desertarse (de la mili)* ‘to desert military service’ are reminiscent of antipassive reflexive structures, since they take a prepositional object; (3) *cenarse* ‘to have dinner’ is reminiscent of transitive verbs of consumption (see Section 6.5).

Fourth, some very sporadic uses, possibly influenced by the existence of other reflexive uses of the same verb, were found. For instance, examples of a stative use of *salirse* and one grooming example of *irse* (see (159) and (160)). These examples might be seen as evidence of how the extension of the RM to new contexts might start, by occupying new positions that are related to older ones, since they appear with a verb that might already take the RM.

- (159) Echaba yo un poco azúcar, un gustico de azúcar quemao y leche
 add.PST.1SG I a little sugar a taste of sugar burn.PTCP and milk
 quemada y así, un gusto muy rico se le salía.
 burn.PTCP and so a taste very fine REFL.3 DAT.3SG go.out.PST.3SG
 ‘I used to add a bit of sugar, a bit of burnt sugar and burnt milk and so, that
 gave it a very fine taste.’ (Leitza, Navarra, COSER 3214)

41. The use of *andarse* meaning ‘irse, marcharse’ in Jérica (Castellón, COSER 1307) may be due to the influence of the Catalan form *anar-se*, but note that the speaker that produced that example was a Spanish monolingual.

42. Note that there is also an anticausative *volver(se)* meaning ‘to turn around’. For more details on the behaviour of intransitive *volverse*, see De Benito Moreno (2015: 7.3.1.2.5).

43. Because *entrar* might be used transitively in some varieties and hence enter the anticausative alternation, I checked whether the presence of transitive *entrar* triggered a higher frequency of the RM. The COSER data show that there is only a partial overlap between the area that shows causative *entrar* and the area that shows *entrarse* (see De Benito Moreno 2015: 7.3.1.2.6).

- (160) [¿Cómo eran las representaciones teatrales?] Había un salón grande
 have.PST.3SG a room big
 ahí abajo y nos íbamos vestidos de pastoras, los
 there down and REFL.1PL go.PST.1PL dress.PTCP of shepherdesses the
 pastores y to eso.
 shepherds and all that
 '[How were the theater plays?] There was a large theatre down there and we
 were dressed as shepherdesses, shepherds, that kind of thing.'
 (Bacares, Almería, COSER 0404)

From these examples it can be seen that the presence of the RM with intransitive verbs shows a lot of variation. Trying to unify all the uses of the RM in one single syntactic-semantic function is most likely impossible, for its behaviour often seems to respond to very specific analogies and its productivity is continuously expanding. This is why the analysis of corpus data will almost always offer counterexamples to aspectual accounts of the RM, which assign one single function to the RM (see also de Benito Moreno 2021). Besides the examples mentioned above, the COSER data offer examples of *existirse* 'to exist', which Sanz & Laka (2002) rule out for being atelic, and of *trabajarse* 'to work', which should not take the RM according to Rigau's (1994) analysis, as it does not have an internal argument. This is why I believe that an account that relies less heavily on a single function of the RM has greater chances of being successful.

5.10 Conclusion

In this chapter I have argued that reflexive intransitive verbs have an analogical origin and that they have arisen from previous functions of the RM. That is, I reject analyses of the RM which propose that this particle shows a clear-cut syntactic function that attaches productively to a number of predicates that share certain characteristics (see also de Benito Moreno 2021). On the contrary, I believe that the RM has increased its productivity and functional range gradually by lexical diffusion (and keeps doing so). According to this hypothesis, the spread of the RM to new contexts is motivated by semantic similarity between the contexts where it used to appear and the new ones.

The advantage of this proposal is that it can explain the heterogeneity of the class, which is seen in at least three aspects. First, the absence of a single meaning that applies to all intransitive verbs that might take the RM is explained by the fact that there are several source constructions for the analogy and that the RM can absorb different semantic values from these. Second, the fact that the alternation

between marked and unmarked forms *sometimes* shows meaning differences and *sometimes* does not fits a gradual extension of the RM. Older contexts have routinised these semantic contrasts, while newer contexts show interchangeability of both forms. Third, the lack of a unitary dialectal distribution is expected if different analogies take place at different places and at different times. On the one hand, my data show that north-western varieties are the least favourable for these analogical uses, which is to be expected, since the productivity of the RM is lower in these varieties. That is, the number of potential source constructions for the analogical extension is reduced in these varieties. In the rest of the territory, however, some reflexive intransitive verbs do show geographically conditioned distributions, but these do not necessarily overlap. For instance, *vivirse* ‘to live’ seems to be a regional use of northeastern varieties: the two examples in my sample were found in La Rioja and Zaragoza, and it also exists in the area of Ribaforada (Navarra) (Ekaitz Santazilia, p.c.). *Entrarse* ‘to go in’ seems to be especially frequent in southwestern varieties, while *volverse* ‘to go back’ is found both in southern and in eastern varieties (see Map 15). If we were to include American varieties, this variation will only be larger. *Tardarse* ‘to take time’ was documented only once in my data, but is very frequent in Mexico (and probably other American countries). The contrasts offered by Maldonado, who is Mexican, often show those diatopic differences: while for him unmarked *tropezar* is almost unacceptable (see (161)), my questionnaire informants almost always prefer the unmarked variant in the very same context. In my opinion, this heterogeneity speaks of a construction in constant lexical diffusion.



Map 15. Geographical distribution of *entrarse* and *volverse* in the COSER data

- (161) Mariangélica ??(se) tropezó con una piedra.
Mariangélica REFL.3 stumble.PST.3SG with a stone
‘Mariangélica stumbled with a stone.’ (Maldonado 1989: 354)

In order to support the analogical hypothesis, I have highlighted a number of semantic similarities between reflexive intransitive verbs and the previous functions of the RM. First, I have noticed that most intransitive verbs that take the RM belong to the same semantic classes as most anticausatives (translational motion verbs and spontaneous events, see Section 5.2). Translational motion verbs constitute the most relevant class here, both in terms of productivity and number of verbs that enter the alternation, which is consistent with the fact that they are also the eldest class attested (Bogard 2006). In other cases, the extension of the RM can be associated with its oldest functions, namely the reflexive (see Section 5.6) or the reciprocal (see Section 5.2).

Second, the fact that the RM with intransitive verbs is frequently associated with animate subjects is also reminiscent of the restrictions that apply to older functions of the RM, since reflexive constructions are only possible with animate subjects, and anticausatives show the same association, no doubt due to the same historical explanation. This does not mean that the RM is impossible with intransitive verbs with inanimate subjects – it only means that its productivity is lower in these contexts. In some cases the presence of the RM is also associated with the presence of an affected dative, again paralleling the situation of marked anticausative verbs (see Section 4.5).

Finally, I have tried to explain the most conspicuous meaning contrasts between marked and unmarked verbs by making a connection between them and some of the most relevant meanings associated with the RM in its previous functions. These meanings are mostly associated either with the role of the subject or with the role of an external causer/participant. This double possibility of the RM explains its contradictory semantic associations. On the one hand, in anticausatives with inanimate subjects, the RM is often associated with situations that have an external cause, which can also be an affected participant – especially when causation is unintentional. This explains the many cases of reflexive intransitive verbs where the RM is associated with events that go against physical or social conventions or are “unexpected” (see Sections 5.4, 5.5 and 5.8), since these normally affect an external participant. On the other hand, with some intransitivised verbs with human subjects (emotion conversive verbs and some anticausatives), the RM is associated with a reduction in the subject’s control, because they reject external agents (see 3.2.1). This explains the behaviour of the RM with body processes analysed in Section 5.6, since they also allow for both interpretations (with and without control) and the RM is more strongly associated with the latter. Finally, the RM seems also to be associated with the opposite notion, i.e. a subject with increased control. This can also be explained by the previous functions of the RM, since this particle is also associated with (both animate and inanimate) subjects that initiate spontaneous

processes on their own. This higher control of the subject in marked stative verbs (see Section 5.7) is also related to the animacy of the subject, as expected by the fact that anticausatives with human subjects are more consistently marked and that these are the subjects with higher control potential.

One of the risks of proposing analogical explanations is that they might appear to be too weak, since they can be used to explain even seemingly contradictory functions with the same origin. This is why I have tried to support the connections between older and newer functions of the RM with quantitative data. But, most importantly, my analogical explanation is intended to produce testable historical hypothesis, which must be confirmed by a careful diachronic analysis that supports the notion that the sources of the analogy are older than the analogical uses.

As said above, a consequence of this analogical explanation is that the RM is not to be seen as a particle with a clearly defined syntactic-semantic function. On the contrary, its distribution across the lexicon and across different contexts of the same lexical item show some degree of variability and idiosyncrasy, also showing different degrees of lexicalisation. Variationist approaches such as this one and those recently undertaken by a number of Mexican authors (Aaron 2003, 2004; Aaron & Torres Cacoullós 2005; Torres Cacoullós & Schwenker 2008; Merino Hernández 2019), are especially suited to uncovering the different contexts that favour the presence of the RM. However, the regularities discussed in this chapter also show that the presence of the RM cannot be studied for each lexical item individually, since groups of items with similar behaviours can be identified. In this sense, a constructional analysis based on similarity (see Bybee & Eddington 2006) is especially fruitful.

The fact that the RM might attach not only to single verbs, but also to more specific constructions, was already mentioned in Section 5.6, where the constructional productivity of the combination of the RM with the adverb *encima* 'above' was discussed. The COSER data show other such examples. For instance, the RM shows some productivity in constructions like that in (162), where the subject is a source and the experiencer is coded as a dative. Semantically, the construction conveys an emotion or feeling, but its syntactic structure recalls spontaneous events (which have a theme subject and might take an experiencer dative). Interestingly, it is not only intransitive verbs such as *apetecer* 'to feel like doing something' or *gustar* 'to like' that can appear in this construction (always unmarked in Standard Spanish), but also transitive verbs such as *desear* 'to wish', which undergoes a conversive change of diathesis to become available for the construction, which requires an intransitive verb. This situation is thus very similar to the one described in 5.6 above, where the reflexive construction imposes a predetermined syntactic behaviour of the verb.

- (162) a. Había un señor a quien... se le gustaba eso.
 have.PST.3SG a man to who REFL.3 DAT.3SG like.PST.3SG that
 ‘There was a man who liked that.’
 (Villalba de Lampreana, Zamora, COSER 4611)
- b. Ahora mismo sólo se me apetece estar con Javi
 now same only REFL.3 DAT.1SG feel.like.PRS.3SG be.INF with Javi
 y con Ana.
 and with Ana
 ‘Right now I only want to be with Javi and Ana.’
 (July 15th 2015, Huelva, Andalucía. Tweet.)
- c. Lo que pasa es que a él se le deseó
 the that pass.PRS.3SG is that to OBL.3SG REFL.3 DAT.3SG wish.PST.3SG
venirse.
 come.INF-REFL.3
 ‘The thing is he wanted to come.’
 (La Atalaya (Villanueva de las Algaidas), Málaga, COSER 3002)

Reflexive transitive verbs

6.1 Introduction

This chapter focusses on the presence of the RM with transitive verbs (reflexive transitive verbs henceforth), following the same hypothesis outlined in Chapter 5: since an aspectual function is too distant in semantic terms from the older functions of the RM, I hypothesise that the presence of the RM is triggered by the association of the RM with semantic features typically found in previous functions of the RM, and that the aspectual constraints found in some examples are a by-product of these factors.

Note, however, that we do not necessarily expect the source of the analogy to be the same for both intransitive and transitive verbs. Since intransitive verbs encode only one verbal participant, we expect reflexive intransitive verbs to have been influenced by anticausative, conversive, de-objective, antipassive verbs or even from direct reciprocal (or reflexive) constructions, which also have a single participant. On the other hand, transitive verbs have at least two participants and their reflexive version is more likely to be connected to transitive reflexive constructions, such as indirect reflexive (and reciprocal) constructions. We thus expect these to be the source of the analogy of the RM that appears with transitive verbs – especially the possessive and benefactive reflexive dative, which are more productive than the recipient reflexive dative (see Section 6.2). This idea is not new and has been explicitly defended by Maldonado (1999, 2000). It is also implied in Rigau's (1994) proposal, where she claims that this RM is a beneficiary (Rigau 1994). However, the COSER data show that the situation is not so simple and that some reflexive transitive verbs have been crucially influenced by reflexive intransitive verbs. I will analyse the different kinds of reflexive transitive verbs and their connection with other reflexive uses in Subsections 6.3 to 6.11. Prior to this, in 6.2 below I will briefly review (some) reflexive indirect constructions, so as to describe the characteristics of the alleged sources of analogy.

6.2 Indirect reflexive constructions

The indirect reflexive label is modelled on the label of the direct reflexive. While in the latter the key concept is coreferentiality between the subject and the direct object, in the former coreferentiality is found between the subject and the indirect object. The identification of different indirect reflexive constructions is subject to debate in the literature, as seems to be the case with the classification of all reflexive constructions. Kemmer defines the prototypical indirect reflexive as follows:

The prototypical **indirect reflexive** situation type can be defined as comprising situations in which (a) there are three participants, an Agent, a Patient, and a Recipient or Beneficiary; (b) the Agent and the Recipient/Beneficiary are coreferential; and (c) the basic verbal event is one in which the Agent and Recipient/Beneficiary normally are distinct entities. (Kemmer 1993: 74, emphasis in original)¹

Kemmer also proposes an indirect middle category, paralleling the difference between direct reflexive and body action middles. That is, according to her approach, the indirect middle “comprises actions that one *normally* or *necessarily* performs for one’s own benefit” (Kemmer 1993: 78). Thus, the semantic difference between the indirect reflexive and the indirect middle can be stated on the same grounds as the difference between the direct reflexive and the body action middle, i.e. the relative distinguishability of the participants in the event:

[T]he indirect reflexive involves more conceptual separation of the Initiator and Endpoint entities than the indirect middle. The indirect middle lacks the expectation of non-coreference that is part of the meanings of ordinary verbs of three-participant semantics. (Kemmer 1993: 81)

She observes, however, that it is not easy to differentiate between these two types, since there is a continuum between situations in which the endpoint is usually the same participant as the initiator and those in which they are typically different entities. I agree and will hence approach the different types of indirect reflexive

1. In Kemmer’s data, most languages appear to mark both direct and indirect reflexives with some form of the RM. However, she notes that many grammars lack information on the indirect reflexive and wonders whether this may be due to the fact that it is so obviously reflexive that it is not considered to be especially relevant. Nevertheless, the case may well be completely the opposite. Galician and Asturian grammars tend to explicitly mention the lack of indirect reflexive constructions in these languages because they are often modelled on grammars of Spanish, a language with a lively indirect reflexive construction. It seems likely, however, that grammars tend not to mention whatever categories do not exist in the language and focus rather on those that do actually exist.

using a classification that is not so open to subjective interpretations, based on the semantic role of the indirect object, partially following Geniušienė (1987).

Geniušienė (1987), who calls the indirect reflexive “dative transitive reflexive”, categorises examples of this construction according to the thematic role of the indirect object. She distinguishes three different possibilities, namely, the indirect object is a recipient, a possessor or a benefactive.² In fact, these three categories can be placed on the continuum between indirect reflexive and indirect middle hypothesised by Kemmer. On the one hand, recipients are at the reflexive end of the continuum, since transfer actions are not normally conducted on oneself. On the other hand, possessors should be at the other end, closer to the indirect middle, since possessive datives are particularly common with body parts (König & Haspelmath 1998; Linzen 2009; Lamiroy 2003) and actions involving body parts are more frequently conducted on oneself. Beneficiaries, however, should be somewhere in the middle of the continuum, since actions like buying or cooking are quite frequently self-directed but also other-directed – sometimes they are even directed at oneself and others at the same time.

If both Kemmer’s proposal and the continuum I have just proposed are right, we would expect that verbs at the reflexive end of the continuum (recipient datives) would present more reflexive marking than those in the centre of the continuum (beneficiaries), and that those at the middle end of the continuum (possessive datives) would show the littlest reflexive marking. However, something else has to be taken into account: the syntactic status of each of these datives.

In the varieties under study, recipients are verbal arguments (Gutiérrez Ordóñez 1999), which means that they need to be encoded within a sentence by means of a prepositional phrase or an unstressed dative pronoun. Possessive datives, although they respond positively to both tests of argumentality and non-argumentality (see Gutiérrez Ordóñez 1999), are also compulsory in the external possession construction (see Section 6.2.3). Beneficiaries, on the contrary, are adjuncts and hence optional – they do not need to be explicitly encoded in the sentence (Gutiérrez Ordóñez 1999). In this study I used the presence/absence of the RM as a parallel of the heavy/light marking described in Kemmer’s (1993) work. However, benefactive datives cannot be measured with the same scale as recipient and possessive datives,

2. Note that Kemmer’s definition of the indirect reflexive did not include possessors as a possible thematic role of the indirect object. This may have been due to the fact that possessive datives seem to be largely restricted to European languages cross-linguistically (König / Haspelmath 1998). She also wonders whether the different thematic roles of the indirect object would make a difference concerning reflexive and middle marking, but she has little available data to test this in her sample.

since they are not verbal arguments and can hence remain unexpressed. This means that we can only make a fair comparison between recipients and possessive datives.

Distinguishing indirect reflexive types according to the thematic role of the dative is not common in work on reflexives in Asturian, Galician or Spanish. However, it is commonly noticed that Galician and Asturian are remarkably permissive with the absence of indirect reflexive pronouns, as compared to Spanish (see also Section 1.2.1). Most authors (Academia de la Llingua Asturiana 2001; D'Andrés 1994; Dubert García 1999; Freixeiro Mato 2002; Álvarez & Xove 2002) agree that the indirect reflexive pronoun is only possible when there is an absolute need to clarify that the subject and the dative are coreferential. In D'Andrés (1994), examples of cases where coreference is not obtained by default in Asturian – and are thus marked – show recipient datives, as opposed to possessive datives, where coreference is interpreted by default and reflexive marking is lacking.

- (163) O médico {diagnosticouse /diagnosticoulle} unha
the doctor diagnose.PST.3SG-REFL.3 diagnose.PST.3SG-DAT.3SG a
hepatite.
hepatitis
'The doctor diagnosed {himself/herself ~ him/her} a hepatitis.'
(Galician, Freixeiro Mato 2002: 155)
- (164) (Asturian, D'Andrés 1994)
- a. Xuan sírvese cerveces.
John serve.PRS.3SG-REFL.3 beers
'John serves himself a beer.'
- b. Enxagües les manes.
rinse.PRS.2SG the hands
'You rinse your hands.'

Before we turn to the study of each specific type, it should be noted that indirect situations do not fit nicely into the approaches that treat the middle marker as a valency reducing device (as Kemmer 1993 already noted), since in these the valency of the verb is kept stable.

In the following Subsections I analyse the COSER data for the presence of the RM with recipient datives (Section 6.2.1), beneficiary datives (Section 6.2.2), possessive datives (Section 6.2.3). Finally, I will discuss some indirect constructions where the subject seems to undergo demotion in Section 6.2.4.

6.2.1 Recipient datives

Recipient datives are verbal arguments that typically appear with verbs of transfer, speech or perception (Gutiérrez Ordóñez 1999), as illustrated in (165). Corpus data show reflexive constructions not to be very frequent with datives of this kind (De Benito Moreno 2015), as expected by the fact that transfer actions typically occur from one participant to another. Despite their scarcity, they seem to be overwhelmingly marked with the RM, with only one unmarked case from 29 examples (3%) in the COSER data, and even here it can actually be explained by the syntactic properties of the example, since instead of taking a dative it uses the prepositional phrase *entre mí* ‘within me’ (De Benito Moreno 2015). With speech verbs, the reflexive may add a slight change of meaning, turning the event into a cognition event, as in (165b). This is probably because speech directed at oneself can be seen as a variant of thinking. Actually, Maldonado (1999) considers that the attachment of the RM to audible speech events produces reflexive constructions, whereas the presence of the RM in internal or mental speech acts produces middle constructions.

- (165) a. Yo me medicinaba así el ocalito
 I REFL.1SG give.as.medicine.PST.1SG so the eucalyptus
 ‘I gave myself the eucalyptus as medicine.’
 (Almadén de la Plata, Sevilla, COSER 3806)
- b. Y no sé si tendría hijos son cosas que a
 and not know.PRS.1SG if have.COND.1SG sons are things that to
 veces te preguntas.
 times REFL.2SG ask.PRS.2SG
 ‘And I don’t know if I had kids, those are the things you sometimes wonder.’
 (Aguinaga, Guipúzcoa, COSER 2001)

The relationship between the indirect reflexive construction with recipient datives and a few reflexive transitive verbs that do not admit the prepositional reinforcement *a sí mismo*, like *explicarse* ‘to understand, lit. to explain-REFL’, is very clear and suggests that these reflexive transitive verbs are modelled on such reflexive constructions. The fact that *explicarse* is favoured in negative contexts (all four examples attested in the COSER data are negated) confirms that the semantically reflexive interpretation is not appropriate:

- (166) [desahucios] Pero yo, ¿lo que hay ahora? Yo no me
 but I the that have.PRS.3SG now I not REFL.1SG
explico...
 explain.PRS.1SG
 ‘[evictions] But that what’s happening now? I don’t get it.’
 (Tablones, Granada, COSER 1834)

6.2.2 Autobenefactives

Beneficiary participants do not allow for such a clear-cut definition as recipient participants. The working definition of the beneficiary role that Kittilä & Zúñiga (2010) provide in their volume on benefactives reads as follows:

The beneficiary is a participant that is advantageously affected by an event without being its obligatory participant (either agent or primary target, i.e. patient). Since normally only animate participants are capable of making use of the benefit bestowed upon them, beneficiaries are typically animate. (Kittilä & Zúñiga 2010: 2)

However, it is not always easy to differentiate recipients of autobenefactives on semantic grounds, since benefaction frequently implies receiving. In order to retain the distinction, I have considered as recipient datives only those that appear with transfer and speech verbs, which require a third argument. It can also be difficult to draw the line between autobenefactive constructions and transitive verbs marked with the RM, since subjects might be seen as natural beneficiaries of at least some actions. In the former I have included constructions where the RM alternates with a non-referential pronoun which can be clearly interpreted as a beneficiary, whereas I subsume the remaining cases under the latter category (reflexive transitive verbs).

As noted above (Section 6.2), beneficiaries are not verbal arguments, which means that they are optionally encoded in the sentence – even if they are not coreferential with the subject. This is seen clearly in (167), where the beneficiary dative is present only in some of the very similar contexts where it is coreferential with the subject.

- (167) [su vestido de novia] Me lo hizo a mí la
 DAT.1SG ACC.3SG make.PST.3SG to OBL.1SG the
 modista. Tanto mi hermana como yo nos hicimos
 dressmaker so.much my sister like I REFL.1PL make.PST.1PL
 camisones de dormí... aquí en casa. Ella nos lo cortaba
 nightshirts of sleep.INF here in house she DAT.1PL ACC.3SG cut.PST.3SG
 y nosotros lo hacíamos. Las cosas así un poquito más
 and we ACC.3SG make.PST.1PL the things so a little more
 delicás pos las hacía ella.
 delicate well ACC.3PL make.PST.3SG she
 '[her wedding dress] The dressmaker made it for me. Both my sister and I made
 nightdresses for ourselves, here at home. She used to cut the fabric for us and
 we made them... The things that were a bit more delicate, she made them.'
- (Valencia del Ventoso, Badajoz, COSER 0728)

Typological evidence about benefactive constructions shows that the expression of self-benefaction (as opposed to non-self-benefaction) is somehow secondary, since there are no languages that have self-benefactive constructions in absence of non-self-benefactive constructions (Smith 2010). This indicates that the agent is understood as the beneficiary of the action (when there is one) *by default*. That is, the semantically most common situation (agent = beneficiary) is unmarked, while less common situations (agent \neq beneficiary) are marked.³ This suggests that when an autobenefactive is explicitly encoded in the sentence it might have some additional meaning to the regular benefactive construction, most likely in terms of emphasis. On the nature of this additional meaning, Boyd (2010) notes:

The semantics of benefactivity in this use [the reflexive] is much more difficult to define than in the case of benefactive arguments in the other senses. This is unavoidable, as the nature of the semantic roles of arguments is to a large extent induced by the governing predicative term, usually a verb. Benefactivity as an additional feature of another role is not defined by the predicate, but depends greatly on pragmatic factors. (Boyd 2010: 332)

The existence of an emphatic nuance due to the higher involvement of the subject caused by a greater difficulty in the verbal action is seen in a number of Spanish examples (see (168a)), as noted by Armstrong (2013). However, it might also be absent in marked autobenefactive constructions (see (168b)).

- (168) a. [construyendo su casa] A levantarme parés, a
to raise.INF-REFL.1SG walls to
levantarme esto, a levantarme lo otro, y a las
raise.INF-REFL.1SG that to raise.INF-REFL.1SG the other and to the
diez me iba.
ten REFL.1SG go.PST1SG
'[when building his house] Raising walls, raising this, raising that, and at
ten I used to leave.' (Salinas, Alicante, COSER 0310)
- b. [la nata] Nos la pringábamos en el pan con un
REFL.1PL ACC.3SG spread.PST.1PL in the bread with a
poquito de azúcar y estaba fenomenal.
little of sugar and be.PST.3SG great
'We used to spread [whipped cream] on the bread with a little bit of sugar
and it tasted great.' (Anaya, Segovia, COSER 3701)

3. Note that while the benefactive construction is considered an argument increasing diathesis calculus, the valency of the verb is not increased in autobenefactive constructions, since two derivations are performed: the benefactive and the indirect reflexive (Kulikov 2011).

Maldonado notes that the degree of proximity between subject and object plays a role in the presence or absence of the autobenefactive RM in Spanish. For instance, the fact that *esfuerzo* ‘effort’ is closer to the subject than *tiempo* ‘tiempo’ would determine the distribution of the RM in the examples of (169) – the RM is required if the object is *esfuerzo*, but not if it is *tiempo*. Note that this distinction seems to be related both to the notion of possessive dative and the higher affectedness of the subject.

(169) (Maldonado 2000)

- a. (Se) ahorró mucho tiempo.
REFL.3 save.PST.3SG much time
‘S/he saved a lot of time.’
- b. Ahorre??(se) el esfuerzo de ir a pagar a la oficina.
save.IMP.3SG-REFL.3 the effort of go.INF to pay.INF to the office
‘Save the effort of going to pay to the station.’

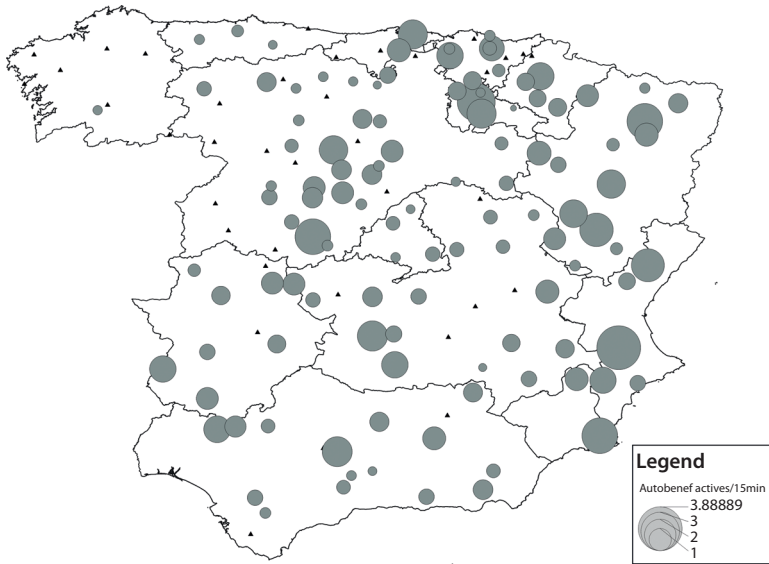
It is also worth noting that, as Kemmer (1993) points out, some verbs that admit autobenefactive participants are typically autobenefactive – they denote actions in which the subject is usually the one that benefits from the action, such as those meaning ‘acquire, choose, get’. As will be seen below (Sections 6.4, 6.5.1), this is also true for Spanish.

Finally, it is important to note that the frequency of the RM with autobenefactive value exhibits geographic variation, as the COSER data show.⁴ Corpus NE documented 487 autobenefactive examples marked with the RM in a total of 113 locations. Since relative frequencies cannot be calculated taking into account unmarked cases, I have standardised these data according to the duration of each interview, in order to chart them on a map.⁵ Map 17 suggests that marked autobenefactive constructions are uncommon in a large area within the northwest, while no clear geographical pattern arises in the rest of the territory. However, it needs to be noted that, despite the standardisation, the fact that only marked examples are considered increases the level of uncertainty.

4. Again, there is also variation across different Spanish-speaking countries, as shown by the fact that some of the examples given by Maldonado (2000) from Mexican Spanish are rather uncommon in my variety of Peninsular Spanish (see De Benito Moreno 2015).

5. The duration of the interviews was computed counting periods of 15 minutes. Fractions were assigned half a point. The standardisation simply involved dividing the total number of autobenefactive marked constructions by this measure of duration, thus obtaining the mean number of autobenefactive marked verbs for a 15-minute period for each interview.

Presence of the RM in autobenefactive constructions (normalised by duration)



Map 16. Autobenefactive constructions marked with the RM by length of the interview in the COSER data

6.2.3 Possessive datives

Possessive datives are a subtype of external possession constructions, where the possessive relation is not coded in the same phrase (König & Haspelmath 1998). That is to say, in constructions of external possession, the possessor and the possessee do not appear in the same constituent. In the varieties under study, the possessor is coded as a dative, while the possessee may appear in different functions: subject of unaccusative verbs (see (170a)), accusative (see (170b)) or oblique objects (see (170c)).⁶ Only in the two latter cases can the dative be reflexive.

- (170) a. Le tiemblan las manos.
 DAT.3SG shake.PRS.3PL the hands
 ‘Their hands are shaking.’

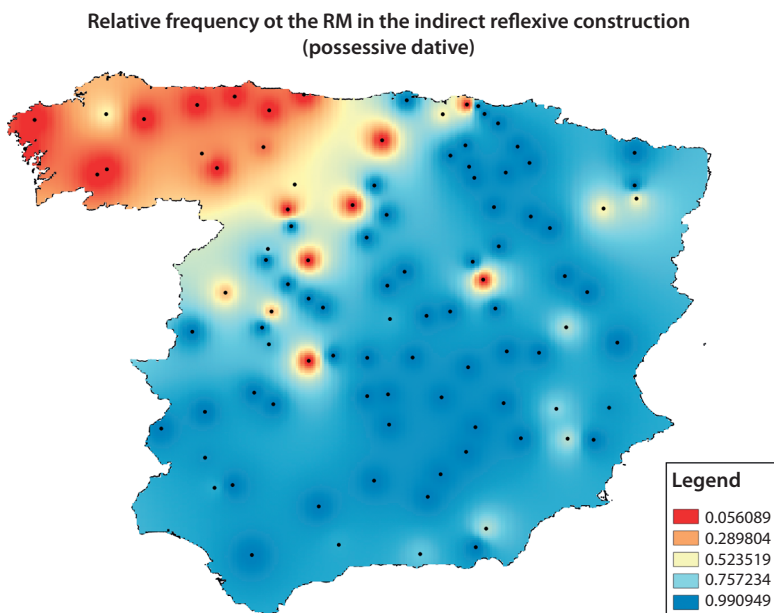
6. The typological survey carried out by König & Haspelmath (1998) on external possession showed that possessive datives are typical of European languages, but seldom found in other languages of the world. European languages also vary according to the productivity of the construction, from languages where it is extremely productive (Spanish, German and Balkan languages, for instance) to languages where it is very rare and only present with oblique arguments (Dutch or English), with intermediate languages such as French (see also Lamiroy 2003).

- b. Le arregló el reloj a Juan. ~ Se arregló el reloj
 DAT.3SG fix.PST.3SG the watch to John ~ REFL.3 fix.PST.3SG the watch
 a sí mismo.
 to REFL.3 same
 'S/he fixed John's watch ~ He fixed his own watch.'
- c. Le miró a los ojos. ~ Se miró a los ojos a
 DAT.3SG look.PST.3SG to the eyes ~ REFL.3 look.PST.3SG to the eyes to
 sí mismo.
 REFL.3 same
 'S/he looked at her/his eyes ~ He looked at his own eyes.'

It has been noted in the literature that when both the internal possession (encoded by genitive case or possessive pronouns) and the external possession construction are possible, the latter usually implies a possessor that is more affected by the verbal event (Rapold 2010; König & Haspelmath 1998; Lamiroy 2003; Linzen 2009). That is to say, reflexive external possession constructions are not only associated with possessive semantics, but also with the notion of affectedness of the subject (since it is also the possessor). Accordingly, it has been proposed that the use of the possessive dative is conditioned by several implicational hierarchies related to affectedness (König & Haspelmath 1998; Lamiroy 2003; Linzen 2009). While these hierarchies do not always coincide, they all agree in placing body parts at the top. Interestingly, it seems that these are the only possessors admitted in external possession reflexive constructions in the varieties under study (Gutiérrez-Ordóñez 1999, De Benito Moreno 2015).⁷

Finally, it is of relevance here to note that the frequency of the RM in reflexive possessive constructions is crucially affected by dialectal parameters. Gramarians have already observed that its presence is rare in Galician (Lantes 1980, Álvarez, Regueira & Monteagudo 1986; Cidrás Escáneo 1991; Dubert García 1999, Álvarez & Xove 2002; Freixeiro Mato 2002) and in Asturian (D'Andrés 1994; Academia de la Llingua Asturiana 2001). The COSER data confirm this for the varieties documented in these areas, but also for a larger area of central- and north-western Castilian dialects. This is clearly shown in Map 17, which charts the frequency of the RM in the 376 examples of reflexive dative possessive constructions in 100 locations.

7. The presence of the RM in external possession reflexive constructions seems to be determined by the direction of the action – the RM is most frequent when the action is directed at the possessee, be it a locative or the direct object of the sentence – and least frequent when the possessee “performs” the action (see Cartagena 1972; De Benito Moreno 2015).



Map 17. Relative frequency of the RM in reflexive possessive dative construction in the COSER data

6.2.4 Indirect detransitivised constructions

In this section I briefly consider a number of indirect reflexive constructions which have a non-reflexive counterpart, but which differ from these (and from the reflexive constructions discussed in Sections 6.2.1–6.2.3) in the fact that their subject is not the agent of the action. In some of them the implication is that the event was caused by an external cause, while the subject depicts an affected participant, while in others the subject becomes a theme that undergoes a change of position. These two situations are reminiscent of anticausative and conversive constructions (see Sections 3.2.1 and 3.2.4), which also have no agentive subjects due to the demotion of the subject of the causative construction. In the indirect constructions that I discuss here, however, there is no promotion to subjecthood by the direct object, because the object coreferential with the subject is the indirect one (while in anticausative and conversive verbs it is the direct object of the transitive construction) – these constructions are still transitive. Crucially, the indirect reflexive constructions addressed here are not semantically reflexive, since the subject is not performing the action on themselves. That is, they are some sort of “indirect

anticausatives” or “indirect detransitivised constructions”.⁸ I will discuss three types of these constructions.

First, a number of these constructions take one of the subject’s body parts as their direct object, being clearly related to reflexive possessive datives. Thus, some of these examples are formally identical to the indirect reflexive possessive (see (171)). The only difference between (171a) and (171b) is the fact that the subject of the former is an agent that voluntarily hurt himself (as indicated by the fact that he apparently did so in order to be cool), while the subject of (171b) is not an agent, but an affected patient, who did not intend to hurt herself. The agentive reading of the subject is cancelled, implying that the subject did not have control over the event (and potentially implying an external cause). This is clearly reminiscent of anticausative and conversive verbs, as well as a number of reflexive intransitive verbs (see Section 5.6). Moreover, these constructions show an affected subject, due to the possessive relationship established between the subject and the body part.

- (171) a. Soy guay y me he pinchado el dedo
 be.PRS.1SG cool and REFL.1SG have.PRS.1SG prick.PTCP the finger
 con la escuadra. Casi me desangro. (?)
 with the set.square almost REFL.1SG bleed.out.PRS.1SG
 ‘Because I’m cool I’ve pricked my own finger with the set square. I almost bled out.’
 (February 19th 2013, Valencia. Tweet.)
- b. Y ENCIMA! Me he pinchado el dedo mientras
 and above REFL.1SG have.PRS.1SG prick.PTCP the finger while
 cosia!! AHG! [Attaches picture of her finger and a rose]
 sew.PST.1SG argh
 ‘And moreover I’ve pricked my finger while sewing!!! Argh!’
 (October 25th 2011. Tweet.)

Second, other examples consist of a light verb and an abstract direct object. Because many of these light verbs are transfer verbs, it can be argued that these constructions are modelled on indirect reflexive constructions with recipient datives. Others, however, seem to be modelled on indirect reflexive constructions with possessive datives, as their direct object is a subject’s body part. They typically express cognition (*{armar(se)/hacer(se)} un lío* ‘to get confused, lit. to make a mess to oneself’, *calentar(se) la cabeza* ‘to overthink, lit. to warm one’s head’, *callar(se) la boca* ‘to shut up, lit. to shut one’s mouth’, *dar(se) una idea* ‘to get a basic idea, lit. to give oneself an idea’) or emotion events (*{dar(se)/pegar(se)} un susto* ‘to get scared, lit. to give

8. Vicente Mateu (2001) seeks to account for the fact that these sentences show two affected participants (linked by a possessive relationship) but no agentive participant by extending the concept of unaccusativity to transitive sentences. See also Fábregas (2021) for a recent review of the possible syntactic analysis.

oneself a scare', *dar(se) una sorpresa* 'to get surprised, lit. to give oneself a surprise'). Because they show a causative non-reflexive counterpart but they do not admit the emphatic reflexive phrase *a sí mismo* (see (172)) they can be assimilated into anticausative constructions, although once again there is no syntactic promotion to subjecthood by the direct object. Depending on the semantics of the predicate, the subject has either an experiencer or an affected patient interpretation.

- (172) a. Perdona, te estoy armando un lío.
 sorry DAT.2SG be.PRS.1SG building a mess
 'Sorry, I'm making you confused.'
- b. ¡Espera!, me estoy armando un lío (??a
 wait REFL.1SG be.PRS.1SG building a mess to
 mí mismo).
 REFL.1SG same
 'Wait! I'm getting confused.'

Third, a verbal phrase that stands out within these constructions is *dar(se) {la/ media} vuelta* 'to turn around, lit. to give oneself the/half turn' (see (173a)).⁹ First, it is the only one in this category that conveys a motion event. Second, while it has a causative unmarked counterpart (see (173b)), the dative in the causative counterpart is typically inanimate, while in the indirect reflexive construction the subject is animate. This seems to be simply a product of how normal life events occur, since it is not implausible to think of an animate indirect object in the causative counterpart (i.e. children).

- (173) a. [los hijos] Si hoy mismo les riñes, se dan
 if today same DAT.3PL tell.off.PRS.2SG REFL.3 give.PRS.3PL
media vuelta y hacen lo que quieren.
 half turn and do.PRS.3PL it that want.PRS.3PL
 'If nowadays you tell [your children] off, they turn around and do whatever they want.' (Valencia del Ventoso, Badajoz, COSER 0728)
- b. [tripa de cerdo] La lavas, le das la
 ACC.3SG clean.PRS.2SG DAT.3SG give.PRS.2SG the
vuelta y luego la salmoreas.
 turn and then ACC.3SG brine.PRS.2SG
 'You clean [the pig's intestine], you turn it around and brine it afterwards.' (Huércanos, La Rioja, COSER 2506)

9. I consider examples of *dar una vuelta* and *dar las vueltas*, meaning 'to go for a walk', to be reflexive transitive verbs and not indirect reflexive constructions, because their meaning ('to go for a walk') does not correspond with an "anticausative" version of the unmarked counterpart.

Finally, I would like to highlight the fact that while the concept of “indirect anticausative” or “indirect detransitivised” construction is somehow unorthodox, it finds further support in the fact that most of these periphrastic expressions seem to be related to an (almost) synonym anticausative verb etymologically related to the noun: *dar(se)/pegar(se) un susto* = *asustar(se)*, *dar(se) la vuelta* = *volver(se)*, *armar(se) un lío* = *liar(se)*, etc.

6.3 Transitive verbs: General considerations

After this brief discussion of indirect reflexive constructions, I turn to a deeper analysis of reflexive transitive verbs on the basis of the COSER data. The distinction between these two types of reflexive constructions is both semantic and syntactic – in reflexive transitive verbs the RM is not semantically interpretable as one of the participants of the event, so it is not interchangeable by a non-coreferential pronoun. In this section I provide a general overview of the reflexive transitive verbs documented in the COSER data, mainly regarding the frequency of the RM and their geographical distribution.

As in the case of reflexive intransitive verbs, collecting reflexive transitive verbs poses some problems concerning the exhaustivity of the data collection method. Therefore, I only collected all the instances of certain verbs that are known to take the RM. As before, I collected some verbs in all the interviews and some others only in subcorpus E, due to their high frequency.

Transitive verbs that were exhaustively collected only in subcorpus E are the following: *saber* ‘to know’, *llevar* ‘to take, bring’, *comer* ‘to eat’, *ver* ‘to see’, *dejar* ‘to leave’, *tomar* ‘to take’, *estudiar* ‘to study’, *mirar* ‘to look’, *ganar* ‘to win’, *aprender* ‘to learn’, *pensar* ‘to think’, *entender* ‘to understand’, *leer* ‘to read’, *encontrar* ‘to find’. Transitive verbs that I collected exhaustively in all the COSER interviews (corpus NE) are: *traer* ‘to bring’, *pasar* ‘to pass’, *creer* ‘to believe’, *beber* ‘to drink’, *recordar* ‘to recall’, *bajar* ‘to take down’ and *subir* ‘to take up’. From all the other transitive verbs I only collected marked examples.¹⁰

With this method, I collected a total of 96 transitive verbs that showed the RM at least once (see Appendix 4). As regards the semantic class of these verbs, the COSER data show that the semantic classes described by Kemmer (1993) for the middle voice are not as well represented as they were with reflexive intransitive

10. Marked examples of some verbs that were collected exhaustively only in subcorpus E, like *dejar*, were also collected in the rest of the interviews.

verbs (see Figure 45).¹¹ Only 45.8% of these transitive verbs (49 in total) belong in the semantic classes proposed by Kemmer, compared to the 71% and 72.4% that fit into one of those categories in reflexive intransitive verbs and in reflexive verbs with change of diathesis (see Sections 5.2 and 3.4 respectively). Moreover, the largest class of reflexive transitive verbs that fits in Kemmer's categories consists of cognition verbs (29/107, 27.1%), which amount to 6% (36/601) of reflexive verbs with a change of diathesis (see Section 3.4). This lack of correspondence is expected if the major source of the analogy does not lie in the middle uses of the RM but in its indirect uses. Globally, this indicates that reflexive transitive verbs are not similar to reflexive middle verbs.

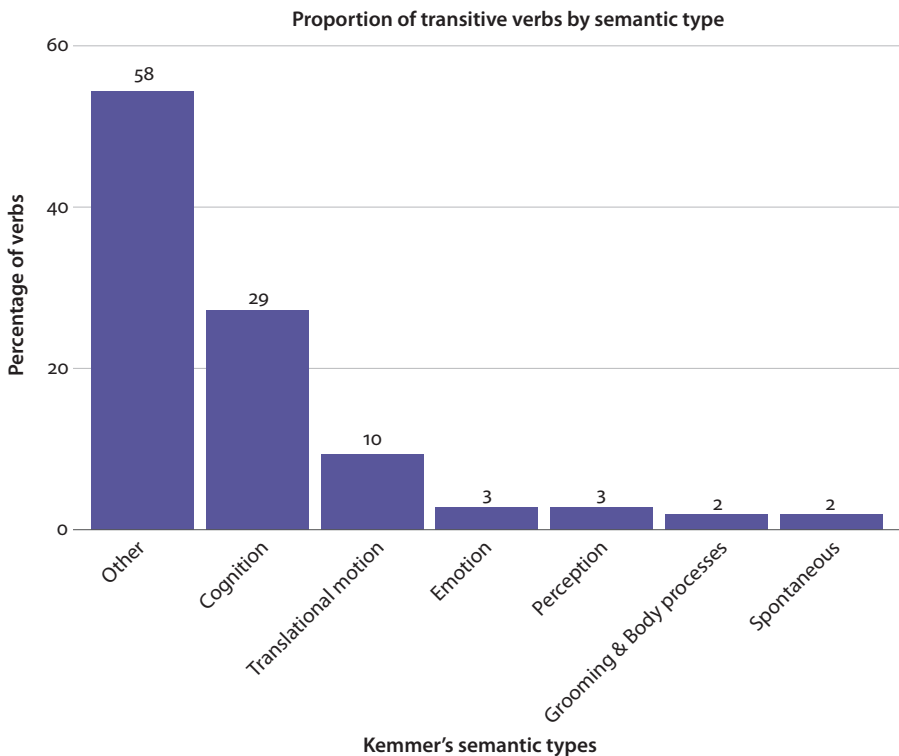


Figure 45. Semantic categories of reflexive transitive verbs in the COSER data

11. Note that the total of verbs represented in Figure 45 is 107. This is due to the fact that a number of these verbs are polysemic and were thus considered in several categories.

The specific verbs included in each type are given in Table 35.¹² Although most of them are not included in Kemmer's categories, we can identify some categories that have been observed in the specific literature on this construction. First, verbs of consumption are a large group, including 13 verbs: *almorzar, beber, comer, devorar, fumar, leer, mamar, merendar, pegar, terminar, tomar, tragar* and *usar*. Second, we also document other verbs with affected (*molturar*) and effected objects (*escribir, echar*).¹³ Third, verbs with an inherent autobenefactive meaning are also common, with 13 verbs: *cobrar, coger, derrochar, ganar, ligar, merecer, pescar, quedar, reunir, sacar, tener, trabajar*. Fourth, there is a small group of verbs that convey the meaning 'to spend time', like *echar, llevar, pasar* and *tirar* (which Nishida (1994) includes within verbs of consumption with an abstract object). Fifth, it is worth noting that there are some categories (grooming and corporal processes, emotion, translational motion) that have not been considered in the literature. Finally, a large number of verbs are not included in any of these semantic groupings, which hints at a high productivity and heterogeneity of the construction.

In order to be able to perform a quantitative analysis of the frequency of the RM and the characteristics of the subject, I will only consider the verbs that were extracted exhaustively from the interviews. These amount to 21 different verbs, whose overall frequency of appearance varies widely (see Figure 46).

In total, 7,494 examples of these 21 verbs were collected, from which only 843 (11.2%) show the RM. This contrasts with intransitive verbs, 18.3% of which were marked (see Section 3.2.6). As with intransitive verbs, when we look at each verb individually there seems to be a lot of internal variation (see Figure 46). While some semantically similar verbs, like *comer* and *beber* (but not *tomar*, also a consumption verb) show similar frequencies of the RM, within cognition verbs the differences are rather large: *saber, entender, estudiar* and *aprender* show a very low proportion of marked examples; *pensar* and *recordar* are about twice as frequently marked; and *creer* shows a far higher proportion of marked examples. The same can be said for verbs of translational motion verbs, since the proportion of the RM with *llevar* is far higher than with *traer*.

12. For the classification I considered the meaning of the verb in the specific context.

13. The presence of the RM in this semantic class was explicitly ruled out by Nishida (1994). The example with *echar* is rather atypical, since its meaning is closer to *hacer* 'to make':

(i) Pues vamos a hilar, y se echaban mantas. Con la lana
well go.PRS.1PL to spin.INF and REFL.3 throw.PST.3PL blankets with the wool
de las ovejas, sí.
of the sheep yes
'Let's spin the wool, and they made blankets, With the wool of the sheep, exactly'

(Mahide, Zamora, COSER 4617)

Table 35. Reflexive transitive verbs by semantic class in the COSER data

Cognition	<i>aprender</i> 'to learn', <i>calcular</i> 'to calculate', <i>callar</i> 'to be quiet', <i>chupar</i> 'to suck', <i>confundir</i> 'to confuse', <i>conocer</i> 'to know', <i>creer</i> 'to believe', <i>cuidar</i> 'to think', <i>dar</i> 'to give', <i>dar cuenta</i> 'to understand', <i>discurrir</i> 'to reflect', <i>encontrar</i> 'to find', <i> echar cuenta</i> 'to imagine', <i>engañar</i> 'to deceive', <i>entender</i> 'to understand', <i>esperar</i> 'to hope', <i>estudiar</i> 'to study', <i>explicar</i> 'to explain', <i>hacer</i> 'to make', <i>imaginar</i> 'to imagine', <i>inventar</i> 'to invent', <i>llevar</i> 'to take', <i>olvidar</i> 'to forget', <i>pensar</i> 'to think', <i>recordar</i> 'to recall', <i>repasar</i> 'to review', <i>saber</i> 'to know', <i>suponer</i> 'to suppose', <i>temer</i> 'to fear'
Emotion	<i>aguantar</i> 'to endure', <i>divertir</i> 'to have fun', <i>reír</i> 'to laugh'
Grooming and body processes	<i>hacer</i> (<i>pis, caca...</i>) 'to take (a piss, a shit...)', <i>llevar</i> (<i>prendas</i>) 'to wear'
Translational motion	<i>bajar</i> 'to go/take down', <i>dar la vuelta</i> 'to turn around', <i>dar una {vuelta/paseo}</i> 'to go for a walk', <i>llevar</i> 'to take', <i>pasar</i> 'to pass, to cross', <i>pegar un paseo/una caminata</i> 'to go for a walk', <i>recorrer</i> 'to cover a distance', <i>subir</i> 'to go/take up', <i>traer</i> 'to bring'
Perception	<i>mirar</i> 'to look', <i>oír</i> 'to hear', <i>ver</i> 'to see'
Spontaneous	<i>cuajar</i> 'to clot', <i>hacer</i> 'to make'
Other	Consumption <i>almorzar</i> 'to have for lunch', <i>beber</i> 'to drink', <i>comer</i> 'to eat', <i>devorar</i> 'to devour', <i>fumar</i> 'to smoke', <i>leer</i> 'to read', <i>mamar</i> 'to suckle', <i>merendar</i> 'to snack', <i>pegar</i> 'to take', <i>terminar</i> 'to finish', <i>tomar</i> 'to take', <i>tragar</i> 'to swallow', <i>usar</i> 'to use'
	Affected objects <i>molturar</i> 'to grind'
	Effected objects <i>escribir</i> 'to write', <i> echar</i> 'to throw'
	Inherently autobenefactive <i>cobrar</i> 'to earn', <i>coger</i> 'to get', <i>derrochar</i> 'to squander', <i>ganar</i> 'to win, to earn', <i>ligar</i> 'to pick up girls/boys', <i>llevar</i> 'to take', <i>merecer</i> 'to deserve', <i>pasar</i> 'to undergo, to enjoy, to suffer', <i>pescar</i> 'to get, take', <i>quedar</i> 'to acquire', <i>reunir</i> 'to have, to put together', <i>sacar</i> 'to get', <i>tener</i> 'to have', <i>trabajar</i> 'to work on something'
	Spending time <i> echar</i> 'to spend', <i>llevar</i> 'to spend', <i>pasar</i> 'to spend', <i>tirar</i> 'to spend'
	Other <i>aguardar</i> 'to wait', <i>apañar</i> 'to fix', <i>arreglar</i> 'to fix', <i>calzar</i> 'to wear shows', <i>cargar</i> 'to load, to carry', <i>carrear</i> 'to carry', <i>chivar</i> 'to rat', <i>correr</i> 'to run', <i>dar</i> 'to give', <i>decir</i> 'to say', <i>dedicar</i> 'to devote', <i>dejar</i> 'to leave', <i>desobedecer</i> 'to disobey', <i>encerrar</i> 'to lock up', <i>encontrar</i> 'to find', <i>equipar</i> 'to equip', <i>haber</i> 'to be (existential)', <i>hacer</i> 'to make, to do', <i>intentar</i> 'to try', <i>jugar</i> 'to play', <i>pelar</i> 'to peel', <i>perder</i> 'to lose', <i>picar</i> 'to bite', <i>querer</i> 'to want', <i>sacar</i> 'to get', <i>tropezar</i> 'to stumble', <i>valer</i> 'to be worth', <i>vender</i> 'to sell'

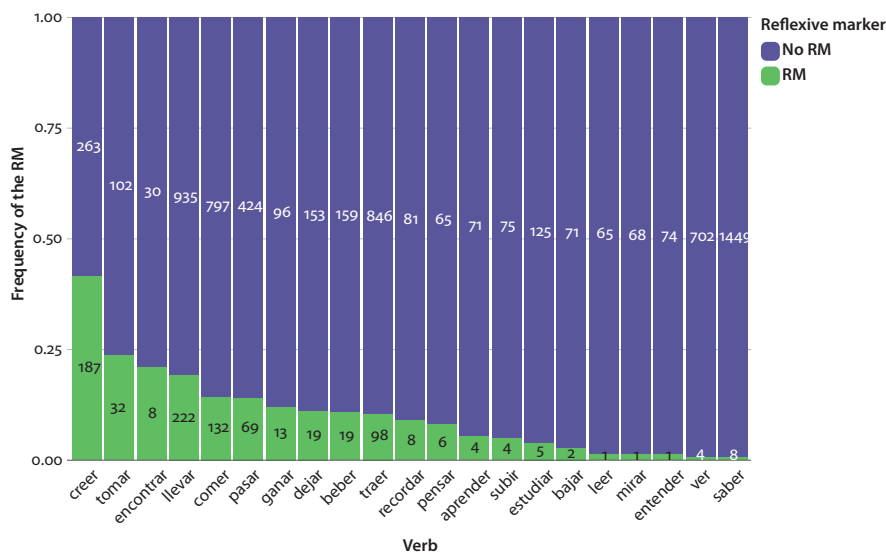
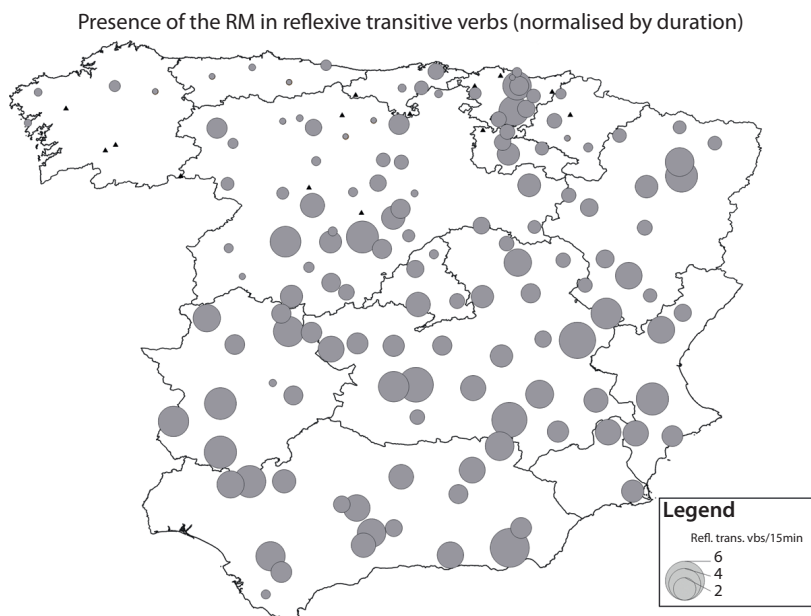


Figure 46. Frequency of the RM in selected (exhaustively collected) transitive verbs by verb in the COSER data

As regards the animacy of the subject, it is worth noting that only nine of these 21 verbs (*beber*, *comer*, *dejar*, *ganar*, *llevar*, *pasar*, *subir*, *tomar* and *traer*) can take inanimate subjects. Moreover, only two of these verbs appeared more than five times with inanimate subjects, namely, *llevar* ($N = 145$) and *traer* ($N = 46$). That is, since transitive verbs show a strong tendency to take animate subjects, it makes little sense to check the effect of this parameter in general.

Finally, let us turn to the geographical distribution of the RM in transitive verbs. Since I was not exhaustive in the collection of the unmarked examples, the relative frequencies of marked examples in relation to all examples cannot be calculated. Accordingly, Map 18 charts only marked examples, whose frequency was normalised twice: (1) in order to account for the fact that not all the interviews are equally long, I divide the total occurrences by the number of 15 minute periods for each interview (fractions were assigned half a point); (2) in order to account for the fact that I collected examples of a larger number of verbs in subcorpus E, I divided the total occurrences in corpus E by the result of dividing the mean frequency of reflexive transitive verbs per location in subcorpus E (mean = 13.47) by the equivalent measure in subcorpus NE (mean = 9.13). It still needs to be born in mind that these figures show larger degree of uncertainty than if unmarked and marked examples could have been compared to each other. Map 18 shows a large area in the northwest where the frequency of the RM is lower. This area is larger



Map 18. Relative frequency of reflexive transitive verbs in the COSER data

than the one relevant for anticausative verbs (see Section 3.2.1), and is reminiscent of the area of lower frequency of the RM found for reflexive possessive datives (see Section 6.2.3) or even some reflexive intransitive verbs (see Chapter 5).

As can be seen from this section, then, the behaviour of reflexive transitive verbs differs in a number of ways from the behaviour of reflexive verbs with change of diathesis. On the one hand, Kemmer's semantic classes are not a good predictor of the presence of the RM when the class is considered globally. On the other hand, the geographical area most resistant to the presence of the RM with transitive verbs (see Map 18 above) is larger.

In what follows I will focus in more detail on some specific transitive verbs which can show the RM, in order to show how they are connected to previous functions of the RM, and to propose that the main sources of analogy are indirect reflexive constructions and (some) reflexive intransitive verbs. Due to the scant data for some groups of verbs (such as those that convey perception, emotion and spontaneous events, or verbs with effected objects), I will not deal with all the possibilities found in the COSER data, but only with the most important in quantitative terms.

6.4 Inherently autobenefactive transitive verbs

The fact that some reflexive transitive verbs show inherently autobenefactive semantics (such as ‘to get’, ‘to acquire’) strongly hints at an analogical influence of the autobenefactive reflexive construction. Some of these verbs were documented only sporadically with the RM, such as *cobrar* ‘to earn’, *coger* ‘to get’, *derrochar* ‘to squander’, *ligar* ‘pick up girls/boys’, *pescar* ‘(to fish’, meaning ‘to get, take’ in this context), *reunir* ‘to have, to put together’, *tener* ‘to have’ and *trabajar* ‘to work on something’.¹⁴

Other inherently autobenefactive verbs, however, were found frequently in the reflexive transitive construction. For instance, although *merecer* ‘to deserve’ was a globally rare verb, with only nine documented examples, six of them (66.7%) show the RM. Similarly, *quedar* can appear in a transitive construction with the meaning ‘to acquire, to keep’ (see (174a)), where it strongly prefers the RM – all nine examples are marked. This transitive construction of *quedarse* is reminiscent of an intransitive construction of *quedar(se)* with a prepositional phrase with *con* ‘with’ or *sin* ‘without’ meaning ‘to acquire, to keep, to loose the possession of’, illustrated in (174b). This intransitive construction, which was disregarded in Section 5.7, is attested 50 times in the COSER data, where is marked 42 times (84%) – all unmarked examples are found either in north-western or neighbouring varieties. The transitive construction is only found in the rest of the territory, albeit infrequently, a finding that fits a historical connection between transitive *quedarse* and intransitive *quedarse con/sin*.

- (174) a. “Aquí faltan veinte duros”. “Pues esos veinte duros me
here lack.PRS.3PL twenty bucks well those twenty bucks REFL.1SG
los he quedado yo”.
ACC.3PL have.PRS.1SG keep.PTCP I
“‘There are twenty bucks missing.’ ‘I’ve taken those twenty bucks.’”
(Arjona, Jaén, COSER 2301)
- b. Nos quedábamos con las cosas menudas. Las morcillejas,
REFL.1PL keep.PST.1PL with the things little the blood.sausages
los choricejos, los jamones, pues se vendían.
the sausages the hams well REFL.3 sell.PST.3PL
‘We kept the small things. Blood sausages, sausages and hams, we sold them.’
(Malagón, Ciudad Real, COSER 1414)

14. *Coger* was extensively documented in the reflexive autobenefactive construction analysed in Section 6.2.2. The uses considered here, however, do not admit the alternation of the RM with a non-coreferential beneficiary, which is why they are considered as instances of the reflexive transitive construction.

Finally, a verb like *ganar* ‘to win, to earn’ can also take the RM, although not as frequently – from the total of 109 examples of transitive *ganar* collected in subcorpus E, only 13 (11.9%) show the RM. I collected all the instances of *ganarse* in the whole corpus, and these included 25 examples of reflexive *ganarse*. Most typically, *ganarse* appears with objects signifying money, but it is also frequent when the object refers to a prize (De Benito Moreno 2015: § 7.3.2.2.5). However, all these uses can also be encoded by unmarked *ganar*, and, since I did not collect exhaustively all examples of *ganar*, this distribution of the type of objects preferred by *ganarse* might be due to the global distribution of these objects with *ganar(se)*.

Some verbs, although not inherently autobenefactive, frequently take the RM in constructions that have an autobenefactive nuance. This is the case with *echar* ‘to take, lit. to take’, which very often takes the RM in constructions like *echar(se) {{la/una} partida/un baile/una carrera}* ‘to play, to dance, to run lit. to throw {{a/the} game / a dance / a race}’, *echar(se) la siesta* ‘to take a nap’, *echar(se) {un café/un trago}* ‘to drink a coffee, something, lit. to throw {a coffee / a sip}’, etc. Of 21 such examples, ten (48%) showed the RM. The reflexive form was the only one recorded in the 24 examples in the construction *echarse novio/a* ‘to get a boyfriend/girlfriend, lit. to throw boyfriend/girlfriend’.¹⁵

The situation of *sacar* ‘to take out’ in contexts meaning ‘obtaining a degree or certificate’ is very similar – it frequently takes the RM (9 examples out of 51, 37%), but it is possible and even more frequent without it (see (175)).

- (175) a. Mi hijo s’ha sacao la carrera.
 my son REFL.3-have.PRS.3SG take.out.PTCP the degree
 ‘My son has obtained his degree.’ (Barrax, Albacete, COSER 0204)
- b. Hoy la mitad de los abogados, si tuviesen que
 today the half of the lawyers if have.PST.SBJV.3PL that
sacar la carrera no la tenían.
 take.out.INF the degree not ACC.3SG have.PST.3PL
 ‘If half of the lawyers had to obtain their degrees today, they couldn’t do it.’ (Porzuna, Ciudad Real, COSER 1417)

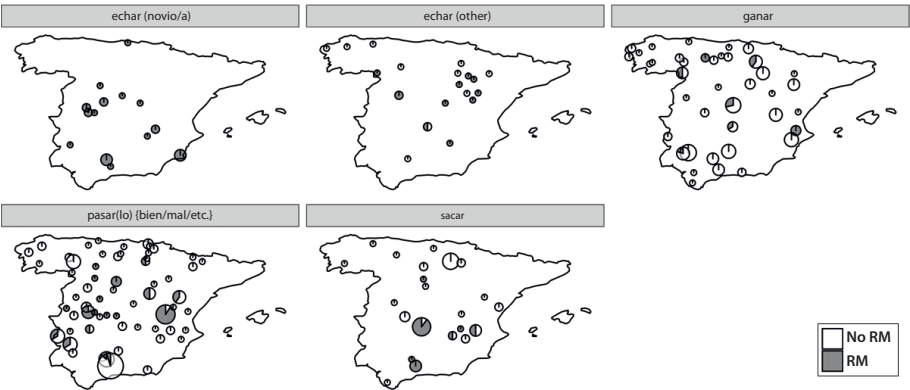
Finally, the behaviour of *pasar(se)* ‘to undergo, to enjoy, to suffer’ is also very interesting. In constructions such as *pasar(se)lo {bien/mal}* ‘to have a {good/bad} time, lit. to pass it {well/bad}’, where it combines with a non-referential clitic object and a typically adverbial adjunct that evaluates the experience conveyed by the verb,

15. The two examples of the construction *hacerse novio/a* ‘to get a boyfriend/girlfriend, lit. to make boyfriend/girlfriend’ were also marked. This construction is a regional use found in the Region of Valencia and is not to be confused with *hacerse novios* ‘to become a couple’, which is an anticausative pseudocopulative construction.

the presence of the RM seems to be restricted to positive experiences. Of 87 examples of this construction with positive adjuncts such as *bien* ‘well’, 36 (41%) were marked. However, when the construction appeared with negative adjuncts such as *mal* ‘bad’ ($N = 52$), or when the example was unmarked regarding the evaluation of the experience, either in questions with *cómo* ‘how’ or in combination with the adverb *así* ‘like that’ ($N = 5$), it was always unmarked. This radical semantic split is interesting because it is compatible with the claim that the RM emphasises the involvement of the agent (see Section 6.2.2) – it seems to be part of human nature that good experiences tend to be presented as a result of the action of the subject, whereas bad experiences are preferably attributed to external circumstances.

Moreover, *pasar(se)* can also appear in less fixed transitive constructions with the meaning ‘to undergo, to enjoy/to suffer’. Of 202 examples in the COSER data, nine (4.5%) showed the RM. Taking into account the behaviour of the RM in *pasar(se)lo* {*bien/mal*}, this low frequency might have to do with the fact that this construction is almost always used to refer to negative experiences such as *pasar* {*hambre / frío / miedo / la guerra*} ‘lit. to pass hunger / cold / fear / the war’ (171/202, 84.7%). It should also be noted that this use of *pasar(se)*, especially when its object is an event, is semantically connected to those discussed in Section 6.6, where *pasar* means ‘to spend time’, where the absence of the RM is also preferred.

To conclude this section, it is worth noting that, despite the scant and irregular documentation of these verbs across the territory, Map 19 shows a consistent lack of marking in north-western varieties in the case of *echar*, *ganar*, *pasar(lo) bien* and *sacar*. Note that, although *echarse (novio/a)* was always marked, the construction was only seldom used in the north of the territory.



Map 19. Frequency of the RM with selected inherently autobenefactive verbs

6.5 Verbs of consumption

Verbs of consumption such as *comer(se)* ‘to eat’, *beber(se)* ‘to drink’ or *leer(se)* ‘to read’ have been the most commonly mentioned in works on aspectual *se* (Nishida 1994; Zagona 1996; Sánchez López 2002; Campanini and Schäfer 2011), due to the fact that they seem to show a strict rejection of non-delimited objects: see (176a). Some authors, however, have noted that such a restriction is not absolute (De la Mora 2011; Romero & Teomiro 2012; Armstrong 2013): see (176b). Moreover, they have also been the focus of more recent quantitative analyses that claim that the aspectual analysis is incomplete (De la Mora 2011; Rivas 2011).

- (176) a. Juan *(se) comió paella.
 Juan REFL.3 eat.PST.3SG paella
 ‘Juan ate paella.’ (Sánchez López 2002: 110)
- b. Sócrates se comió veneno
 Socrates REFL.3 eat.PST.3SG poison
 ‘Socrates ate poison.’ (Romero & Teomiro 2012)

The COSER data document enough examples of *comer(se)*, *beber(se)* and *tomar(se)* so as to enable an in-depth look at the contexts that favour the presence of the RM.¹⁶ When we look at the global distribution of the RM with these three verbs by geographical area (see Figure 47), the first thing we note is that north-western varieties mostly reject the RM with these verbs. Accordingly, in the remainder of this section I will disregard the data of these varieties. In the rest of the territory, the most frequent reflexive form is *tomarse* (which appeared in 40.5% of all the instances of *tomar*), followed by *comerse* (29.7%). *Beberse* is the least frequently reflexive verb (18.6%).¹⁷ This situation is reminiscent of the one found for several groups of semantically close intransitive verbs, with different verbs showing different productivity of the RM.

16. All three verbs were collected exhaustively, *beber* in all the interviews (as it was the least frequent) and *comer* and *tomar* only in subcorpus E.

17. The distribution of my data differs from that of Rivas (2011), who finds a far greater frequency of *comerse* (110/172, 64%) and *beberse* (24/78, 31%). The frequency of *tomarse*, however, is very similar in both corpora (75/197, 38%). In De la Mora (2011), these verbs are ordered following the same hierarchy as in Rivas (2011), but with far lower relative frequencies (*comerse*, 36% > *tomarse*, 22% > *beberse*, 16%).

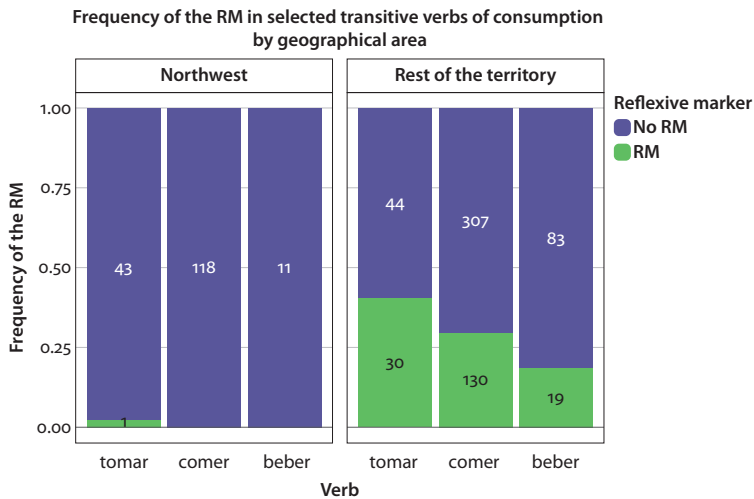


Figure 47. Frequency of the RM with *comer(se)*, *beber(se)* and *tomar(se)* in the COSER data

Absolute uses of *comer* and *beber* (that is, intransitive uses meaning ‘to have lunch’, ‘to ingest food’, ‘to drink alcohol’ or ‘to ingest fluids’) are disregarded here (note that they were included in 6.3, where all uses were considered).

As regards the properties of the object in these examples, it has been noted that bare nouns and, in general, non-delimited objects are not used with reflexive transitive verbs. Figure 48 summarises the data for *comer(se)*, *beber(se)* and *tomar(se)* found in the rest of the territory in this regard.¹⁸ On the one hand, these data show that the difference between delimited and non-delimited direct objects is quite relevant

18. Operationalising this distinction is not straightforward. I will follow Nishida’s (1994) definitions, since they are the most detailed ones (they are rather similar to the criteria followed by De la Mora 2011). For this author, “[nominals] compatible with SE_q refer to entities that are quantitatively delimited, whereas those not compatible with SE_q refer to entities that do not have a clear upper bound.” (Nishida 1994: 431) That is, mass terms and bare plurals are never quantitatively delimited. Count nouns and terms of measurement, however, can be quantitatively delimited, but are not necessarily so, since the interpretation of the sentence must be taken into account. I have given primacy to the formal criterion, since subtle semantic notions are often difficult to grasp. That is, it must be borne in mind that formally non-delimited objects are always non-delimited semantically, while formally delimited objects are typically, but not always, delimited semantically. I have classified as non-delimited direct objects (1) mass nouns, (2) bare plurals, (3) nominals quantified with the adverbs *más* ‘more’, *menos* ‘less’, *mucho(s)* ‘much/many’, *poco(s)* ‘little/few’ and *tanto* ‘so much’, because they do not have an inherent limit; (4) partitive objects, and (5) sentential objects, which “cannot measure the event” (Sanz and Laka 2002: 321, my translation). Null pronouns have been considered a category on their own, since there is geographical variation as to their referential properties in Peninsular Spanish

(6/786, 0.8%) of inanimate subjects were documented in the whole territory. At any rate, inanimate subjects are at least possible even with the RM, as shown in (178).¹⁹

- (178) a. [lomo, costillares, en agua] Se mete pa que se
REFL.3 put.in.PRS.3SG for that REFL.3
escorran porque la sal ya se la han
drain.PRS.3PL because the salt already REFL.3 ACC.3SG have.PRS.3PL
tomao, pa que después no estén salaos.
have.PTCP for that after not be.SBJV.3PL salty
‘They [cured meat products] are put into water so that they drain (remove)
the salt, because they have already taken the salt in, so that they are not
too salty afterwards.’ (Fuentes Claras, Teruel, COSER 4117)
- b. Al primer templo cristiano consagrado en Málaga se
to.the first temple Christian consecrate.PTCP in Malaga REFL.3
lo come la humedad.
ACC.3SG eat.PRS.3SG the humidity
‘Humidity is eating the first consecrated Christian temple in Malaga.’
(Ignacio A. Castillo, *La opinión de Málaga*, “A Santiago se la come la
humedad”, December 10th 2013, [http://www.laopiniondemalaga.es/
malaga/2013/12/15/santiago-come-humedad/639228.html](http://www.laopiniondemalaga.es/malaga/2013/12/15/santiago-come-humedad/639228.html))

A generalised linear mixed model based on the data of the rest of the territory (see Table 36) confirms that having a delimited object (as opposed to a non-delimited one, null objects were removed from the data, since they are ambiguous and scarce) favours the presence of the RM. The model also suggests that the differences between verbs are not too great.

Table 36. Generalised linear mixed model. Presence of the RM in the rest of the territory depending on the delimitation of the object [reference level: Non-delimited] (fixed factor) and the place and verb lexeme (random factor)

Effect	Term	Estimate	Std. error	Statistic	Conf. low	Conf. high
fixed	(Intercept)	−4.799	0.729	−6.581	−6.228	−3.369
fixed	Delimitation of the object: Delimited	4.08	0.615	6.628	2.873	5.286
random (COSERID)	sd (Intercept)	2.118	NA	NA	NA	NA
random (Verb)	sd (Intercept)	0.495	NA	NA	NA	NA

19. Due to the scarcity of examples in the imperative ($N = 10$ in the rest of the territory) it makes little sense to test the effect of this parameter, which could be interesting as a means of investigating the effect of the characteristics of the subject (higher involvement or agentivity) in the presence of the RM with verbs of consumption.

- b. sà-vá-s-íyù tá Yüzù
 drink-APPL-drink-1SG OBJ home.brew
 'I drank all the beer.'

The same implicature is found when the RM *zich* is attached to *eten* 'to eat' in the Dutch spoken in Herleen (Cornips & Hulk 1996):

- (181) Jan eet zich een boterham
 Jan eat.PRS.3SG REFL.3 a sandwich
 'Jan eats a (whole) sandwich completely.'

These implicatures are reminiscent of the link between reflexive transitive verbs and delimited objects, since only with delimited objects can the implicature of 'consuming the whole object' arise. Moreover, such implicatures are also related to the notion of an emphasis on the involvement of the subject and the impressiveness of the event, since consuming the whole object requires greater involvement than consuming just a part of it and is more impressive (see also De la Mora 2011; Armstrong 2013). In Section 6.5.1 I explore this notion of emphasis in autobenefactive constructions and other reflexive transitive constructions, arguing that such emphasis is most likely the historical source of the association between the RM and delimited objects in verbs of consumption.

6.5.1 Emphasis in autobenefactives and reflexive transitive verbs

Can the association between the RM in consumption verbs be explained by the behaviour of the RM in autobenefactive constructions? In Section 6.2.2 I have proposed that the presence of an autobenefactive pronoun may add emphasis to the sentence, since they are typically redundant (in the sense that the default reading of a given sentence in terms of benefaction is generally that the subject acts for its own benefit). Since the RM is coreferential with the subject, the expectation is that the added emphasis would affect the representation of the agent. Actually, the existence of emphasis related to higher agentivity has been mentioned in several studies on the RM with transitive verbs (Armstrong 2013; Maldonado 2000). However, the fact that such an emphasis is not present in all examples of reflexive transitive verbs has led to a focus on the properties of the object to explain the presence of the RM. The preeminence of the delimitation of the object as opposed to the agentivity of the subject is clear in the case of consumption verbs. At the end of Section 6.5 I have argued that both notions (higher involvement of the agent and delimitation of the object) can be seen as two facets of the same implicature. In this section I argue that the association with delimited objects can already be observed in autobenefactive constructions.

In some (but not all) marked autobenefactive constructions there is a clear will to highlight a variety of vague and emphatic meanings, typically related to a high degree of involvement of the subject. As the autobenefactive examples given in (182) show, this high involvement of the subject is due to the characteristics of the events, which are understood to be especially hard to achieve. That is, they are what Armstrong (2013) calls “impressive events” – he calls the RM in such constructions agentive reflexive clitics. While every event can be emphasised, regardless of its telicity, the “impressiveness” of the event often arises either from the fact that it implies a list of different objects (see (182a) and (168a) above) or from the fact that the event involves a whole object (see (182b)). In such contexts, the complexity of the subject’s actions and/or the subject’s hard work, are highlighted. Because emphasis on the involvement of the subject is related to the characteristics of the event, it can be transferred to the characteristics of the object. I believe that the association with the delimitation of the object arises from the second type of contexts.

- (182) a. Te criabas pollos, gallinas, conejos
 REFL.2SG bred.PST.2SG chicken hens rabbits
 [y yo me dejaba mis hijas pequeñas allí en un capazo y me iba de
 buena mañana a coger hierba pa los conejos...]
 ‘You had to breed chickens, hens, rabbits [and I used to leave there my little
 girls in a basket and go early in the morning to get grass for the rabbits...]’
 (Enguera, Valencia, COSER 4310)
- b. Mi mujer se ha bordao ella toda su dote
 my wife REFL.3 have.PRS.3SG embroider.PTCP she all her dowry
 y se la ha hecho ella y a mis
 and REFL.3 ACC.3SG have.PRS.3SG make.PTCP she and to my
 hijas también.
 daughters too
 ‘My wife embroidered her whole dowry and she also made it, and for my
 daughters too.’ (Porzuna, Ciudad Real, COSER 1417)

That is, I argue in the exact opposite direction from Sanz & Laka (2002), in that they claim that the RM attaches to accomplishments, and that the fact that it also takes agentive subjects is a by-product of the fact that accomplishments are always agentive. I propose that the clitic originally appeared in emphatic contexts related to an agentive subject, as predicted by its autobenefactive function, and that the tendency to take delimited objects is a by-product of the (secondary) association of these emphatic contexts to delimited objects. Regarding the case of consumption verbs, this prediction would entail that the RM appeared first in contexts where the event was hard to achieve, such as eating/drinking large quantities of food/liquid, so as to highlight the impressive achievement by the agent. A typical example would

be eating/drinking all of something, which would require delimited objects. In this sense I claim that the association with delimited objects has come about through the same process that caused the association between the marked form *irse* and *de*-source complements.

I argue that the existence of reflexive transitive verbs (as opposed to autobenefactive constructions) where the high involvement of the subject is quite clear, such as (183), shows the intermediate step between emphatic autobenefactive constructions and non-emphatic reflexive transitive verbs of consumption. Armstrong (2013) distinguishes two types of reflexive transitive constructions, namely, agentive reflexive clitics (emphatic) and transitive *se* constructions (non-emphatic). At least with some verbs (like verbs of consumption), I consider them to be two different stages in the analogical extension of the RM across transitive verbs, the latter being more routinised than the former.

- (183) a. [“Los españoles, los mismos lobos, na más que con diferentes collares”.]
 Eso me lo he dicho yo delante de los
 that REFL.1SG ACC.3SG have.PRS.1SG say.PTCP I in.front.of of the
 cuatro feligreses,
 four parishioner
 [porque se pelean por los partidos, pero yo no me peleo...]
 ‘[“Spaniards, same dogs, different collars”] I’ve said that in front of the four
 parishioners, [because they fight for the political parties, but I don’t].’
 (Povedilla, Albacete, COSER 0222)
- b. [Yo miedo no lo he conocido nunca. ...] Cuántos días a las doce
 how.many days to the twelve
 de la noche, ¡oy!, sí, me he encerrao las gallinas.
 of the night oh yes REFL.1SG have.PRS.1SG lock.PTCP the hens
 ‘[I’ve never been afraid. ...] How many days I had to shut the hens at
 midnight!’
 (Urriés, Zaragoza, COSER 4720)

Emphatic reflexive transitive verbs are often found in colloquial expressions (see (184)). This is not surprising, since situations of communicative immediacy (see Koch & Oesterreicher 2007 for this concept), where colloquial speech is typically used, foster the use of intensifying and emphatic constructions. Thus, I believe that a study of the distribution of reflexive transitive constructions in different registers could offer support for the hypothesis that they are historically linked, the expectation being that both emphatic and non-emphatic reflexive transitive constructions would combine more often with the RM in situations of communicative immediacy than in situations of communicative distance.

(185) (Maldonado 2000)

- a. (??Me) leí la última novela de García Márquez.
REFL.1SG read.PST.1SG the last novel of García Márquez
'I read García Márquez's last novel.'
- b. (??/*Me) leí la última novela de García Márquez en
REFL.1SG read.PST.1SG the last novel of García Márquez in
una sentada.
one sitting
'I read García Márquez's last novel in one sitting.'

6.6 Verbs of spending time

Nishida (1994) includes constructions like *pasarse un día entero* 'to spend a whole day' among verbs of consumption, proposing that they only differ from the ones in Section 6.5 in the semantics of the object – while verbs like *comer(se)*, *beber(se)* and *tomar(se)* take a "spatial" object, the verbs in this section take a temporal object. The COSER data attest four different reflexive transitive verbs in this category: *echar(se)*, *tirar(se)* 'to throw', *llevar(se)* 'to take' and *pasar(se)* 'to pass' (see (186)).²¹ That is, two motion verbs (*llevar* and *pasar*) and two verbs of throwing (*echar* and *tirar*).

- (186) a. Allí en Logrosán, allí íbamos nosotros a por abono. Tres
there in Logrosán there go.PST.1PL we to for fertiliser three
días nos echábamos.
days REFL.1PL throw.PST.1PL
'We went there, to Logrosán, for fertiliser. We needed three days for that.'
(La Nava de Ricomalillo, Toledo, COSER 4214)
- b. Antes, madre mía, se llevaban los pobres to'l verano
before mother mine REFL.3 take.PST.3PL the poor all-the summer
trillando, limpiando.
threshing cleaning
'Before, oh my god, poor people, they used to spend the whole summer
threshing, cleaning.' (Valencia del Ventoso, Badajoz, COSER 0728)
- c. Nos pasábamos dos días prácticamente de fiesta haciendo la
REFL.1PL pass.PST.1PL two day almost of party making the
matanza.
slaughter
'We used to spend almost two days celebrating the pig slaughter.'
(Anaya, Segovia, COSER 3701)

21. *Echar(se)* and *llevar(se)* were collected in subcorpus E, while *pasar(se)* and *tirar(se)* were collected in the whole corpus.

- d. Y, claro, se tiran años, y años, y años
 and clear REFL.3 throw.PRS.3PL years and years and years
 de novios...
 of boyfriends
 ‘And, well, they spent years and years and years as boyfriend and girl-
 friend...’ (Portillo de Solórzano, Cantabria, COSER 1219)

Again, Figure 49 shows that north-western varieties essentially reject the presence of the RM in these transitive verbs. Only one reflexive example was found, with the verb *tirarse*. Interestingly, only this verb has specialised the reflexive form for these contexts, while the other three are most frequently unmarked in these temporal contexts.

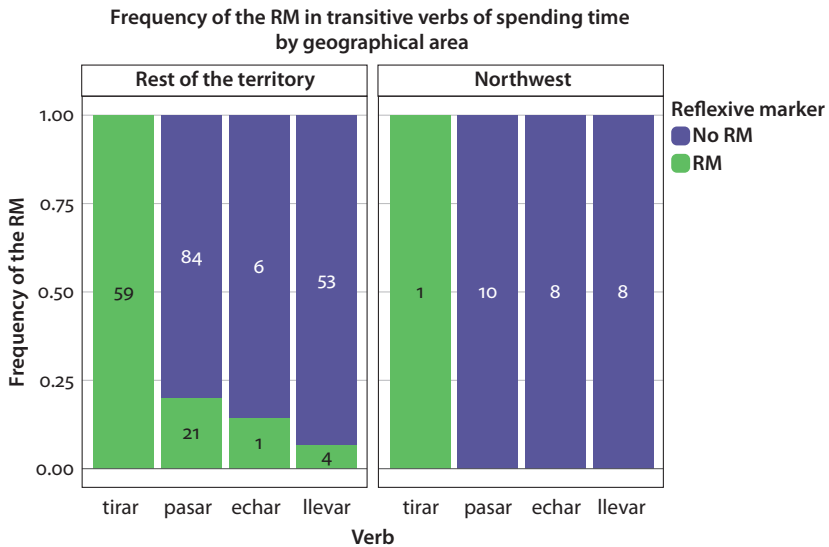


Figure 49. Frequency of the RM with transitive verbs with a temporal direct object

One might wonder if the fact that *tirarse* must be reflexive in this context suggests that it is not a real transitive verb in such examples, but that it derives rather from the anticausative uses of *tirar* with the RM. Whether this is the case or not from the diachronic point of view can only be determined in future historical studies, since it is clear that nowadays the temporal argument is a direct object, since it can be replaced by accusative pronouns:²²

²². As noted in Section 5.7, this is not the case for *esperar(se)* ‘to wait’, which justifies including it with reflexive intransitive verbs, contrary to Nishida (1994).

- (187) Se de alguien que esta noche se la tira
 know.PRS.1SG of someone that this night REFL.3 ACC.3SG throw.PRS.3SG
 haciendo deberes y estudiando
 doing homework and studying
 ‘I know of someone who is going to spend the night doing homework and
 studying.’
 (March 30th 2014. Tweet.)

Although Nishida (1994) includes them in her account, these verbs have little to do with the prototypical reflexive transitive verb presented in the literature and with the aspectual characteristics that the generative literature (included Nishida 1994) ascribe to them. First, they are not accomplishments, but activities. Semantically, they are far more similar to stative verbs like *estar* ‘to be’ than to verbs of consumption. Actually, it is worth recalling that the most frequent context for the marked form *estarse* is its combination with temporal adjuncts (see Section 5.7).²³ Second, their subjects are better characterised as experiencers instead of agents. Finally, despite the few examples found with non-delimited direct objects, there seem to be no restrictions regarding the presence of the RM related with this factor (see Figure 50).

Introspectively, the presence of the RM seems to highlight the involvement of the subject. Objectively, however, this is difficult to prove. My data attest no instances of these verbs in the imperative, which would allow us to check for the association between the RM and agentivity, and do not provide enough examples of inanimate subjects with these verbs as to check whether the animacy of the subject affects the distribution of the RM (see Figure 51).

However, some clues can be found that suggest that the presence of the RM with these verbs has to do with emphasising the experience undergone or the effort made by the subject. On the one hand, the direct object of *tirarse* is very frequently marked as an impressive amount of time by means of intensifiers (such as *todo/a* ‘all’, *hasta* ‘even’, etc.), clear exaggerations, or repetitions (as in (186d) above) – this was the case 37 times out of 58 in the rest of the territory (64%). On the other hand, the distribution of the RM with *pasar(se)* seems to be affected by the presence of other adjuncts besides the temporal direct objects. When only the temporal direct object is present (see (188a)), the interpretation of the subject is less agentive, since “spending time” is something we all do all the time, no matter whether we want to do it or not. In these cases the RM was attested in four cases out of 36 (11%) in the rest of the territory. When other adjuncts (locative, modal, comitative or gerund phrases) are present (as in 188b)), the subject is felt to be more involved in

23. A further similarity is that *estarse* seems only to admit temporal adjectives with a durative interpretation – the same interpretation that the temporal direct objects discussed in this section take. All the examples documented in my data ($N = 37$) had such an interpretation, except for one, which was ambiguous. My random sample of examples of unmarked *estar*, on the contrary, shows both durative (10/17) and punctual (7/17) temporal adjuncts.

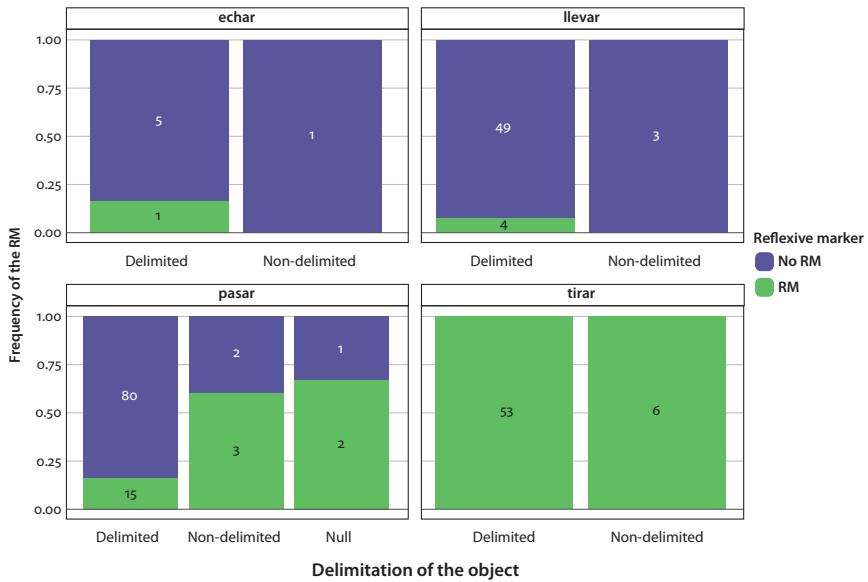


Figure 50. Frequency of the RM with transitive verbs with a temporal direct object by the characteristics of the object in the rest of the territory

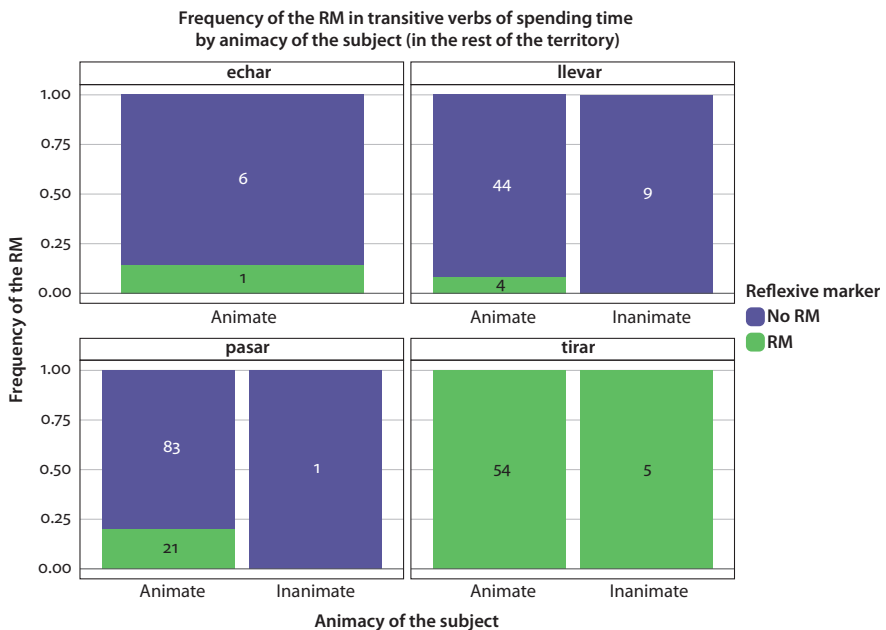


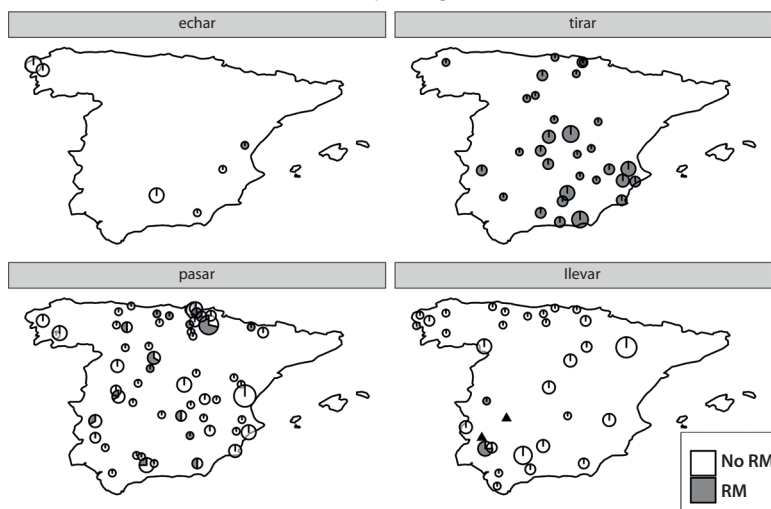
Figure 51. Frequency of the RM with transitive verbs with a temporal direct object by animacy of the subject in the rest of the territory

the depicted action (and the imperative is more acceptable). In these cases, the RM appeared twice as frequently: in 14 out of 64 examples (22%). Although the figures are low, they are very suggestive and could be confirmed with more data: differentiating between types of adjuncts could be also fruitful, since gerund phrases are typically more agentive than locative, modal or committative phrases. My data attest 5/8 (62%) marked cases with gerund phrases, as opposed to 9/56 (16%) marked cases with other adjuncts.

- (188) a. To'l mundo viene a pasar los días esos.
all-the world come.PRS.3SG to spend.INF the days those
'Everyone comes to spend those days.' (Liétor, Albacete, COSER 214)
- b. Me pasaba las noches cosiendo.
REFL.1SG spend.PST.1SG the nights sewing
'I used to spend the nights sewing.' (Azcona, Navarra, COSER 3203)

Finally, it is worth noting that the geographical distribution of the reflexive variants of these verbs is not uniform (see Map 20). The COSER data are too scant to make any generalisations regarding *echar(se)*. As for *tirarse*, it is quite widespread but seems to be especially infrequent in both north-western and north-eastern varieties. *Pasarse*, on the other hand, shows a more restricted distribution, being found only in a central vertical strip of the territory, despite the fact that unmarked *pasar* is found everywhere. Finally, while unmarked *llevar* is found throughout the study area, *llevarse* was only found in Extremadura and Huelva (the two black triangles indicate examples of *llevarse* in locations not included in subcorpus E).

Presence of the RM in transitive verbs of spending time



Map 20. Frequency of the RM with transitive verbs with a temporal direct object in the COSER data

6.7 Verbs of cognition

Another class of transitive verbs which often show the RM but do not fit well with the aspectual requirement of the delimitation of the object are cognitive verbs. Since the attachment of the RM is rather productive in this class, it is worth noting that cognition events are one of Kemmer's (1993) typically middle categories, although she focusses on intransitive verbs. The COSER data document 466 examples of the reflexively marked transitive construction with cognition verbs, where 29 different verbs were documented: *aprender(se)* 'to learn', *calcular(se)* 'to calculate', *callar(se)* 'to be quiet', *chupar(se)* 'to suck', *confundir(se)* 'to confuse', *conocer(se)* 'to know', *creer(se)* 'to believe', *cuidar(se)* 'to think', *dar(se)* 'to give', *dar cuenta(se)* 'to realise', *discurrir(se)* 'to reflect', *encontrar(se)* 'to find', *echar(se) cuenta* 'to imagine', *engañar(se)* 'to deceive', *entender(se)* 'to understand', *esperar(se)* 'to hope', *estudiar(se)* 'to study', *explicar(se)* 'to explain', *hacer(se)* 'to make', *imaginar(se)* 'to imagine', *inventar(se)* 'to invent', *llevar(se)* 'to take', *olvidar(se)* 'to forget', *pensar(se)* 'to think', *recordar(se)* 'to recall', *repasar(se)* 'to review', *saber(se)* 'to know', *suponer(se)* 'to suppose', *temer(se)* 'to fear'. However, most of these verbs were documented only sporadically: fifteen of these verbs were found only once in the reflexive construction and only six verbs were found more than ten times.

From the perspective of the likely origin of the RM, cognition verbs are not a homogenous class. As noted in Section 6.2.1, the RM in *explicarse* is clearly related to indirect reflexive construction with recipient datives. On the other hand, the reflexive form of a verb like *recordar(se)* has a clear analogical model in its synonym verb *acordarse* 'to remember', which is a non-reversible reflexive verb in Spanish and an antipassive marked verb in Galician (see Section 3.2.7). Moreover, *recordar(se)* was also found in the antipassive construction (see Section 3.2.5). The COSER interviews attest only eight instances of reflexive *recordarse* (8/89, 9%) in the transitive construction. These few examples seem to have the same meaning as the unmarked version of the verb (see (189)).²⁴

- (189) Yo me recuerdo que tuve una gran suerte.
 I REFL.1SG remember.PRS.1SG that have.PST.1SG a big luck
 'I remember that I was very lucky.' (Vega de Pas, Cantabria, COSER 1232)

In the following Subsections I will discuss in more detail some of the most relevant reflexive transitive verbs of cognition in the corpus, classifying them in three different categories. I will first address *saber(se)*, *aprender(se)* and *estudiar(se)* (Section 6.7.1). The relevance of this group of verbs of cognition lies in the fact

24. Most instances were collected in the same locality (Almázcara, León, COSER 2637). Actually, this speaker uses the phrase *me recuerdo* systematically, on 29 occasions, most of which were disregarded for being considered pet phrases.

that *saber(se)* is the most frequently mentioned verb of cognition in studies on reflexive transitive verbs. In Section 6.7.2 I turn to *creer(se)* and *imaginar(se)*, since, despite having received less attention in the literature, they are very frequent reflexive transitive verbs of cognition. Finally, in Section 6.7.3 I discuss briefly the case of *dar(se) cuenta*, also a very frequently reflexive construction which stands out in that it relies on a light verb. Interestingly, these verbs show important differences regarding the frequency of the RM:

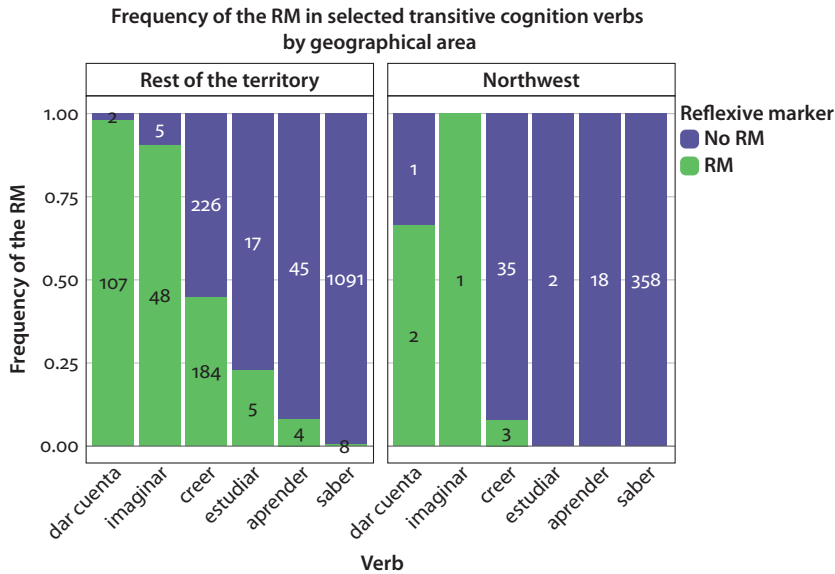


Figure 52. Frequency of the RM with *aprender*, *creer*, *imaginar* and *saber* in the COSER data

Saber(se), *aprender(se)* and *estudiar(se)* were collected exhaustively only in subcorpus E. As there were so few marked cases, instances of these verbs with the RM were collected in the whole corpus: these extra examples, of course, are not included in Figure 52. Absolute uses of *aprender* and *estudiar* (‘to go to school, to study’) were excluded. All the other verbs were collected exhaustively in all the interviews.

Again, the first thing we notice from Figure 52 is the fact that north-western varieties are rather resistant to the presence of the RM with these verbs. However, the verbs that show no reflexive examples in north-western varieties, that is, *estudiar(se)*, *aprender(se)* and *saber(se)*, are only very seldom marked in the rest of the territory. As can be seen in Figure 52, although *saber* is impressively frequent in global terms, the examples of the marked form *saberse* only amount to 0.7% of the cases in the rest of the territory. This is especially noteworthy, since *saber(se)* has

very often been discussed in works on aspectual *se*, while other cognition verbs like *aprender*, *creer* or *imaginar* have received far less attention.²⁵

6.7.1 *Saber(se)* ‘to know’, *aprender(se)* ‘to learn’, *estudiar(se)* ‘to study’

The distribution of the RM shown in Figure 52 above does not support a hypothetical higher agentivity in reflexive forms. While the subject of *saber(se)* ‘to know’ is never an agent (as shown by the resistance of this verb to the imperative form), which might explain why it is the most resistant verb to the RM, the subject of *aprender(se)* ‘to learn’ and *estudiar(se)* ‘to study’ is a clear agent, but they are the second least frequently marked verbs.

A crucial difference that groups *creer(se)* ‘to believe’ and *imaginar(se)* ‘to imagine’ together on the one hand, and *estudiar(se)*, *aprender(se)* and *saber(se)* on the other, is that only the latter can be used with objects that convey skills, such as speaking a language or dancing (see below). The use of the RM in such contexts is not preferred – this does not support an association between agentivity and the RM with these verbs either, since knowledge of a skill implies a higher involvement of the subject than knowledge of a fact. Actually, in some languages, knowledge of (at least some) skills is not conveyed by the correlate of *saber* ‘to know’, but of *poder* ‘can’, showing that they are conceived of as capabilities of the subject:

- (190) a. I can dance.
 b. I know who she is.
 c. አማርኛ አልችልም (Amharic)
 Amarīgna alčīlim
 Amharic can.PRS.NEG.1SG
 ‘I don’t speak Amharic’.
 d. ማን እንደሆነች አላውቅም (Amharic)
 man īndāhonāč alawk’im
 who COMP.be.3SG.FEM know.PRS.NEG.1SG
 ‘I don’t know who she is’.

25. Nishida (1994) does include the example *creerse sus palabras* in her study on aspectual *se*, translating it as ‘believing (all) his words’. In my opinion, the addition of *all* seems not to be particularly justified.

As noted above, *saber(se)* is often mentioned in studies on aspectual *se*, which have argued that while *saber* is a state, *saberse* is either an accomplishment (Sanz & Laka 2002), a quasi-state (Zagona 1996), that it only admits delimited objects (Nishida 1994), or that the RM in *saberse* is a marker of the culmination of the change of state that occurs between not knowing something and knowing something (De Miguel & Fernández Lagunilla 2000). While I have argued elsewhere against this conception (De Benito Moreno 2015, 2021), the data confirm the intuition on which these arguments are based, that is, that *saberse* is preferred (although, again, not strictly) with (some kinds of) delimited direct objects. In fact, the examples given by these authors always present the direct object *la lección* ‘the lesson’. Of the 33 examples of reflexive *saberse* collected in the NE corpus, 25 (76%) refer to knowledge typically learnt (and known) by heart, such as knowledge acquired at school or knowledge of songs, rhymes, etc. (see (191)). However, the RM is not required in these contexts (see (192)) and neither is it exclusively associated with such contexts (see (193)). No examples of *saberse* were found in contexts of knowing a skill.

- (191) Jamás lo he olvidao. Lo tengo aquí,
 never ACC.3SG have.PRS.1SG forget.PTCP ACC.3SG have.PRS.1SG here
me lo sé de memoria.
 REFL.1SG ACC.3SG know.PRS.1SG of memory
 ‘I have never forgotten it. I have it here, I know it by heart.’
 (Villaconejos de Trabaque, Cuenca, COSER 1636)
- (192) Pos si lo sé de memoria, cuántas veces se
 well if ACC.3SG know.PRS.1SG of memory how.many times DAT.3SG
 lo digo a ella.
 ACC.3SG say.PRS.1SG to OBL.3SG
 ‘I know it by heart, how many times I’ve told her!’
 (Moraleja, Cáceres, COSER 1014)
- (193) ¡Qué me sé yo que de ánde eran, aquellos!
 what REFL.1SG know.PRS.1SG I that of where be.PST.3PL those
 ‘What do I know, where were they from!’
 (San Mateo de Gállego, Zaragoza, COSER 4719)

I also collected all forms of marked *aprenderse*, since so few examples were found in subcorpus E. An examination of the examples of *aprenderse* reveals that it takes similar objects as those described above for *saberse*. From 21 examples of *aprenderse*, 14 (67%) refer to pieces of memorised speech or items of knowledge, often learnt at school (see (194)), but, again, the RM is not required in such contexts. Example (194) also illustrates nicely that, although *saberse* and *aprenderse* appear in very similar contexts, using one does not imply using the other. Interestingly, however, two examples of *aprenderse* were found in contexts of learning a skill (see (195)).

- (194) Yo he ido a los sermones de los curas y me
 I have.PRS.1SG go.PTCP to the sermons of the priests and REFL.1SG
los he aprendió. Sí, los sé, claro.
 ACC.3PL have.PRS.1SG learn.PTCP yes ACC.3PL know.PRS.1SG clear
 'I've been to the sermons of the priests and I've learnt them. Yes, I know them,
 of course.' (Povedilla, Albacete, COSER 0222)
- (195) a. Y conducir un coche a la primera me lo
 and drive.INF a car to the first REFL.1SG ACC.3SG
aprendo, pero para conducirlo no valgo.
 learn.PRS.1SG, but for drive.INF-ACC.3SG not serve.PRS.1SG
 'Driving a car, I'll learn how to do it fast, but I'm not good for driving it.'
 (Villacónes de Trabaque, Cuenca, COSER 1636)
- b. Es de Navarra. ¡Y cómo se ha aprendió el
 be.PRS.3SG of Navarra and how REFL.3 have.PRS.3SG learn.PTCP the
andaluz, Dios mío, que yo me creí que
 Andalusian God mine that I REFL.1SG believe.PST.1SG that
 era andaluza!
 be.PST.3SG Andalusian
 'She is from Navarra. How well she speaks Andalusian, my God, I thought
 she was from Andalusia!' (La Atalaya, Málaga, COSER 3002)

Although even scant, the data from *estudiarse* show a similar picture. Of 7 total marked examples, only one does not refer to content that needs to be learnt by heart, but rather to knowledge of a skill (speaking a dialect): interestingly it is the continuation of Example (195b) above (see (196)).

- (196) Esa es que se lo tiene que habé estudio, el
 that is that REFL.3 ACC.3SG have.3SG that have.INF.3SG study.PTCP the
 andaluz.
 Andalusian
 'She must have studied it, Andalusian.' (La Atalaya, Málaga, COSER 3002)

Overall, I believe that these three verbs illustrate quite clearly how the RM expands across contexts and verbs on the basis of analogy, triggering secondary associations with aspectual notions. Although scant, the data suggest that the RM combines most frequently with *estudiar(se)*, less frequently with *aprender(se)*, and least frequently with *saber(se)*. According to their semantics, *estudiar(se)* and *aprender(se)* are the most likely to refer to content learnt by heart (and academic content in general).

Moreover, *estudiar(se)* is the least likely to refer to skills. As shown in (197), *estudiar(se)* requires that the skill is encoded by a noun phrase (rejecting infinitives, contrary to *saber(se)* and *aprender(se)*), triggering the implicature that it is being learnt in an academic context. That is, of these three verbs, *estudiar(se)* is

the one most closely associated with academic contexts, followed by *aprender(se)* and *saber(se)*, whose association with such contexts is far weaker, since they have a broader meaning. Although my data are scant, there are two hints that the use of the RM is more advanced with *estudiar(se)* than with *aprender(se)*: (1) *estudiarse* is more frequent than *aprenderse*; and (2) the tendency to use the RM in the typically reflexive contexts described above is slightly stronger with *estudiar(se)*: in the rest of the territory, subcorpus E attests eight such examples of *estudiar(se)*, five of which (62%) are marked, and eleven examples of *aprender(se)*, four of which (36%) are marked.

- (197) a. {Saber /aprender /estudiar} inglés.
 know.INF learn.INF study.INF English
 ‘To {know (how to speak)/learn/study} English.’
 b. {Saber /aprender a /??estudiar} bailar.
 know.INF learn.INF to study.INF dance.INF
 ‘To {know/learn/??study} how to dance.’
 c. {??Saber /aprender /estudiar} baile.
 know.INF learn.INF study.INF dance.SG
 ‘To {??know/learn/study} dance.’

Of course, although these figures must be taken with caution, they are very suggestive. It seems to me that the association between the RM and these verbs is the product of the association of the RM and the “impressive” effort one must make in order to learn some academic topic by heart, normally in a short period of time. Although learning a skill might be a more agentive event and require more effort overall, it often requires larger periods of times, being less noteworthy in terms of discourse. On the one hand, the connection of the RM with contexts of this kind provides a reasonable connection with emphatic autobenefactives. On the other hand, it explains why it is more common with *estudiar(se)* and *aprender(se)* and why reflexive *saberse* is so restricted – its appearance with this verb is the product of both paradigmatic and syntagmatic analogy with *estudiar(se)* and *aprender(se)*. The presence of the RM does not respond to strict syntactic or semantic requirements, such as a given aspectual behaviour. On the contrary, its association with some contexts is only statistical (i.e. it is never compulsory) and it triggers some implicatures that can be used in more atypical contexts, as can be seen in the examples (195b) and (196), where the Andalusian speaker is clearly impressed by the non-Andalusian subject’s competence in Andalusian. Again, I am proposing here a diachronic hypothesis on the basis of synchronic data – only a historical study can confirm or reject such hypothesis.

6.7.2 *Creer(se)* ‘to believe’, *imaginar(se)* ‘to imagine’

I now turn to *creer(se)* ‘believe’ and *imaginar(se)* ‘to imagine’, two verbs whose semantics clearly differ from the semantics of *saber(se)* ‘to know’, *aprender(se)* ‘to learn’ and *estudiar(se)* ‘to study’ and which do not take the kind of direct objects that are preferred by the latter. Moreover, *creerse* and *imaginarse* admit both sentential and null objects, all of which would be rejected if the RM had an aspectual function (Nishida 1994):

- (198) a. Las mozas cuando me ven se creen que
the girls when ACC.1SG see.PRS.3PL REFL.3 believe.PRS.3PL that
no estoy casao, como no ven el anillo.
not be.PRS.1SG marry.PTCP like not see.PRS.3PL the ring
‘Girls think I’m not married when they see me, because they don’t see the ring.’
(Burgohondo, Ávila, COSER 604)
- b. Pero con cualquiera no se bailaba, ¿eh?, no, no te creas.
but with anyone not REFL.3 dance.PST.3SG huh no not REFL.2SG
believe.PRS.2SG
‘But you didn’t dance with just anyone, no, don’t think that.’
(San Francisco de Olivenza, Badajoz, COSER 723)
- c. Usté imáginese que beben más y les
you imagine.IMP.3SG-REFL.3 that drink.PRS.3PL more and DAT.3PL
da por venir a mi casa.
give.PRS.3SG for come.INF to my house
‘Imagine that they drink more and they feel like going to my place.’
(Navalmoral de la Mata, Cáceres, COSER 1015)
- d. [la máquina de llenar los chorizos] Es la misma de siempre,
be.PRS.3SG the same of always
de cuando yo me criaba, así que imáinate.
of when I REFL.1SG grow.PST.1SG imagine.IMP.2SG-REFL.2SG
‘[The machine for filling the sausages] is the same one I always had, since I was growing up, so imagine.’
(Valencia del Ventoso, Badajoz, COSER 0728)

Introspectively, the difference between *creerse* and *creer* has to do with the veracity of the beliefs, *creerse* being associated with false beliefs (an appropriate translation in English would be ‘to buy’), while *creer* is agnostic towards the veracity of the belief (see also De Benito Moreno 2015; Anvari, Maldonado & Soria Ruiz 2019). This is especially clear in examples such as (199a), where the speaker asks rhetorically about the beliefs of the interviewer and specifies his expected answer, the implicature being that the answer is wrong. That is, by asking the rhetorical question

the speaker aims to contradict the hearer. Actually, *creer(se)* was found in thirty questions with the interrogative pronoun *qué* ‘what’ in the COSER data (in all the territory), all of which took the RM and bore the implicature that the hearer was wrong in their beliefs.²⁶ This does not mean that *¿qué crees?* ‘what do you think?’, without the RM, is ungrammatical – on the contrary, it is completely acceptable. However, such a question is not rhetorical and really asks about the beliefs of the hearer – the nature of the COSER interview does not foster such a kind of questioning, since the informant is the main speaker and what matters is their opinion. It is very rare that they show interest in the lives or opinions of the interviewers. Overall, in the rest of the territory, the reflexive form *creerse* is almost as frequent as its unmarked forms – it appears in 184 of 410 cases (45%).

- (199) [antes] ¿Qué te crees, que hay un mes de
 what REFL.2SG believe.PRS.2SG that have.PRS.3SG a month of
 vacaciones como ahora o qué?
 holidays like now or what
 ‘[before] Do you really you think that there was a month of vacation, like
 nowadays?’ (Luzuriaga, Álava, COSER 0107)

Let us turn to *imaginar(se)*. As seen in Figure 52 above, the marked form is far more frequent than the unmarked one – 48 of 53 cases (91%) show the RM in the rest of the territory. There is no clear difference between the marked and unmarked forms. Of course, according to its own meaning, *imaginar(se)* implies that its objects is not true, but imagined.

Although more infrequent in global terms, *suponer(se)* ‘to suppose’ shows a similar situation to *imaginar(se)*, with no clear semantic difference between the marked and the unmarked form and the implication in all cases that its object is not true. The presence of the RM is as frequent as with *creer(se)*, appearing in eight of 17 cases (47%) in the whole territory (only one of those examples was found in the north-western varieties and it showed the RM). Again, this verb allows the combination of the RM and sentential objects:

- (200) Supónte que la finca es esta y llegas aquí
 suppose.IMP.2SG-REFL.2SG that the farm is this and arrive.PRS.2SG here
 [A-Inn] en mulo.
 [A-Inn] in mule
 ‘Suppose that this one is the farm [we’re talking about] and you arrive here on
 a mule.’ (Zas, La Coruña, COSER 2403)

26. Detecting whether such an implicature is intended in contexts other than questions in corpus data is often too difficult, preventing a quantitative analysis in other contexts.

In my opinion, the presence of the RM with these cognition verbs is not related to autobenefactives, but to reflexive constructions with a predicative (see (201a)). In such cases it is clearly implied that the subject does not have the quality conveyed by the predicative, that is, that the subject is wrong in their belief. Thus, these contexts might have been the origin of the association between the RM and the notion of wrong beliefs. It must be noted that these reflexive constructions are not indirect, but rather direct constructions, since the RM is the direct object. However, the fact that they have a second argument (the predicative) and that they alternate with a(n) (unmarked) ditransitive construction such as the one in (201b) can explain the extension of the RM to reflexive transitive constructions.

- (201) a. Se {cree /imagina /supone}
 REFL.3 believe.PRS.3SG imagine.PRS.3SG suppose.PRS.3SG
 muy interesante.
 very interesting
 ‘S/he believes/imagines/supposes herself/himself to be very interesting.’
- b. {Cree /imagina /supone} que es
 believe.PRS.3SG imagine.PRS.3SG suppose.PRS.3SG that is
 muy interesante.
 very interesting
 ‘S/he believes/imagines/supposes herself/himself to be very interesting.’

6.7.3 *Dar(se) cuenta* ‘to realise’: Cognition light verb structures

Syntactically, *dar(se) cuenta* ‘to realise, lit. to give account’ belongs in a larger group of reflexive transitive verbs that encode cognition events, all of which consist of a light verb and a noun: *dar(se) maña* ‘to be skillful, lit. to give skill’, *echar(se) cuenta* ‘to imagine, lit. to throw account’, *hacer(se) una idea* ‘to get a basic idea, lit. make an idea’. These are highly reminiscent of some of the uses described above under the “indirect detransitivised constructions” label (such as *armar(se) un lío* ‘to get confused’ or *darse una idea* ‘to get a basic idea’, see Section 6.2.4). The main difference between the examples considered here and those considered in Section 6.2.4 is that the former do not have a non-correferential counterpart (see (202)). Nevertheless, many of the indirect detransitivised verbs have a rather abstract meaning that can function as the source of the analogy for the verbs in this section.

- (202) a. Lo leeré bien, a ver si {me /??te} hago
 ACC.3SG read.FUT.1SG well to see.INF if REFL.1SG DAT.2SG do.PRS.1SG
una idea.
 an idea
 ‘I’ll read it carefully in order to {get/??give you}an idea.’

- b. No {me/ ^{??}te} di cuenta de que Juan ya
not REFL.1SG DAT.2SG give.PST.1SG count of that John already
había llegado.
have.PST.3SG arrive.PTCP
'I didn't realise (??you) that John had already arrived.'

The COSER data do not attest many examples of these structures and the few examples of *darse maña* ($N = 1$), *echarse cuenta* ($N = 2$) and *hacerse una idea* ($N = 3$) are all marked with the RM. *Dar(se) cuenta* is a far more frequent structure in the corpus, with a total of 112 examples, most of which appear in the rest of the territory ($N = 109$). Only three examples of this construction were unmarked, one in north-western varieties (1/3, 33%) and two in the rest of the territory (2/109, 2%):

- (172) [lobos] Los pastores cuando dan cuenta lo mejor por un
the shepherds when give.PRS.3PL count the better for one
lao y se la hicieron por otro.
side and DAT.3PL ACC.3SG do.PST.3PL for other
‘When the shepherds realise [that there are wolves], maybe they go check one
area and the wolves have done something in another area.’
(Ledantes, Cantabria, COSER 1212)

6.8 Corporal processes

A number of examples of reflexive transitive verbs are instances of *hacer* ‘to make, to do’ (also used as a light verb) in scatological contexts. In these structures, the light verb *hacer* is followed either by a noun that refers to corporal fluids (*caca* ‘poo’, *pis/pipí* ‘piss’) or by a euphemistic prepositional phrase which encodes the same meaning (*de vientre*, ‘bowel movement’, lit. ‘of gut’).

These verbal phrases are periphrastic synonyms for intransitive verbs that also encode scatological processes, such as *cagar(se)* and *mear(se)* (see Section 5.6), and which are clearly the source of the analogy. Actually, the transitive constructions show a similar picture regarding the lack of control of the subject in imperfective tenses to the one discussed above for the equivalent intransitive verbs. Accordingly, the presence of the RM allows for encoding the incipency of the corporal process (see (203)). Of 22 examples of these transitive constructions documented in the COSER data (all found in the rest of the territory), 16 (73%) took the RM. All 16 appeared in contexts where the subjects shows reduced control over the action (16/17, 94%). The five examples where the subject controlled the action were unmarked. All in all, this situation is very similar to the one described for intransitive corporal processes. It seems clear that they should all be considered under the same class of the RM, regardless of the valency of the verb.

- (203) Había un muchacho, que se estaba haciendo de vientre y
 have.PST.3SG a boy that REFL.3 be.PST.3SG doing of gut and
 no se atrevía a pedirle permiso al maestro para
 not REFL.3 dare.PST.3SG to ask.INF-DAT.3SG permission to.the teacher for
 ir a hacer sus necesidades...
 go.INF to do.INF his needs
 ‘There was a boy, who needed to poo and didn’t dare to ask permission to the
 teacher to go poo.’ (Mas de los Mudos, Valencia, COSER 4301)

6.9 Verbs of finding and forgetting: Lack of intentionality

In this section I focus on reflexive transitive verbs of finding and forgetting, such as *encontrar(se)* ‘to find’ and *dejar(se)* ‘to leave’, where the presence of the RM seems to be associated with unintentional events.²⁷ Thus, these verbs do not show the meaning typically associated with autobenefactive constructions, since the presence of the RM in autobenefactive constructions tends to side with higher agentivity and intentionality, as already noted. In Maldonado’s words: “[t]he inherent benefactive meaning of the *se* clitic contributes the intentionality, which is directed towards one of the participants and is lacking in most bivalent verbs” (Maldonado 1999: 208, my translation). As shown in Figure 53, the presence of the RM with these verbs is almost wholly restricted to the rest of the territory, where *encontrarse* is far more frequent than *dejarse*.

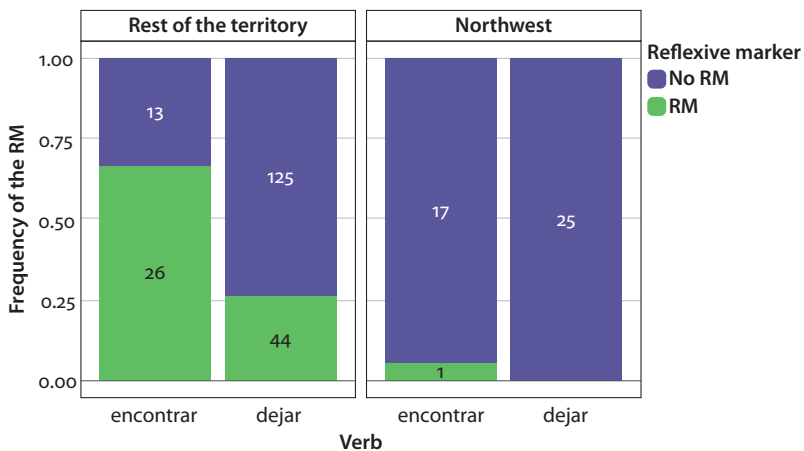


Figure 53. Frequency of the RM in transitive verbs of finding and forgetting by geographical area

27. Both verbs were collected exhaustively in subcorpus E. *Encontrar(se)* might also have a reciprocal (‘to meet’) or a cognition (‘to find, to believe’) meaning, none of which are included in this section. Similarly, I disregarded the examples of *dejar(se)* meaning ‘to let, to allow’.

Introspectively, the marked form *encontrarse* seems to be specialised in conveying unplanned or unintentional (and hence non-agentive) events of finding, where the agent was not purposefully looking for whatever they found (as in (204)). When *encontrar(se)* appears in such contexts in the rest of the territory, it is overwhelmingly marked: 23/26 (88%). In contexts where the subject was looking purposefully for the object, the presence of the RM was not only rarer (3/13, 23%) but also seems to be connected to the notion of unplanned finding, since in the three marked examples, the subject was looking for the direct object (a person) but found them – unexpectedly – dead. That is, while the notion of unplanned finding is present in all marked examples, the unmarked variant is used to express both planned and unplanned events.²⁸

- (204) A una pareja me he encontrao yo al pie de la
 to a couple REFL.1SG have.PRS.1SG find.PTCP I to.the foot of the
 sierra esta.
 mountain.rage this
 ‘I’ve run into a couple, at the foot of the mountains.’

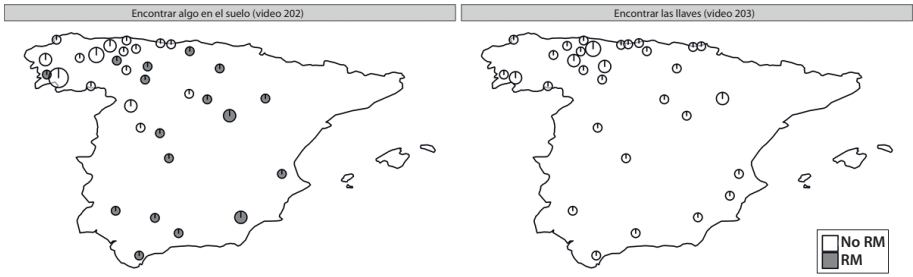
(Villaconejos de Trabaje, Cuenca, COSER 1636)

The questionnaire data confirm both the tendency of *encontrarse* to be used for unplanned events and the geographical differences found above. The questionnaire includes two videos designed to obtain examples of *encontrar(se)*. Video 202 shows a woman walking, who sees something on the floor and picks it up. Video 203 shows the same woman desperately looking for something in a wooden case and finally finding her keys. As shown in Map 21, no instance of marked *encontrarse* was produced to describe video 203, illustrating that situations where the event of finding is intentional are highly resistant to the presence of the RM. On the contrary, unintentional findings such as the one depicted in video 202, are typically described by *encontrar* in north-western varieties and by *encontrarse* in the rest of the territory.

Interestingly, my data also attest sporadic examples of other verbs (such as *tropezar* ‘to stumble’ and *ver* ‘to see’) in the reflexive transitive construction with a similar meaning of unplanned finding.²⁹

28. Due to the low frequency of inanimate subjects (5/39, 13% for *encontrar(se)* and 5/169 (3%) for *dejar(se)*) and of non-delimited direct objects (7/39, 18% for *encontrar(se)* and 15/168 (8.9%) for *dejar(se)*) with these verbs in the rest of the territory, it is difficult to assess whether there are relevant differences in the behaviour of the RM depending on these parameters.

29. *Ver(se)* has often been mentioned in work on aspectual *se* (see Nishida 1004, De Miguel & Fernández Lagunilla 2000; Campanini & Schäfer 2011) because it can also appear in reflexive transitive constructions reminiscent of consumption verbs (*ver(se) la película* ‘to watch the movie’). However, due to the characteristics of my corpus, no examples were found in my data.



Map 21. Frequency of the RM with *encontrar(se)* in videos 202 and 203 in the questionnaire

- (205) a. Parece que estamos en un campo, venga a
 seem.PRS.3SG that be.PRS.1PL in a field come.SBJV.3SG
tropezate personas, personas, personas...
 stumble.INF-REFL.2SG people people people
 ‘It is like being in the field, running into more and more people.’
 (Canredondo, Guadalajara, COSER 1902)
- b. Y llego yo allí y me veo a la señora
 and arrive.PRS.1SG I there and REFL.1SG see.PRS.1SG to the lady
 en la cama.
 in the bed
 ‘And I arrive there and find the lady in bed.’
 (Yebra, Guadalajara, COSER 1921)

A similar meaning split for *dejar(se)* as that found in *encontrar(se)* is expected, since *dejarse* has been claimed to convey unintentional or unplanned events of leaving, as in examples like (206), where its meaning is closer to ‘to forget’ (Cano 1981; Armstrong 2013).

- (206) a. Me dejé las llaves en la casa (#a propósito).
 REFL.1SG leave.PST.1SG the keys in the house to purpose
 ‘I left the keys at home (#on purpose).’ (Armstrong 2013)
- b. Juan (*se) dejó el libro en la biblioteca él.
 John REFL.3 leave.PST.3SG the book in the library he
 ‘John left the book in the library himself.’
 (De Miguel & Fernández Lagunilla 2000)

My data confirm this claim only partially. All 24 examples where *dejar(se)* conveyed an unplanned event of leaving in the rest of the territory show the RM (see (207a)). Regarding examples that convey planned events of leaving, only 20 of 145 (14%) were marked (see (207b, c)), showing that the unmarked form is greatly preferred in these contexts. That is, my data suggest that *dejar* is always agentive, while *dejarse* may convey both agentive and unintentional events.

- (207) a. Que por poco me dejo la blusa allí en la puerta
 that for little REFL.1SG leave.PRS.1SG the blouse there in the door
 ‘I almost forgot the blouse there in the door.’
 (Barrax, Albacete, COSER 0204)
- b. Y ese se dejó a la mujer y se fue
 and that REFL.3 leave.PST3SG to the wife and REFL.3 go.PST3SG
 con una.
 with one
 ‘That one left his wife for another women.’
 (Povedilla, Albacete, COSER 0222)
- c. Yo antes de irnos pa abajo me dejo la
 I before of go.INF-REFL.1PL to down REFL.1SG leave.PRS.1SG the
 mesa arreglá.
 table fix.PTCP
 ‘I leave the table ready before we leave.’ (Bacares, Almería, COSER 0404)

That is, in contrast to *encontrar(se)*, where the marked form is the one whose meaning is more restricted, the behaviour of *dejar(se)* points to a higher degree of routinisation of the RM with this verb, according to which it is already obligatory in the presumably original contexts, i.e. unintentional events. It seems that the use of the RM with *dejar(se)* in planned events of leaving is geographically marked, being especially frequent in eastern varieties (see Map 22). Unfortunately, unplanned events of leaving with *dejar* were not documented in the whole territory and were actually more frequent in the east, so it is difficult to know whether this geographical distribution is an indication that these are two different kinds of uses of the RM.

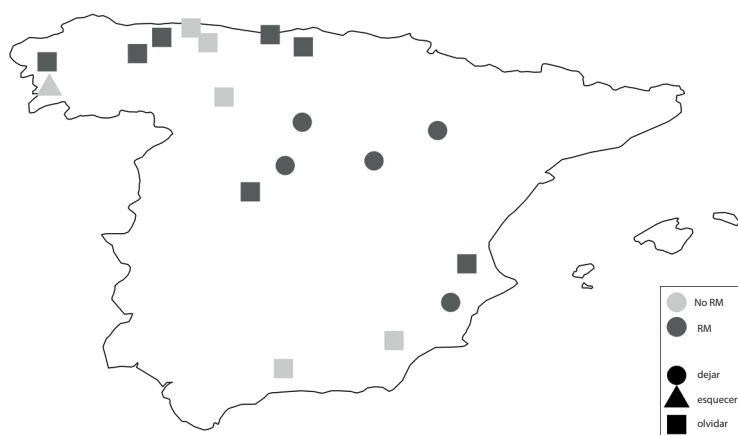


Map 22. Frequency of the RM with *dejar(se)* depending on the intentionality of the subject

Again, my data attest sporadic examples of other reflexive transitive verbs with a similar meaning of intentional leaving, such as *olvidar(se)* ‘to forget’. Moreover, video 76 in the questionnaire showed a man searching his pockets and unable to open a door, since his pockets are empty – he has forgotten his keys. This video

documented three different verbs (*dejar(se)* ‘to leave’, *olvidar(se)* ‘to forget’ and *esquecer* ‘gal. to forget’) in the transitive construction. Map 23 shows that the presence of the RM is preferred in these constructions and that it is even found in north-western varieties.

- (208) Y cuando recién me he confesado: “¡Oy! ¡Que
and when recently REFL.1SG have.PRS.1SG confess.PTCP oops that
me he olvidao aquello!”
REFL.1SG have.PRS.1SG forget.PTCP that
‘And just after having been to confession: “Oops, I’ve forgotten something!”’
(Tramaced, Huesca, COSER 2229)



Map 23. Presence of the RM in video 76 (*Olvidar las llaves* ‘to forget the keys’)

The association of the RM with unintentional events suggests that these verbs are different from those discussed in Section 6.4 to 6.6 and that they are not related to autobenefactive constructions. On the contrary, there are reminiscent of some of the indirect detransitivised constructions discussed in 6.2.4, where the RM is associated with unintentional events. A number of connected constructions depicting cognition events were also discussed in Section 6.7.3 – we must bear in mind that the meaning of *dejarse* with unintentional events is close to cognition verbs, since its meaning is similar to ‘to forget’ – indeed, *olvidar(se)* ‘to forget’ also participates in this construction, as just mentioned. While *encontrar(se)* is not semantically related to cognition events, unintentional events of finding are semantically related to unintentional events of leaving and forgetting objects (since they depict the exact opposite situation), which provides a credible analogical path for the association of the RM with this verb.

6.10 Transitive reflexive verbs of translational motion

Reflexive transitive verbs of movement have typically been ignored in work addressing reflexive transitive verbs (even Nishida (1994), who gives the most detailed enumeration and classification of reflexive transitive verbs, overlooks them). However, they are not at all rare and I collected 330 instances of reflexive transitive verbs of movement in the COSER interviews, of nine different verbs: *bajar(se)* ‘to go/take down’, *dar(se) la vuelta* ‘to turn around’, *dar(se) una {vuelta/paseo}* ‘to go for a walk’, *llevar(se)* ‘to take’, *pasar(se)* ‘to pass, to cross’, *pegar(se) un paseo/una caminata* ‘to go for a walk’, *recorrer(se)* ‘to cover a distance’, *subir(se)* ‘to go/take down’ and *traer(se)* ‘to bring’. Because most of these verbs are documented in the reflexive transitive construction only sporadically (less than five times each, except for *darse una {vuelta/paseo}*, which was found 13 times), in what follows I will focus on *llevar(se)* and *traer(se)*, whose reflexive form was far more frequently documented. The behaviour of the RM with these verbs seems to be clearly related to its behaviour with intransitive verbs of translational motion.

Llevar and *traer* are the two most frequent reflexive transitive verbs of movement documented in the COSER interviews and showed the RM on 201 and 98 occasions respectively. As shown in Figure 54, the relative frequency of these verbs is very low in north-western varieties. In rest of the territory, *llevarse* was almost three times as frequent as *traerse* (200/684, 29.2% and 95/836, 11.4%, respectively). *Llevar(se)* and *traer(se)* are semantically very similar and differ from each other by virtue of a deictic nuance – while *traer(se)* implies that the movement is directed towards the speaker, *llevar* has the opposite implication. On the one hand, this means that *llevar(se)* is more frequent than *traer(se)*, for its meaning is less restricted. On the other hand, the relationship between the two verbs parallels the one between *ir(se)* and *venir(se)* – while *venir(se)* encodes a movement directed towards the speaker, *ir(se)* shows the opposite implication. Crucially, Figure 54 shows that differences in the productivity of the RM with *llevar(se)* and *traer(se)* parallel those with *ir(se)* and *venir(se)*.

Interestingly, the animacy of the subject does not seem to have an effect in the distribution of the RM with *llevar(se)* and *traerse* (or not in the expected direction – see Figure 55, but note the low figures with inanimate subjects). With some translational motion intransitive verbs, the reflexive variant was less frequent with inanimate subjects (see Section 5.4). This was not the case for *ir(se)*, however, where no effect of the animacy of the subjects was found. Similarly to their intransitive counterparts, the imperative seems to favour the presence of the RM (see Figure 56), although the data are again scant.

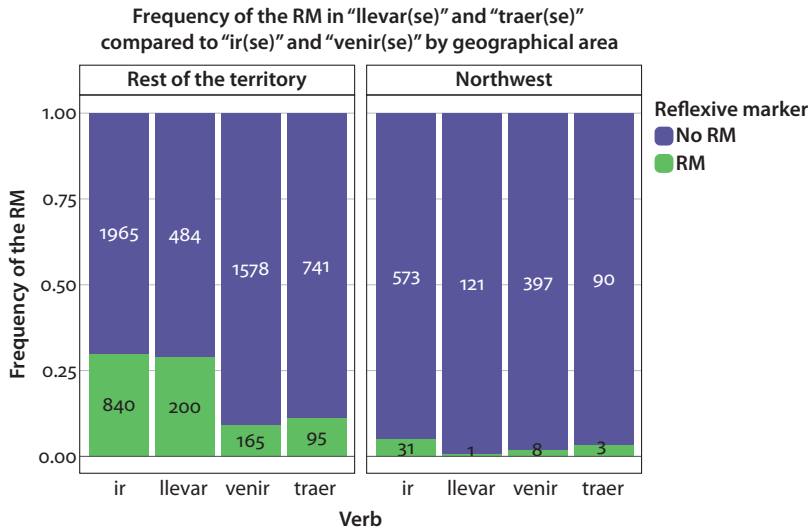


Figure 54. Frequency of the RM with *llevar(se)* and *traer(se)* compared to *ir(se)* and *venir(se)* in the COSER data

Llevar was collected only in subcorpus E, as it was more frequent, while *traer* was collected in all the COSER interviews.

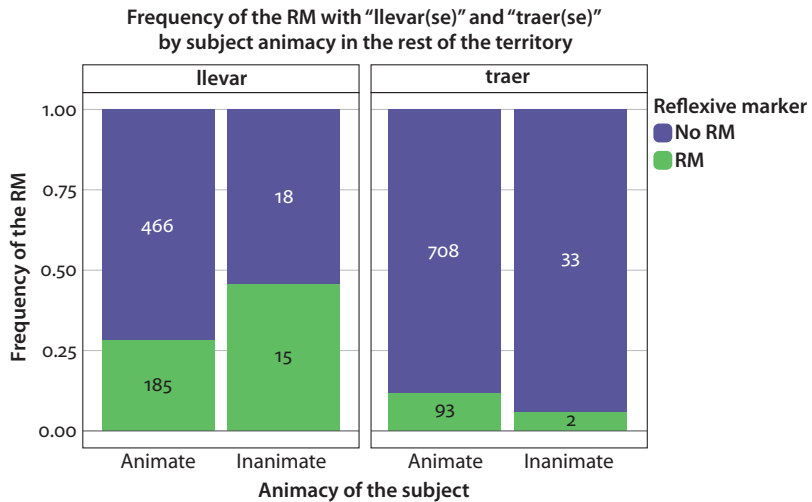


Figure 55. Frequency of the RM with *llevar(se)* and *traer(se)* by animacy of the subject in the rest of the territory

In north-western varieties, the six marked examples have an animate subject.

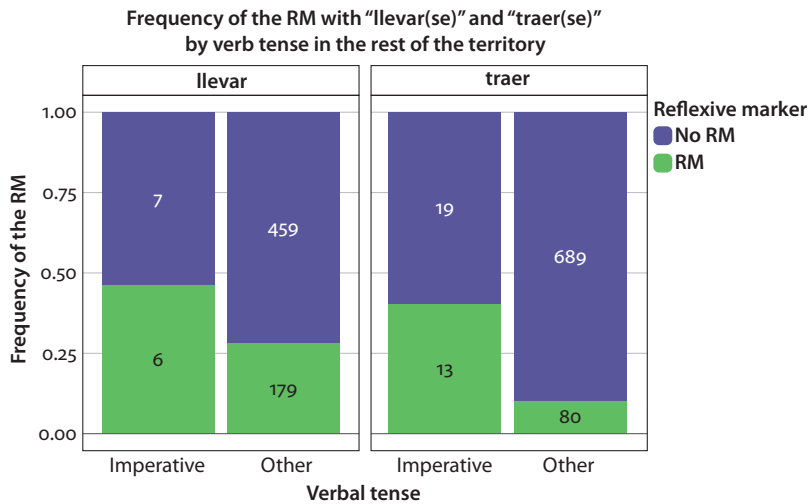


Figure 56. Frequency of the RM with *llevar(se)* and *traer(se)* by verbal tense in the rest of the territory

Let us now turn to one of the aspects that was discussed in Section 5.4 for intransitive reflexive verbs, namely, the different locative phrases that these verbs take. For *llevar(se)*, most contexts show similar frequencies of the RM (see Figure 57). It is tempting to see a link between the higher frequency of the RM with source locative phrases with the preposition *de* and the behaviour of *ir(se)*, but the numbers are rather low in this category and do not allow for establishing such a parallel. A larger sample would serve to establish the degree of the connection between the two verbs.

As regards *traer(se)*, Figure 58 shows that the RM is not favoured by explicitly coded source locative phrases (which are far more frequent here than in the case of *llevar(se)*) and that the RM is more common when a goal locative phrase is encoded. The distribution of the RM with *traer(se)* by type of locative complement is very similar to the distribution of the RM with its intransitive counterpart *venir(se)* (see Figure 58).

An important difference between transitive and intransitive verbs of translational motion concerns the presence of a dative. In Section 5.4 it was noted that the marked form *irse* is the only one found in the COSER data with datives that encode either the (affected) source of a movement, the (affected) cause of a physical movement, or an affected possessor with *ir(se)*. Moreover, many reflexive intransitive verbs show a positive association with the presence of a dative. However, the presence of a dative shows a very strong effect on the distribution of the RM for both *llevar* and *traer*, but in the other direction completely: the presence of the

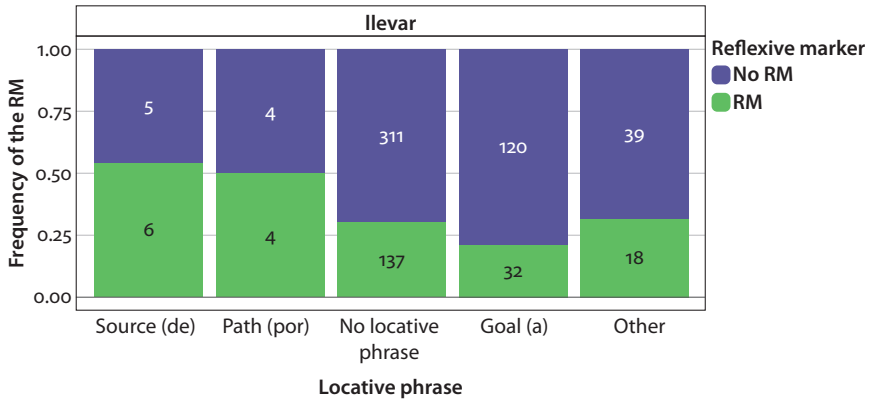


Figure 57. Frequency of the RM with *llevar*(se) by locative phrase in the rest of the territory

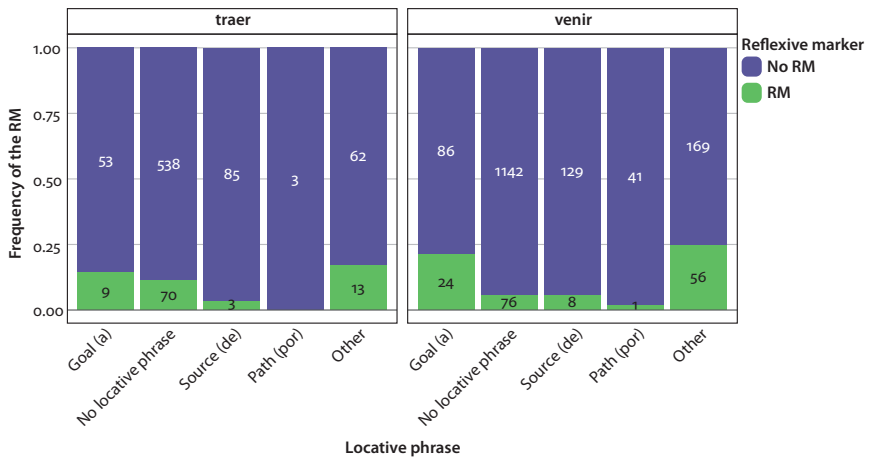


Figure 58. Frequency of the RM with *traer*(se) and *venir*(se) by locative phrase in the rest of the territory

dative seems to preclude the possibility of the appearance of the RM (see Figure 59), despite the fact that *llevar*(se) might take datives that encode the (affected) source of a movement (and indeed did so in ten cases), as in (209):³⁰

30. The most common type of datives that appear with *llevar*(se) ($N = 34$) and *traer*(se) ($N = 106$) are recipient datives, which are not compatible with the intransitive counterparts *ir*(se) and *venir*(se).

- (209) ¿Colchones de lana? Sí. Eso ya te llevan, los colchones
mattresses of wool yes that already DAT.2SG take.PRS.3PL the mattresses
de lana ya te los llevan, no sé lo que
of wool already DAT.2SG ACC.3PL take.PRS.3PL not know.PRS.1SG the that
harán con ellos.
make.FUT.3PL with OBL.3PL
‘Wool mattresses? Yes, if you have them, they take them away, they do take the
wool mattresses away, I don’t know what they do with them.’
(Menagaray, Álava, COSER 0109)

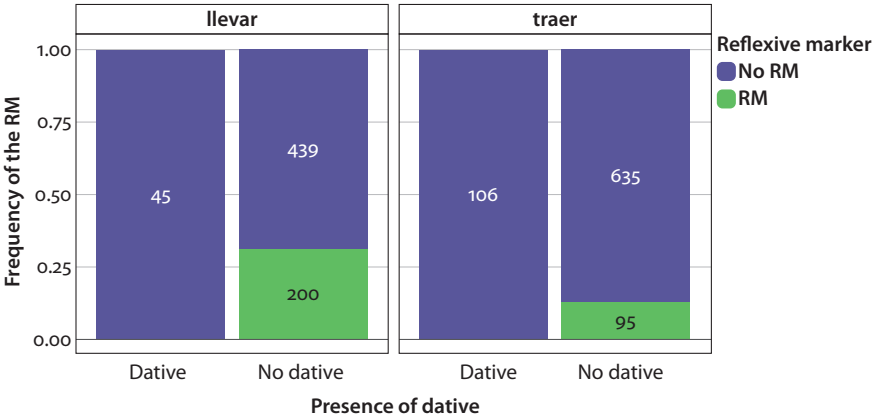


Figure 59. Frequency of the RM with *llevar(se)* and *traer(se)* by presence of a dative in the rest of the territory

That is, the presence of a dative affects the distribution of the RM differently, depending on whether the verb is transitive or intransitive. In my opinion, this is related to the structural model of these constructions. Reflexive intransitive verbs are modelled on middle constructions, where the RM indicates a reduction of valency and is originally an accusative clitic. Therefore, it can combine easily with dative participants. On the contrary, reflexive transitive verbs are modelled on indirect reflexive constructions, where the RM is originally a dative clitic and hence shows a high resistance to combining with other dative participants. Note, however, that the presence of a dative does not completely rule out the presence of the RM, since a few examples were found in subcorpus NE:³¹

31. When the direct object is also a clitic, one might think that it is the cluster of three different clitics that prevents the presence of the RM. However, the majority of examples with a dative clitic showed a non-pronominal object – 36/43 (84%) in the case of *llevar* and 69/106 (65%) in the case of *traer*, so this circumstance cannot alone be responsible for the absence of the RM when a dative is present.

- (210) [vestido] El de mi madre, que se nos lo llevaron.
 the of my mother that REFL.3 DAT.1PL ACC.3SG take.PST.3PL
 ‘My mother’s [dress], that they took away.’ (Bandaliés, Huesca, COSER 2207)

A generalised linear mixed model based on the data of the rest of the territory (see Table 37) confirms that the verb being in the imperative is a factor that favours the presence of the RM and that *traer(se)* takes the RM in a far lower proportion than *llevar(se)*. Having an animate subject seems to slightly disfavour the presence of the RM. Unfortunately, adding the presence of a dative as another factor prevents the model from properly converging, so this factor was left aside.

Table 37. Generalised linear mixed model. Presence of the RM in the rest of the territory depending on the animacy of the subject [reference level: Inanimate], the verb tense (imperative vs. other) [reference level: Other], the lexeme [reference level: *llevar*] (fixed factors) and the place (random factor)

Effect	Term	Estimate	Std. error	Statistic	Conf. low	Conf. high
fixed	(Intercept)	-0.755	0.38	-1.985	-1.5	-0.01
fixed	Subject animacy: Animate	-0.748	0.353	-2.121	-1.439	-0.057
fixed	Verbal tense: Imperative	1.053	0.371	2.838	0.326	1.781
fixed	Verb: <i>traer</i>	-1.18	0.182	-6.499	-1.536	-0.824
random (COSERID)	sd (Intercept)	1.267	NA	NA	NA	NA

Overall, the parallels between the distribution of the RM with *llevar(se)* and *traer(se)* on the one hand, and semantically similar intransitive motion verbs like *ir(se)* and *venir(se)* on the other, are clear. This suggests a connected origin for all these constructions, as also proposed for reflexive transitive and intransitive verbs conveying corporal processes (see Section 6.8). The intransitive-transitive distinction, however, is not trivial and has a clear effect on the co-occurrence of RMs and dative pronouns.

For other verbs of translational motion in the reflexive transitive construction, this parallel behaviour with their intransitive counterparts is not found. This is the case, for instance, with transitive *bajar(se)* ‘to take down’ and *subir(se)* ‘to take up’. The first thing that must be noted is that they are rather infrequent compared to their intransitive counterparts (see Figure 60).

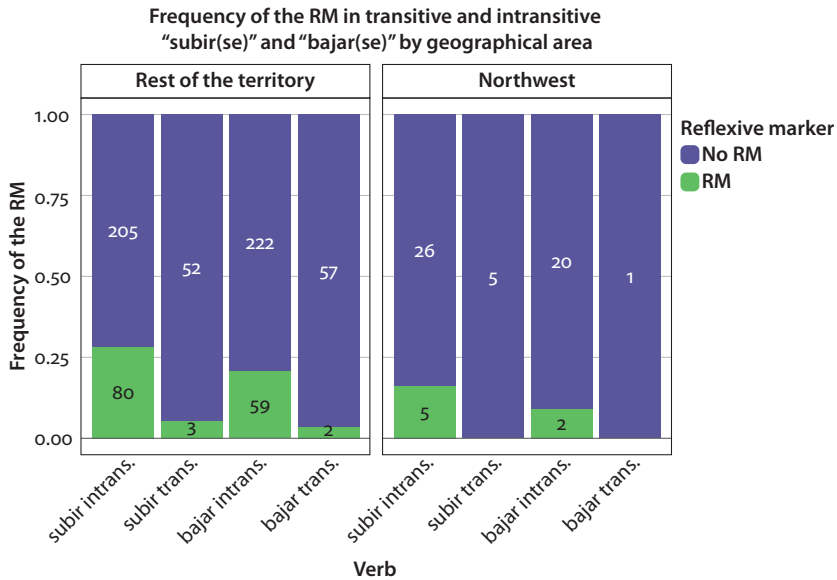


Figure 60. Frequency of the RM with transitive and intransitive *subir(se)* and *bajar(se)* in the COSER data

However, it must be noted that the contextual associations of the RM with intransitive *subir(se)* and *bajar(se)* was found to be of a different nature from verbs like *ir(se)* and *venir(se)*, since it is related to a higher involvement of the subject. This results in the higher frequency of the RM with one specific kind of motion. This difference between two types of movements is not found in transitive *subir(se)* and *bajar(se)* and hence cannot influence the frequency of the RM. When their direct objects are the moved entity, their meaning simply conveys ‘taking something up/down’³² (see (211)). That is, in these transitive uses they depict movement along a path, which is the type of movement that shows higher resistance to the attachment of the RM with their intransitive counterparts. Such a difference might account for the asymmetrical distribution of the RM with transitive and intransitive *subir(se)* and *bajar(se)*.

32. When their direct object is the path of the movement, however, they seem to be connected to the uses of the RM found with consumption (and some cognition) verbs and will be briefly discussed in Section 6.10.1.

- (211) a. Como también se va a Huesca, pues se
 like too REFL.3 go.PRS.3SG to Huesca well REFL.3
baja mucha gente las cosas.
 take.down.PRS.3SG many people the things
 ‘Since they go to Huesca, a lot of people take their things there.’
 (Tramaced, Huesca, COSER 2229)
- b. Se le olvidó a mi madre de subirse
 REFL.3 DAT.3SG forget.PST.3SG to my mother of take.up.INF-REFL.3
 las tijeras y le decía a mi abuela: “Tía, me
 the scissors and DAT.3SG say.PST.3SG to my grandmother aunt DAT.1SG
 subas las tijeras”.
 take.up.SBJV.2SG the scissors
 ‘My mother forgot to take the scissors up with her and she asked my grand-
 mother: “Please, bring me the scissors up here.”’
 (Fuentes Claras, Teruel, COSER 4117)

Moreover, it has to be said that all transitive movement verbs mentioned thus far allow for the codification of non-argumental recipient datives. That is, their structure has room for referential datives which can, in theory, be coreferential with the subject. And hence, their structure has room for reflexive datives. Such reflexive structures, however, are rare, and the use of the stressed reflexive prepositional phrase *a sí mismo* with these verbs, although grammatical, is intuitively uncommon – these verbs seem to prefer other prepositions to encode their non-argumental reflexive recipients (see (212)). At any rate, it would be interesting to analyse whether the presence of prepositional phrases with *para* or *con* such as the ones in (212c) or even the presence of a possessive coreferential with the subject in the direct object (as in (212d)) favour the presence of the RM, which would hint at an influence of true indirect reflexive constructions in the distribution of the RM with these verbs. Unfortunately my data are too scant to check for this – in a sample of 100 marked examples of *llevarse* together with all the marked examples of *traerse* ($N = 95$), only 11 examples with possessives (and none with prepositional phrases with *para* or *con*) are found.

- (212) a. Te lo llevé (a ti).
 DAT.2SG ACC.3SG take.PST.1SG to OBL.2SG
 ‘I took it to you.’
- b. Me lo llevé (#a mí mismo).
 REFL.1SG ACC.3SG take.PST.1SG to REFL.1SG same
 ‘I took it (#to myself).’
- c. (Me) lo llevé {para mí /conmigo}
 REFL.1SG ACC.3SG take.PST.1SG for REFL.1SG with.me
 ‘I took it {for / with} me.’

- d. (Me) llevé mis cosas.
 REFL.1SG take.PST.1SG my things
 'I took my things.'

Finally, it is worth mentioning another transitive movement verb whose behaviour might have been influenced by its intransitive counterpart, i.e. *pasar(se)* 'to pass, to cross'. Only three examples (from 30, 10%) of transitive *pasarse* with a movement meaning are documented in the rest of the territory.³³ Interestingly (but also anecdotally, since there are so few examples), two of these cases (see (213)) refer to crossing a limit, one of the most typical contexts of the intransitive verb of motion *pasarse* (see Section 5.4).

- (213) Y no podías pasarte ese límite, porque cuántos,
 and not can.PST.2SG pass.INF-REFL.2SG that limit because how.many
 cuántas se han pasado el límite ese.
 how.many REFL.3 have.PRS.3PL cross.PTCP the limit that
 'And you were not supposed to cross that limit, because how many, how many
 have crossed that limit.'

(Santiuste de San Juan Bautista, Segovia, COSER 3712)

In conclusion, the data in this section show that reflexive transitive verbs of motion show a very similar behaviour to reflexive intransitive verbs of motion, suggesting that their origins are related. In the following subsection I briefly discuss verbs of translational motion in which the direct object depicts the path of the movement, whose behaviour is more reminiscent of reflexive transitive verbs related to auto-benefactive constructions.

6.10.1 Transitive reflexive verbs of translational motion with a direct object that conveys a path

Some verbs of translational motion that are documented sporadically in my data (such as *recorrer(se)* 'to cover a distance', *subir(se)* 'to climb/go up' and *bajar(se)* 'to climb/go down') are reminiscent of the uses of the RM which I argue developed from autobenefactive constructions (that is, reflexive consumption verbs, verbs of spending time and some cognition verbs, such as *saberse* and *aprenderse*). The direct object of these verbs conveys the path of the movement and, intuitively, the RM seems to be associated with impressive actions. Only four marked examples are found in my data, none of them in north-western varieties. In the rest of the territory, *recorrer(se)* shows the RM in three out of eight (38%) examples and *subir(se)* is marked only once in seven (14%) cases. The few examples of *recorrerse* clearly

33. There are only two examples of transitive *pasar* with a movement meaning in north-western varieties, neither of which is marked.

show the association with noteworthy achievements (see the clear hyperbaton in (214a)). The only example of *subirse* in this context, however, shows no trace of this emphatic nuance (see (214b)).

- (214) a. To los años por esa fecha en marzo o así suele estar
all the years for that date in March or so used.to.PRS.3SG be.INF
en Alemania, pero se ha recorrido yo creo que el
in Germany but REFL.3 have.PRS.3SG cover.PTCP I think that the
mundo casi.
world almost
'Every year, around March or so, she is usually in Germany, but she has
been in almost the whole world, I think.' (Ermua, Vizcaya, COSER 4503)
- b. [las mujeres, para mirar por debajo de la falda] Aquí antes estábamos
here before be.PST.1PL
esperando que se subieran las escaleras.
waiting that REFL.3 go.up.PST.SBJV.3PL the stairs
'Before we used to be waiting for them [the women] to climb the stairs [in
order to look under their skirts].' (Higuera, Albacete, COSER 0211)

In this respect, it is also worth noting that there are cases of "light verb + noun" structures where the noun is morphologically related to a movement verb and can be broadly understood as a path. These structures are *dar(se) un paseo/una vuelta* and *pegar(se) un paseo/una caminata*, all broadly meaning 'to go for a walk'. All the examples in the COSER data are documented in the rest of the territory. While *dar(se)* 'to give' is a much more neutral option to convey this meaning, *pegar(se)* 'to hit' is both more colloquial and more emphatic, typically referring to a noteworthy walk (see (215)). The two examples found in my data with *pegar(se)* showed the RM, which is also the only option I find acceptable introspectively. When asked, however, Galician speakers confirm that unmarked examples of *pegar(se)* in these contexts do not sound foreign to them.

- (215) [Las mujeres se cuidan mucho.] Se pegan cada paseo por ahí de
REFL.3 hit.3PL each walk for there of
la órdiga.
the INTERJECTION
'[Women take care of themselves.] They go for very long walks.'
(El Burgo, Álava, COSER 0103)

The RM is definitely optional with *dar(se)* – my data attest 13 marked examples out of 23 (57%). Moreover, not all the marked cases emphasise how impressive the walk is (see (216)), as happens with verbs where the presence of the RM is more routinised, such as consumption or cognition verbs. At any rate, more data on these kind of verbs and structures are needed to reach more definitive conclusions.

- (216) [las gallinas] Mάma, no las quites que con ese...
 mum not ACC.3PL remove.SBJV.2SG that with that
 quehacerico vas y vienes a... y te das
 housework go.PRS.2SG and come.PRS.2SG to and REFL.2SG give.PRS.2SG
 un paseo.
 a walk
 ‘Mum, don’t take them [the hens] away, with the work they give, you come and
 go... and you go for a walk.’ (Beratón, Soria, COSER 3924)

6.11 Allobenefactives

Finally, I would like to briefly mention some examples documented in Spanish (although none in the COSER or the questionnaire data), where, very surprisingly, the RM is used to mark non-self-benefactives. As far as I know, these have not previously been mentioned in the literature. For lack of a better term, I have named this construction “allobenefactive”. They are found with verbs meaning ‘to pay’ or ‘to invite’ (*convidar* ‘to invite’, *invitar* ‘to invite’, *pagar* ‘to pay’, *mandar* ‘to send’), as shown in (217), as well as with the speech verb *contar* ‘to tell’ (see (218)). Although all the examples presented below come from Twitter, it is crucial to note that this use is not restricted to online social networks and is not uncommon at all in at least some registers of colloquial Spanish.

- (217) a. [Respuesta a: Dentro de una semana (16 de mayo) cumpliremos 3 años sin saber aún muy bien cómo hemos conseguido resistir. Pero qué alegría y qué orgullo.]
CONVIDARSE A ALGO CHAVALES
 invite.IMP.3PL-REFL.3 to something guys
 ‘[Answer to: In one week (May 16th) we’ll celebrate our third anniversary nor really knowing how we’ve managed to survive. But how happy and proud we are.] Buy us something, guys!’ (May 12th 2014. Tweet.)
- b. Felicidades @usuario que pases un dia.... aqui con los
 happiness @user that spend.PRS.2SG a day here with the
 unicornios.... jajaja invitate a algo.... 🐾
 unicorns hahaha invite.IMP.2SG-REFL.2SG to something
 [adjunta un archivo GIF de unicornios que bailan y desean feliz cumpleaños]
 ‘Happy birthday, @user, I hope your day is... with the unicorns, hahaha, buy us something... [A GIF file of dancing unicorns wishing happy birthday is attached.]’ (January 4th 2017, Madrid. Tweet.)

- c. Feliz cumpleaños @usuario! págate unos “algos”,
happy birthday @user pay.IMP.2SG-REFL.2SG some somethings
no? :P
no
‘Happy birthday, @user, buy us a few “somethings”, won’t you?’
(January 17th 2013. Tweet.)
- d. [Respuesta a: Ya tengo acabado mi póster folclero cantabresco, pero no
os lo puedo enseñar porque está embargado]
Mandaté unos sobaos “El Macho” y
send.IMP.2SG-REFL.2SG some sponge.cakes El Macho and
dejaté de posters y de historias.
leave.IMP.2SG-REFL.2SG of posters and stories
[Answer to: I have already finished my Cantabrian folk poster, but I can’t
show it to you, because it is sequestered.] ‘Send us some Cantabrian sponge
cakes “El Macho” and stop talking about posters and stuff like that.’
(May 7th 2013. Tweet.)
- (218) Pues parece que, drogas o no drogas, aquí va a dormir
well seem.PRS.3SG that drugs or not drugs here go.PRS.3SG to sleep.INF
Rita, ¿Qué os contáis?
Rita what REFL.2PL tell.PRS.2PL
‘So it seems that, with or without drugs, I won’t be able to sleep, what do you
have to tell?’
(October 5th 2014. Tweet.)

In these cases, which are used in colloquial speech in Peninsular Spanish,³⁴ the use of the RM might be a strategy to express politeness. This is apparent when the RM appears with a verb in the imperative, where they can be used to mitigate the fact that the speaker is asking for something. The RM, because of its connection with autobenefactives, allows for presenting the situation as beneficial to the subject too, presenting the request as less intrusive or threatening. In other tenses, they seem to highlight the generosity of the subject in sharing their own benefit – this is also a natural implicature with imperatives. The examples in (219) (also found in the COSER data and Twitter) could be seen as a bridging context between autobenefactives and allobenefactives, since in them the beneficiary is not only the subject, but also one or more third persons (that is, they are examples of a shared-benefit interpretation, in Smith’s (2010) terms).

34. Unfortunately, I do not know whether they are also used in American varieties of Spanish. Since the COSER data do not offer the appropriate contexts of appearance, I also do not know whether they are restricted to some Peninsular Spanish varieties.

- (219) a. Mátate un conejo, Jesús, que nos le
 kill.IMP.2SG-REFL.2SG a rabbit Jesus that REFL.1PL ACC.3SG
 vamos a comer.
 go.PRS.1PL to eat.INF
 ‘Kill a rabbit, Jesus, we’ll eat it.’
 (Manzanares el Real, Madrid, COSER 2910)
- b. yo tenía una apuesta con mi ex, el primero que se
 I have.PST.1SG a bet with my ex the first that REFL.3
 tirase un pedo se pagaba una cena.
 throw.PST.SBJV.3SG a fart REFL.3 pay.PST.3SG a dinner
 ‘I had a bet with my ex-boyfriend, the first one to fart in front of the other
 would have to buy dinner.’ (February 1st 2014. Tweet.)
- c. -Buenos días princesa, ¿te apetece un colacao? –
 good morning princess DAT.2SG please.3SG a chocolate.drink
 Siiii – Pues prepárate dos
 yes well prepare.IMP.2SG-REFL.2SG two
 ‘– Good morning, princess, do you want to drink a hot chocolate? – Yes!!!
 – Well, then prepare two.’ (December 12th 2013. Tweet.)

A quantitative analysis of these examples is needed to achieve a better understanding of them and their relationship with other reflexive transitive constructions – it must be noted that this section is largely based on my introspective judgements.

6.12 Conclusion

In this chapter I have discussed reflexive transitive verbs. As far as I am aware, this is the most comprehensive analysis of reflexive transitive constructions in the varieties under study, thanks to the exhaustive analysis of corpora. I have shown the limitations of aspectual analyses based on the delimitation of the direct object or the culmination of the event. Such analyses cannot account for all the verbs documented in the reflexive transitive structure and, moreover, some important generalisations are left out, such as the association of reflexive cognition transitive verbs with direct objects of a very specific kind, which might be restricted to small semantic fields (as in the case of some cognition verbs like *saber* ‘to know’, *aprender* ‘to learn’ and *estudiar* ‘to study’). In my opinion, these facts can only be explained from a diachronic perspective, taking into account the many different functions of the RM that affect its presence with transitive verbs, as well as the gradual expansion of this particle across similar contexts via analogy. Accordingly, I have focussed on showing the links between different classes of reflexive transitive verbs and other

reflexive constructions (mostly reflexive indirect constructions, but also reflexive intransitive verbs and the middle uses of the RM), in order to propose a hypothesis that relies heavily on the analogical origin of all reflexive verbs with no change of valency. I have claimed that the different sources of analogy explain the different contextual associations of the RM depending on the verb class it attaches to, and have argued against a unitarian account of the function of the RM in reflexive transitive verbs.

Moreover, the data on reflexive transitive constructions resemble the data on reflexive intransitive constructions in that the RM shows variable frequency depending on the lexeme even with verbs of similar meaning. This suggests a high degree of lexicalisation in the combination of the RM with verbs where no change of diathesis is found.

Conclusions

7.1 A variationist study of the middle voice

In this book I have presented a comprehensive study of the uses of the reflexive marker that can be considered under the “middle voice” label in varieties of Spanish and Galician, this from a variationist perspective. Such a study is novel in several ways: (1) Although several comprehensive accounts of the uses of the RM in these varieties have been given (see Section 1.2), not many have been based on the exhaustive analysis of corpus data. Probably because of this, the classification criteria used for grouping different functions of the RM have often been unclear or unsystematic. A notable exception is Cartagena (1972), to whom I am greatly indebted for the classification used in the current study. (2) Variationist accounts of reflexive constructions have traditionally been very rare. One early exception is provided in Garita & Rojas (1993), who address the distribution of the RM in a number of anticausative verbs. More recently, a number of variationist studies on the distribution of the RM with a few specific intransitive and intransitive verbs have been published, especially but not exclusively with a focus on Mexican Spanish (Aaron 2003, 2004; Aaron & Torres Caocullos 2005; Torres Caocullos & Schwenter 2008; Rivas 2011; De la Mora 2011; Merino Hernández 2019). This book is the most ambitious variationist account of these uses to date, considering the large number of constructions included herein. (3) Although some recent research focussing on the use of the RM with a few verbs (Torres Caocullos & Schwenter 2008; De la Mora 2011) includes large diatopic differences (such as comparing Peninsular Spanish vs. American Spanish), no study has used fine-grained dialectal data from a continuous territory. The advantage of such an approach, as adopted in this book, is that it can be used for establishing fine-grained hypotheses on the relatedness and the relative chronology of different uses.

The design of the study, of course, has not considered novelty as an end in itself, but rather as a means – my goal here has been to provide a synchronic account of the functions of the RM in the middle voice that is consistent with our knowledge of the historical development of the RM and linguistic change in general. That is, I have adopted the opposite approach to most synchronic accounts, which Rigau (1994) summarises perfectly: “*Whatever the historical origin of the presence of the*

pronoun in the lexical representation of these verbs might be, grammar must explain what the syntactic and semantic consequences of such an incorporation are” (Rigau 1994: 29, my translation, emphasis added). In my opinion, such consequences can only be fully understood by taking into account the historical origin of the presence of the pronoun. Proposals with a synchronic focus have restricted their analyses to partial samples of the data, leaving aside verbs and contexts that do not easily fit within those proposals. Historical explanations, however, should be powerful enough to explain the data in full, which is why exhaustive corpus analysis is crucial (see also De Benito Moreno 2015, 2021).

Therefore, my synchronic variationist approach has served to establish hypotheses on how the RM expanded across functions and contexts. That is, by comparing the relative frequency of the RM in different functions, context and areas, and by assuming that the RM must be older wherever it is more frequent, I have proposed diachronic hypotheses accordingly (a summary of this approach is given in Section 7.3). Moreover, I have compared the behaviour of the RM depending on similar parameters across functions in order to support my proposal that analogy is the most important cognitive mechanism underlying the expansion of the RM. Because analogy can be conceived of as a very powerful mechanism, it can also be seen as almost trivial – this is why the quantitative comparison of contexts is crucial and must be further supported with historical studies in the future.

In the remainder of this chapter I summarise the findings of this study. Section 7.2 deals with the three research questions that have to do with the synchronic behaviour of the RM: What is the nature of the middle voice in these varieties? What is an appropriate typology of middle constructions in these varieties? What factors explain the variable presence of the RM? In Section 7.3 I summarise my historical proposal for the expansion of the RM.

7.2 The nature of the middle voice in the varieties under study

Chapter 3 was devoted to a discussion of the nature of the middle voice in the Ibero-Romance varieties under study, since there are several interpretations of this category in the literature, according to which it can be understood as a derived voice (which marks diatheses changes) or as a basic voice (associated with a given semantic content). Accordingly, I have performed two quantitative analyses on my data, one following Kulikov’s (2011) description of the middle voice as a derived voice, and another one following Kemmer’s (1993) account of the middle voice as a basic voice. My data on Ibero-Romance varieties clearly support a syntactic interpretation of the function of the RM, since its productivity in marking changes of diatheses where an argument is demoted or deleted is far higher than

its productivity in marking middle event types. That is, the middle voice in these varieties is a derived voice. This suggests that the most appropriate typology of reflexive constructions in the varieties here must take into account what change of diathesis the RM is associated with. Six relevant changes of diathesis have been identified in this book: anticausative, absolute, de-objective, conversive, antipassive (Section 3.2), plus indirect detransitivization (Section 6.2.4). Moreover, the RM can also appear with both intransitive and transitive verbs that undergo no change of valency – different functions of the RM can be identified here depending on the verb's semantic profile (Chapters 5 and 6).

This function of the RM as an intransivising device is clearly linked to the direct reflexive construction (see Section 1.3), where coreference between the subject and the direct object means that there is a single referent for the two arguments of the transitive verb. In the reflexive direct construction, the RM is a referential pronoun and the subject performs the verbal action intentionally, which requires that it is an animated entity. As an intransivising device, the RM is no longer a pronoun and requires no intentionality from the subject, even allowing for inanimate subjects. That is, this change entails the reanalysis of the reflexive pronoun into a previously non-existent category, since the only other elements that Ibero-Romance languages use to mark changes of diathesis are auxiliary verbs – a radical categorial innovation in the sense of De Smet's (2009) study.

The extension of the reanalysis was gradual and cannot be said to have been completed – the use of the RM in intransivising diatheses is not completely systematic, as was also noted in Chapter 3. In the diatheses where the object is demoted or deleted, the use of the RM is highly dependent on the lexical item, while the use of the RM reaches its peak in the conversive diathesis, suggesting that these verbs were the first ones to acquire the RM. The strong semantic association of this diathesis with emotion events, where the subject is most often conceived of as an experiencer instead of as an agent, might be an indication that the bridging contexts between the reflexive pronoun and the intransivising device were those that allow for both an agentive and non-agentive reading of the subject (such as *golpearse* 'to hit oneself', see Section 6.2.4 for their indirect reflexive counterparts).

Reflexive marking of anticausative verbs is not as systematic as with conversive verbs, which allows for a more detailed study of the factors that favour and disfavour the use of the RM. Such an analysis is provided in Chapter 4. On the one hand, I tested two hypothesis that have been proposed in the literature, namely, that the distribution of the RM is affected either by the distinction between externally and internally caused events or by the aspectual properties of the predicate. Regarding the former, the quantitative analysis suggests that both the proportion of transitive uses (as tested by Vivanco 2016) and the presence of a dative participant interpreted as an external cause or as an affected participant favour the presence of the RM.

Regarding the latter, I have found no quantitative confirmation of the role of telicity and atelicity in the presence of the RM.

Besides testing these hypotheses, I have tested the effect of the animacy of the subject, since the fact that the source construction (the reflexive direct construction) is restricted to animate subjects indicates that this should be an important factor that affects the extension of the reanalysis. My data suggest that this is indeed the case, since animate subjects favour the presence of the RM in anticausative verbs. Surprisingly, animacy has often been disregarded in synchronic analyses of anticausative verbs, despite having been highlighted in historical studies of the evolution of the RM (Monge 1955).

Finally, my data also show important regional differences in the use of the RM as an intransivising device, north-western varieties being very resistant to the presence of the RM. Although most of my data from these regions come from north-western varieties of Spanish (but some also from Galician), the fact that these are bilingual regions suggests that the varieties of Spanish spoken in the northwest are affected by Galician and Asturian, although my data are not suited to exploring the effects of language contact. In all these varieties (Galician, Asturian and the Spanish spoken in those regions) the RM shows a less advanced stage of the extension of the reanalysis – although the RM is used as a intransivising device, it is largely restricted to constructions with animate subjects, where it is less systematic than in other varieties of Spanish.

7.3 The RM in verbs with no change of diathesis

Chapters 5 and 6 are devoted to the analysis of the RM that attaches to verbs with no change of diathesis, that is, constructions where the RM cannot be interpreted as a marker of the middle derived voice. The RM in such constructions poses interesting theoretical challenges, since its categorial status is harder to understand. The RM that functions as an intransivising device, although a categorial innovation, can be seen as an element that absorbs one of the arguments of the verb, which is similar to what a pronoun (i.e. its original function) does. The RM that appears with intransitive and transitive verbs with no change of valency, however, has no clear syntactic function. Much of the recent literature has been devoted to answering the question as to what this element does, either proposing that the RM is a marker of a specific aspectual structure of the verb (Rigau 1994; Nishida 1994; Zagona 1996; De Miguel & Fernández Lagunilla 2000; Sanz & Laka 2002; Sánchez López 2002; McCready & Nishida 2008; Basilico 2010; Campanini & Schäfer 2011), a marker of semantic-pragmatic notions such as counter-expectation (Maldonado 1988; Aaron 2003, 2004; Aaron & Torres Cacoullos 2005; Torres Cacoullos & Schwenter 2008)

and higher involvement of the agent (Maldonado 2000; Armstrong 2013), or, more recently, a marker of high transitivity (Rivas 2011; De la Mora 2011).

The comprehensive analysis of the marked verbs found in a corpus suggests, however, that these proposals are oversimplistic. This is so because the RM that appears with these verbs cannot be reduced to a single synchronic function (see also Maldonado 2000; Armstrong 2013). On the contrary, I believe that the presence of the RM with these verbs must be explained taking into account the previous functions of the RM (the direct and indirect reflexive, the reciprocal and the detransitivising uses) and the contextual associations it shows in such functions (with animate subjects, some semantic classes of verbs, the existence of an external cause in the case in anticausatives, a higher involvement of the agent due to the impressiveness of the event in autobenefactives, the lack of control of the agent in conversive verbs and indirect detransitivised constructions, etc.). As can be seen, the already large number of reflexive constructions provides a large number of contextual associations that allow the RM to be associated even with antonymous contexts (higher involvement and lack of control of the agent, for instance). Moreover, innovative uses of the RM might be influenced by several previous functions (instead of just a single one). In this sense, it is unclear whether these constructions belong in the middle voice, not only because they do not fit within the cluster of valency-reducing diatheses that make up the core of this voice, but because they are also connected to the reflexive and reciprocal diatheses, which are not traditionally considered middle.

All in all, this makes the distribution of the RM extremely complex and the question of its categorial status even more so, since it is possible to identify a large number of different constructions where the RM appears with different contextual associations. My goal, however, has not been to answer that question, but to try to understand how the synchronic distribution of the RM came about, under the hypothesis that it extended gradually from old contexts to new ones, thanks to both iconic (i.e. paradigmatic) and indexical (i.e. syntagmatic) analogy (Fischer 2008). That is, because the RM has not only expanded to new verbs, but it has done so gradually, affecting first those contexts that showed similar characteristics to ones where it had already appeared, it has acquired new meanings via contextual association (see Torres Cacoullous & Schwenter 2008). Interestingly, some of these new meanings appear to be quite lexicalised, while others are very productive. Two different mechanisms seem to be at work here. On the one hand, we find “gradual category-internal change”, according to which the boundaries of a new category are gradually extended through minor semantic changes, by bringing about new instances (De Smet 2009). That is what happens when the RM is combined with a new verb, a new type of subject, etc. On the other hand, we find automation, that is, the creation of less schematic constructions from more abstract constructions through repetition (De Smet 2009). This is what happens when some reflexive transitive

verbs develop a restriction for delimited objects, or when some reflexive intransitive verbs become associated with specific locative prepositions, for instance. These associations, in turn, become new potential sources of analogical extension and might affect other verbs. Accordingly, we find (1) more lexicalised uses of the RM, where it might even become compulsory in certain contexts through automation (*se fue de aquí* ‘s/he leaves this place’), and which I claim to be the original uses of the RM, (2) less lexicalised uses of the RM, where its presence is simply possible and might even be felt as expletive ((*se*) *va al mercado* ‘s/he goes to the market’, (*se*) *bebe un café por la mañana* ‘s/he drinks a coffee in the morning’) – which I deem to be examples of the extension of the RM across contexts, and (3) contexts that the RM has not (yet?) reached, where its presence is felt to be inappropriate ((?*se*) *bebió cerveza* ‘s/he drunk beer’).

In such a picture, the proposals of previous authors to identify a regular and systematic function of the RM are attempts to identify the commonalities that the more lexicalised uses and the inappropriate contexts show, leaving the less lexicalised uses aside. But disregarding those contexts is problematic, for at least two reasons:

1. There is no reason to believe that the extension of the RM stops once it develops some compulsory restrictions: a number of verbs admit and even prefer the RM in every context, suggesting that they might become compulsorily reflexive (i.e. *reír(se)* ‘to laugh’, see de Benito Moreno accepted). When Nishida notes that for some speakers “the pairs like *comer/comerse*, *fumar/fumarse*, *tomar/tomarse*, etc., have come to be in complementary distribution with respect to the partitive/non-partitive contrast”, while for other speakers they “are not entirely in complementary distribution yet at the current stage of Spanish” (Nishida 1994: 443, emphasis added), she seems to assume the complementary distribution as an endpoint. However, corpus data suggest otherwise, since variable reflexive marking in other contexts is constantly found. This is consistent with the status of the RM as the single member of a radically new category – because it is not modelled on previous elements of a given category, there are no clear limits to the extension of the reanalysis (see De Smet 2009 for a discussion on the creation of new grammatical categories).
2. Because these restrictions typically appear first in one verb and subsequently develop afterwards (hence more slowly) with other similar verbs, the less lexicalised uses of one reflexive verb might correspond to the more lexicalised uses of another one, etc. This can be seen both in my study and in other variationist accounts: Rivas (2011: 399) notes that, in his data, “*beberse* [‘to drink-REFL.’] is mainly restricted to examples in which the total affectedness of the direct object is indicated explicitly”, but this was not so for other consumption verbs such as *comer(se)*, *tomar(se)* or *tragar(se)*.

A different approach to the same problem is adopted by Armstrong (2013), who identifies two formally identical constructions: the “agentive reflexive clitic construction”, which has three main properties: “(i) it requires an agent, (ii) it adds a conventional implicature of *willful intent* to the agent and (iii) it imposes an aspectual restriction on the VP where it appears (= only accomplishments)” (Armstrong 2013: 82) and the “transitive SE clitic” construction, which “(i) do[es] not systematically assign one particular role to their subject, (ii) ha[s] no set aspectual value associated with them” (Armstrong 2013: 83). In my view, this distinction clearly illustrates the distinction between more lexicalised and less lexicalised uses of the RM – the agentive reflexive clitic construction refers to the original uses of the RM with some transitive verbs, linked to its autobenefactive function in the indirect reflexive construction, while the transitive SE clitic construction refers to the new uses of the RM, which is expanding to other contexts of those verbs already affected. Differentiating the two constructions obscures the connection between them.

In conclusion, the nature of the RM with verbs with no change of diathesis makes it especially difficult to propose a unified function and also to propose several differentiated functions of the RM. Thus, I believe an account based on historically plausible analogical connections to be more appropriate. Accordingly, Chapters 5 and 6 have been devoted to the comparison of the distribution of the RM in these verbs with its distribution in other functions across the relevant hypothesised contexts (semantic classes, animacy of the subject, agentivity of the subject, lexicalised meanings, etc.).

On the one hand, thanks to the exhaustive analysis of corpus data, I have been able to discuss reflexive verbs that have not been considered in previous accounts, such as symmetric intransitive verbs (Section 5.2), reflexive intransitive and transitive verbs encoding physical corporal processes (Sections 5.6 and 6.8), spontaneous uses of *pasar(se)* ‘to pass’ (Section 5.8), reflexive transitive verbs of motion (Section 6.10), and allobenefactives (Section 6.11). On the other hand, my data have shown that no single previous function of the RM can account for the presence of the RM with these verbs, since the direct reflexive, the reciprocal construction, the detransivising diatheses, as well as all indirect reflexive constructions, might act as the main source of analogy for some reflexive intransitive and transitive verbs. Even reflexive intransitive verbs might be the source of analogy for reflexive transitive verbs (cf. reflexive transitive verbs of motion and reflexive transitive verbs encoding physical corporal processes).

Finally, the spatial distribution of reflexive verbs with no change of diathesis also supports the role of analogy. On the one hand, north-western varieties show a very low productivity of these verbs. This is to be expected, since both detransivising functions of the RM and at least some indirect reflexive constructions, such as

the possessive reflexive dative, are also less productive in these varieties, meaning that the sources of the analogy are reduced. On the other hand, the geographical distribution of these verbs is not consistent, i.e. while some of these verbs are restricted to specific areas, no single area that favours the presence of the RM in these verbs can be found. This is consistent with the highly lexically determined productivity of these verbs and with the lack of a single function – the potential for new analogies is found in all the territory and, since they do not arise simultaneously but affect each verb individually, a single focus of the innovations is not to be expected.

References

- Aaron, Jessi Elana. 2003. Me salí a caminar: Pronominal constructions with intransitive motion verbs in northern New Mexican Spanish. In Lotfi Sayahi (ed.), *Selected Proceedings of the First Workshop on Spanish Sociolinguistics*, 123–133. Somerville, MA: Cascadilla Proceedings Project.
- Aaron, Jessi Elana. 2004. The gendered use of *salirse* in Mexican Spanish: Si me salía yo con las amigas, se enojaba. *Language in Society* 33(4). 585–607.
<https://doi.org/10.1017/S0047404504334044>
- Aaron, Jessi Elana & Rena Torres Cacoullos. 2005. Quantitative measures of subjectification: A variationist study of Spanish *salir(se)*. *Cognitive Linguistics* 16(4). 607–633.
<https://doi.org/10.1515/cogl.2005.16.4.607>
- Academia de la Llingua Asturiana. 2001. *Gramática de la Llingua Asturiana*. Oviedo: Academia de la Llingua Asturiana.
- Alarcos Llorach, Emilio. 1970. Valores de se en español. In Emilio Alarcos Llorach, *Estudios de gramática funcional del español*, 156–165. Madrid: Gredos.
- Alarcos Llorach, Emilio. 1994. *Gramática de la lengua española*. Madrid: Espasa.
- Alcina Franch, Juan & José Manuel Bleca. 1975. *Gramática española*. Barcelona: Ariel.
- Almela Pérez, Ramón, Pascual Cantos & Aquilino Sánchez. 2005. *Frecuencias del español. Diccionario y estudios léxicos y morfológicos*. Madrid, Universitas.
- Álvarez, Rosario, Xosé Luís Regueira & Xosé Henrique Monteagudo. 1986. *Gramática galega*. Vigo: Galaxia.
- Álvarez, Rosario & Xosé Xove. 2002. *Gramática da lingua galega*. Vigo: Galaxia.
- Anvari, Amir, Mora Maldonado, and Andrés Soria Ruiz. 2019. The Puzzle of Reflexive Belief Construction in Spanish. In M.a Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally, and Cristina Real-Puigdollers (ed.), *Proceedings of Sinn Und Bedeutung* 23, 57–74. Barcelona: Universitat Autònoma de Barcelona.
- Arce–Arenales, Manuel, Melissa Axelrod & Barbara A. Fox. 1994. Active voice and middle diathesis: A cross-linguistic perspective. In Barbara A. Fox & Paul J. Hopper (eds.), *Voice. Form and Function*, 1–21. Amsterdam/Philadelphia: John Benjamins.
<https://doi.org/10.1075/tsl.27.03arc>
- Armstrong, Grant. 2013. Agentive reflexive clitics and transitive *se* constructions in Spanish. *Borealis* 2(2). 81–128. <https://doi.org/10.7557/1.2.2.2526>
- Arús, Jorge. 2006. Perspectiva sistémico-funcional de los usos de ‘se’ en español. *Revista Signos* 39(61). 131–159. <https://doi.org/10.4067/S0718-09342006000200001>
- Ávila López-Pedraza, Julio. 2009. Análisis de los usos de se desde la gramática descriptiva. *Hispania* 92(4). 791–799.
- Babcock, Sandra S. 1970. *The syntax of Spanish reflexive verbs*. The Hague/Paris: Mouton.
<https://doi.org/10.1515/9783110874761>
- Baldi, Philip. 1979. A structural ambiguity in Latin. *Classical Philology* 74(1). 49–52.
<https://doi.org/10.1086/366468>

- Basilico, David. 2010. The *se* clitic and its relationship to paths. *Probus* 22(2). 271–302.
<https://doi.org/10.1515/prbs.2010.010>
- Bassols de Climent, Mariano. 1956. *Sintaxis latina*. Madrid: CSIC.
- Bello, Andrés. 1981 [1847]. *Gramática de la lengua castellana. Edición crítica de Ramón Trujillo*. Tenerife: Instituto universitario de lingüística Andrés Bello.
- Benavides, Carlos. 2010. El clítico *se* en español y la estructura léxico-conceptual. *RILCE* 26(2). 261–288.
- Bickel, Balthasar. 2007. Typology in the 21st century: Major current developments. *Linguistic Typology* 11. 239–251. <https://doi.org/10.1515/LINGTY.2007.018>
- Bobes Naves, María del Carmen. 1974a. Construcciones castellanas con *se*. *Análisis transformacional*. *RSEL* 4(1). 87–127.
- Bobes Naves, María del Carmen. 1974b. Construcciones castellanas con *se*. *Análisis transformacional (continuación)*. *RSEL* 4(2). 301–326.
- Bogard, Sergio. 1999a. Construcciones antipasivas en español. *Nueva Revista de Filología Hispánica* 47(2). 305–327. <http://www.jstor.org/stable/40300119>. (9 February, 2022.)
- Bogard, Sergio. 1999b. Duplicación y clausura argumental: Dos funciones del clítico reflexivo en español. *Español actual* 71. 41–48.
- Bogard, Sergio. 2006. El clítico *se*. Valores y evolución. In Concepción Company Company (dir.), *Sintaxis histórica de la lengua española. Primera parte: La frase verbal*, vol. 2, 755–870. México: UNAM/FCE.
- Bohnenmeyer, Jürgen, Melissa Bowerman & Penelope Brown. 2001. Cut and break clips. In Stephen C. Levinson & Nick J. Enfield (eds.), *Manual for the field season*, 90–96. Nijmegen: Max Planck Institute for Psycholinguistics. <http://fieldmanuals.mpi.nl/volumes/2001/cut-and-break-clips/>. (9 February, 2022.)
- Boyd, Raymond. 2010. A ‘reflexive benefactive’ in Chamba-Daka (Adamawa branch, Niger-Congo family). In Fernando Zúñiga & Seppo Kittilä (eds.), *Benefactives and malefactives. Typological perspectives and case studies*, 331–349. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.92.14boy>
- Britain, David. 2004. Geolinguistics and linguistic diffusion. In Ulrich Ammon, Norbert Dittmar, Klaus Mattheier & Peter Trudgill (eds.), *Sociolinguistics: International Handbook of the Science of Language and Society. (Handbücher zur Sprach- und Kommunikationswissenschaft 3.1.)*, 34–48. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110141894.1.134>
- Bull, William E. 1950. *Quedar* and *quedarse*: A study of contrastive ranges. *Language* 26(4). 467–480. <https://doi.org/10.2307/410398>
- Bull, William E. 1952. The intransitive reflexive: *Ir* vs. *irse*. *The Modern Language Journal* 36(8). 382–386.
- Bybee, Joan. 1985. *Morphology. A study of the relation between meaning and form*. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.9>
- Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511612886>
- Bybee, Joan. 2007. Introduction. In Joan Bybee, *Frequency of use and the organization of language*, 5–22. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195301571.003.0001>
- Bybee, Joan. 2015. *Language change*. Cambridge: Cambridge University Press.
- Bybee, Joan, & David Eddington. 2006. A usage-based approach to Spanish verbs of ‘becoming’. *Language* 82(2). 323–355. <https://doi.org/10.1353/lan.2006.0081>

- Bybee, Joan & Paul Hopper. 2001. Introduction to frequency and the emergence of linguistic structure. In Joan Bybee & Paul Hopper (eds.), *Frequency and the emergence of linguistic structure*, 1–24. Amsterdam/Philadelphia: John Benjamins.
<https://doi.org/10.1075/tsl.45.01byb>
- Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago and London, University of Chicago Press.
- Bybee, Joan & Sandra Thompson. 2007. Three frequency effects in syntax. In Joan Bybee, *Frequency of use and the organization of language*, 269–278. New York: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780195301571.001.0001>
- Campanini, Cinzia, & Florian Schäfer, Florian. 2011. Optional *se*-constructions in Romance: Syntactic encoding of conceptual information. Talk at *Generative Linguistics in the Old World* 34. University of Wien.
- Cano, Rafael. 1981. *Estructuras sintácticas transitivas en el español actual*. Madrid: Gredos.
- Cartagena, Nelson. 1972. *Sentido y estructura de las construcciones pronominales en español*. Concepción: Universidad de Concepción.
- Cennamo, Michela, & Elisabetta Jezek. 2011. The anticausative alternation in Italian. In Giovanna Massariello Merzagora & Serena Dal Maso (eds.), *I luoghi della traduzione. Le interfacce*, 809–823. Roma: Bulzoni.
- Cennamo, Michela. 1999. Late Latin pleonastic reflexives and the unaccusative hypothesis. *Transactions of the Philological Society* 97(1). 103–150. <https://doi.org/10.1111/1467-968X.00046>
- Chambers, John Kenneth & Peter Trudgill. 1980. *Dialectology*. Cambridge: Cambridge University Press.
- Cidrás Escáneo, Francisco Antonio. 1991. As construcións pronominais en galego. Proposta de clasificación e caracterización. In Mercedes Brea & Francisco Fernández Rei (coords.), *Homenaxe ó profesor Constantino García*, 101–125. Santiago de Compostela: Universidade de Santiago de Compostela.
- Cidrás Escáneo, Francisco Antonio. 2003. Alternancias transitivo–intransitivo e formación de predicados en galego. *Verba* 30. 81–115.
- Colquhoun, David. 2017. The reproducibility of research and the misinterpretation of p-values. *Royal Society Open Science* 4(12). 171085. <https://doi.org/10.1098/rsos.171085>
- Cooreman, Ann. 1994. A functional typology of antipassives. In Barbara Fox & Paul J. Hopper (eds.), *Voice. Form and Function*, 49–88. Amsterdam/Philadelphia: John Benjamins.
<https://doi.org/10.1075/tsl.27.05coo>
- Cornips, Leonie & Aafke Hulk. 1996. Ergative reflexives in Heerlen Dutch and French. *Studia Linguistica* 50(1). 1–21. <https://doi.org/10.1111/j.1467-9582.1996.tb00340.x>
- Corominas, Joan & José A. Pascual. 1981. *Diccionario crítico etimológico castellano e hispánico*. Madrid: Gredos.
- COSER= Fernández-Ordóñez, Inés. 2005–. COSER. *Corpus Oral y Sonoro del Español Rural*. www.uam.es/coser. (9 February, 2022.)
- Coseriu, Eugenio. 1958. *Sincronía, diacronía e historia*. Montevideo: Universidad de la República.
- Coseriu, Eugenio. 1973. *Sincronía, diacronía e historia*. Madrid: Gredos.
- Coseriu, Eugenio. 1977. “‘Tomo y me voy’. Un problema de sintaxis comparada europea”, in *Estudios de Lingüística Románica*. Madrid: Gredos, 79–151.
- Creissels, Denis. 2010. Benefactive applicative periphrases. In Fernando Zúñiga & Seppo Kittilä (eds.), *Benefactives and malefactives. Typological perspectives and case studies*, 29–69. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.92.02cre>

- Cuervo, Rufino José. 1954 [1907]. *Notas a la gramática de la lengua castellana de D. Andrés Bello*. In Rufino José Cuervo, *Obras*, vol. I, 159–179. Bogotá, Instituto Caro y Cuervo.
- Cuervo, Rufino José. 1994 [1886]. *Diccionario de construcción y régimen de la lengua castellana*. Bogotá: Instituto Caro y Cuervo.
- D'Andrés, Ramón. 1994. Los pronomes reflexivos átonos n'asturianu. *Lletres Asturianas* 51. 7–22.
- D'Introno, Francesco, Verónica González & Javier Rivas. 2007. Aspectos sintácticos y semánticos del pronombre se. *Boletín de lingüística* XIX(28). 5–25.
- De Benito Moreno, Carlota. 2015. Constructions with se: A variationist and dialectal perspective. Madrid: Universidad Autónoma de Madrid PhD thesis.
- De Benito Moreno, Carlota. 2021. Is there really an aspectual se in Spanish? *Folia Linguistica* 55(1). 195–230. <https://doi.org/10.1515/flin-2020-2074>
- De Benito Moreno, Carlota. accepted. From (semi-)oppositional to non-oppositional middles: The case of Spanish reir(se). *Special issue of STUF. Language Typology and Universals* (edited by Guglielmo Inglese).
- De Benito Moreno, Carlota & Ana Estrada Arráez. 2018. Aproximación metodológica al estudio de la variación lingüística en las interacciones digitales. *Revista Estudios del Discurso Digital* 1. 74–122. <https://doi.org/10.24197/redd.1.2018.74-122>
- De la Mora, Juliana. 2011. A quantitative approach to variable se-marking in Spanish ingestive verbs. Columbus, OH: The Ohio State University PhD thesis.
- De Miguel, Elena & Marina Fernández Lagunilla. 2000. El operador aspectual se. *Revista Española de Lingüística* 30(1). 13–43.
- De Smet, Hendrik. 2009. Analyzing reanalysis. *Lingua* 119. 1728–1755. <https://doi.org/10.1016/j.lingua.2009.03.001>
- De Smet, Hendrik. 2012. The course of actualization. *Language* 88. 601–633. <https://doi.org/10.1353/lan.2012.0056>
- DeMello, George. 1997. Morir vs. morirse. In Josse de Kock & George DeMello, *Lengua escrita y habla culta en América y España*, 99–125. Salamanca: Ediciones Universidad de Salamanca.
- Detges, Ulrich & Richard Waltereit. 2002. Grammaticalization vs. reanalysis: A semantic-pragmatic account of functional change in grammar. *Zeitschrift für Sprachwissenschaft* 21(2). 151–195. <https://doi.org/10.1515/zfsww.2002.21.2.151>
- Dimitriadis, Alexis. 2008. Irreducible symmetry in reciprocal constructions. In Ekkehard König & Volker Gast (eds.), *Reciprocals and reflexives. Theoretical and Typological Explorations*, 375–409. Berlin/New York: Mouton de Gruyter.
- Di Tullio, Ángela. 2012. La construcción de la lectura agentiva del se no argumental. In Valeriano Bellostá von Colber & Marco García García (eds.), *Aspectualidad – Transitividad – Referencialidad. Las lenguas románicas en contraste*, 69–85. Frankfurt: Peter Lang.
- Domínguez Oroña, María Beatriz. 2015. El uso del clítico se en construcciones causativas y anticausativas en gallego. Talk given at XXX Congreso Internacional de la Asociación de Jóvenes Lingüistas. 7–9 October, Centro Universitario CIESE, Comillas, Cantabria. <https://minerva.usc.es/xmlui/handle/10347/13667>. (9 February, 2022.)
- DRAE = RAE. 2015. *Diccionario de la Lengua Española*. <https://www.rae.es/>. (9 February, 2022.)
- Dubert García, Francisco. 1999. *Aspectos do galego de Santiago de Compostela*. Anexo 44 de Verba. Santiago de Compostela: Universidad de Santiago de Compostela.
- Elvira, Javier. 1998. *El cambio analógico*. Madrid: Gredos.
- Elvira, Javier. 2002. Sobre el desarrollo de la pasiva refleja en español medieval. In María Teresa Echenique Elizondo & Juan Pedro Sánchez Méndez (eds.), *Actas del V Congreso Internacional de la Historia de la Lengua Española*, vol. 1, 597–608. Madrid: Gredos.

- Enrique-Arias, Andrés & Miriam Bouzouita. 2013. La frecuencia textual en la evolución histórica de los clíticos pronominales en español. *Iberoromania* 77(1). 29–46.
<https://doi.org/10.1515/ibero-2013-0046>
- Ernout, Alfred & François Thomas. 1984. *Syntaxe Latine*. Paris: Éditions Klincksieck.
- eseuTen11: *European Spanish Web 2011* corpus. <https://app.sketchengine.eu/>. (9 February, 2022.)
- Estrada Arráez, Ana & Carlota de Benito Moreno. 2016. Variación en las redes sociales: Datos twilectales. *Revista Internacional de Lingüística Iberoamericana* 28. 77–111.
- Euskaltzaindia. 1987–2005. *Orotariko Euskal Hiztegia*. http://www.euskaltzaindia.eus/index.php?option=com_oeh&view=frontpage&Itemid=413&lang=eu. (9 February, 2022.)
- Fábregas, Antonio. 2021. Se in Spanish: Properties, structures, analyses. *Borealis* 10(2). 1–235.
<https://doi.org/10.7557/1.10.2.5934>
- Fernández Ramírez, Salvador. 1986. *Gramática española. 4. El verbo y la oración. (Volumen ordenado y completado por Ignacio Bosque)*. Arco/Libros: Madrid.
- Fernández-Ordóñez, Inés. 1999. Léismo, laísmo y loísmo. In Ignacio Bosque & Violeta Demonte (dirs.), *Gramática descriptiva de la lengua española*, vol. 1, 1317–1397. Madrid: Espasa.
- Fernández-Ordóñez, Inés. 2011. *La lengua de Castilla y la formación del español*. Madrid: Real Academia Española. http://www.rae.es/sites/default/files/Discurso_Ingreso_Ines_Fernandez_Ordenez.pdf. (9 February, 2022.)
- Fischer, Olga. 2008. On analogy as the motivation for grammaticalization. *Studies in Language* 32(2). 336–382. <https://doi.org/10.1075/sl.32.2.04fis>
- Frajzyngier, Zygmunt. 2000. Domains of point of view and coreferentiality: System interaction approach to the study of reflexives. In Zygmunt Frajzyngier & Traci S. Curl (eds.), *Reflexives: forms and functions*, 125–152. Amsterdam/Philadelphia: John Benjamins.
<https://doi.org/10.1075/tsl.40.06fra>
- Freixeiro Mato, Xosé Ramón. 2002. *Gramática da lingua galega. II Morfosintaxe*. Vigo: Edicións A Nosa Terra.
- García Fernández, Luis. 2015. Some reflections on verb with clitic increase. Verbs of motion. In Elisa Barrajón López, José Luis Cifuentes Honrubia & Susana Rodríguez Rosique (eds.), *Verb Classes and Aspects*, 264–287. Amsterdam & Philadelphia: John Benjamins.
<https://doi.org/10.1075/ivitra.9.12gar>
- García García, Juan Carlos. 1989. Comportamiento del /se/ en asturiano. *Lletres Asturianas* 31. 61–68.
- García-Miguel, José María. 1985. La voz media en español: Las construcciones pronominales con verbos transitivos. *Verba* 12. 307–343.
- García-Miguel, José María. 2001. Algunas motivaciones en la tipología de las variaciones de diátesis: Sistemas actanciales y polisemia de los morfemas de voz. In Augusto Soares da Silva (ed.), *Linguagem e cognição: A perspectiva da lingüística cognitiva*, 375–400. Braga: Associação Portuguesa de Lingüística/Universidade Católica Portuguesa. <http://weba575.webs.uvigo.es/jmgm/public/silva01.pdf>. (9 February, 2022.)
- García-Miguel, José María. 2010. Alternancias de diátesis y construcciones. Talk given at XXXIX Congreso Internacional de la Sociedad Española de Lingüística (Santiago de Compostela, 1–4 feb 2010). http://weba575.webs.uvigo.es/jmgm/public/SEL2010_Garcia_Miguel.pdf. (9 February, 2022.)
- García, Erica C. 1975. *The role of theory in linguistic analysis*. Amsterdam/Oxford: North Holland.

- Garita, Flor & Marta Rojas. 1993. Perdida de *se* en algunos verbos considerados tradicionalmente como transitivos, 201–214. In *Memoria del IV Congreso Costarricense de Filología, Lingüística y Literatura*. Universidad de Costa Rica.
- Geniušienė, Emma. 1987. *The typology of reflexives*. Berlin/New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110859119>
- Goldberg, Adele E. 2002. Surface generalizations: An alternative to alternations. *Cognitive Linguistics* 13(4). 327–356. <https://doi.org/10.1515/cogl.2002.022>
- Gómez Torrego, Leonardo. 1992. *Valores gramaticales de se*. Madrid: Arco/Libros.
- González Vergara, Carlos. 2006. Las construcciones no reflexivas con *se*. Una propuesta desde la gramática del papel y la referencia. Madrid: Universidad Complutense de Madrid PhD thesis.
- Gordon, Matthew J. 2005. Research aims and methodology / Forschungsziele und Methodologie. In Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill (eds.), *Sociolinguistics*, vol. 2, 955–965. Berlin & New York: De Gruyter. <https://doi.org/10.1515/9783110171488.2.7.955>
- Gutiérrez Ordóñez, Salvador. 1999. Los dativos. In Ignacio Bosque & Violeta Demonte (dirs.), *Gramática descriptiva del español*, vol. 2, 1855–1932. Madrid: Espasa.
- Haas, Walter. 2010. A study on areal diffusion. In Peter Auer & Jürgen E. Schmidt (eds.), *Language and space. An international handbook of linguistic variation*, 649–667. Berlin & New York: Mouton de Gruyter.
- Haiman, John. 1983. Iconic and economic motivation. *Language* 59(4). 781–819. <https://doi.org/10.2307/413373>
- Haiman, John. 1994. Iconicity. In Ron E. Asher (ed.), *The Encyclopedia of Language and Linguistics*, vol. 3, 1629–1633. Oxford/New York/Seoul/Tokyo: Pergamon Press.
- Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620553>
- Haspelmath, Martin. 1993. More on typology of inchoative/causative verb alternations. In Bernard Comrie & Maria Polinsky (eds.), *Causatives and transitivity*, 87–120. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/slcs.23.05has>
- Haspelmath, Martin. 2008. A frequentist explanation of some universals of reflexive marking. *Linguistic Discovery* 6(1). 40–63. <https://doi.org/10.1349/PS1.1537-0852.A.331>
- Heine, Bernd. 2000. Polysemy involving reflexive and reciprocal markers in African languages. In Zygmunt Frajzyngier & Traci S. Curl, (eds.), *Reciprocals. Forms and functions*, 1–29. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.41.02hei>
- Heredia, José Ramón. 1999. Sobre construcciones pronominales ‘medias’ en español. In María Carmen Losada Aldrey, José F. Márquez Caneda & Tomás E. Jiménez Juliá (coords.), *Español como lengua extranjera, enfoque comunicativo y gramática: Actas del IX congreso internacional de ASELE*, 551–560. Universidad de Santiago de Compostela: Servicio de Publicaciones. https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/09/09_0554.pdf. (9 February, 2022.)
- Hernández Terrés, José Miguel. 1984. Un tipo de construcciones pronominales en español: Los verbos con incrementación pronominal obligatoria. *ELUA* 2. 137–153. <https://doi.org/10.14198/ELUA1984.2.05>
- Hernández, César. 1966. Del *se* reflexivo al impersonal. *Archivum* XVI. 39–66.
- Hernanz, Maria Lluïsa. 1999. El infinitivo. In Ignacio Bosque & Violeta Demonte (dirs.), *Gramática descriptiva de la lengua española*, vol. 2, 2197–2356. Madrid: Espasa.
- Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2). 251–299. <https://doi.org/10.1353/lan.1980.0017>

- Inglese, Guglielmo. 2021. Towards a typology of middle voice systems. *Linguistic Typology*, aop. <https://doi.org/10.1515/lingty-2020-0131>
- Inglese, Guglielmo. Forthcoming. *The rise of middle voice systems: a study in diachronic typology*. Ms, KU Leuven.
- Instituto de la Lengua Gallega. 1975. *Gallego 1*. Santiago de Compostela: Universidad de Santiago de Compostela.
- Itkonen, Esa. 2005. Concerning the synthesis between intuition-based study of norms and observation-based study of corpora. *SKY Journal of Linguistics* 18. 357–377.
- Itkonen, Esa. 2013–4. On explanation in linguistics. *Energeia* 5. 10–40. <https://energeia-online.org/zeitschrift/2013/Itkonen-on-explanation-in-linguistics.html>. (9 February, 2022.)
- Janic, Katarzyna. 2016. *L'antipassif dans les langues accusatives*. Bruxelles: Peter Lang. <https://doi.org/10.3726/978-3-0352-6598-9>
- Jiménez-Fernández, Ángel Luis & Mercedes Tubino. 2014. Variación sintáctica en la causativización léxica. *Revista de la Sociedad Española de Lingüística* 44(1). 7–38.
- Kabatek, Johannes. 2000. *Os falantes como linguistas. Tradición, innovación e interferencias no galego actual*. Vigo: Xerais.
- Kailuweit, Rolf. 2011. Romance anticausatives: A constructionist RRG approach. In Wataru Nakamura (ed.), *New Perspectives in Role and Reference Grammar*, 104–133. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Kailuweit, Rolf. 2012. Construcciones anticausativas: El español comparado con el francés. In Valeriano Bellostá von Colbe & Marco García García (eds.), *Aspectualidad – Transitivity – Referencialidad. Las lenguas románicas en contraste*, 133–158. Frankfurt: Peter Lang.
- Kemmer, Suzanne. 1993. *The middle voice*. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.23>
- Kemmer, Suzanne. 1994. Middle voice, transitivity and the elaboration of events. In Barbara Fox & Paul. J. Hopper (eds.), *Voice. Form and function*, 170–230. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.27.09kem>
- Kittilä, Seppo & Fernando Zúñiga. 2010. Introduction. Benefaction and malefaction from a cross-linguistic perspective. In Fernando Zúñiga & Seppo Kittilä (eds.), *Benefactives and Malefactives. Typological perspectives and case studies*, 1–28. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.92.01kit>
- Klaiman, Miriam H. 1991. *Grammatical voice*. Cambridge: Cambridge University Press.
- Klein, Philipp. 1987. Syntax and semantics of Spanish se. *Hispanic Journal* 9. 149–162.
- Knjazev, Jurij P. 2007. Lexical reciprocals as a means of expressing reciprocal situations. In Vladimir P. Nedjalkov (ed.), *Reciprocal constructions*, vol 1, 115–146. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.71.07knj>
- Koch, Peter & Wolf Oesterreicher. 2007. *Lengua hablada en la Romania: Español, francés, italiano*. Madrid: Gredos.
- König, Ekkehard & Volker Gast. 2008. Reciprocity and reflexivity – description, typology and theory. In Ekkehard König & Volker Gast (eds.), *Reciprocals and reflexives. Theoretical and Typological Explorations*, 1–31. Berlin/New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110199147.1>
- König, Ekkehard & Martin Haspelmath. 1998. Les constructions à possesseur externe dans les langues d'Europe. In Jack Feuillet (ed.) *Actance et valence dans les langues de l'Europe*, 525–606. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110804485.525>
- Kulikov, Leonid. 2011. Voice Typology. In Jae Jung Song (ed.), *The Oxford Handbook of Linguistic Typology*, 368–398. Oxford: Oxford University Press.

- Kulikov, Leonid. 2013. Middle and reflexive. In Silvia Luraghui & Claudia Parodi (eds.), *The Bloomsbury companion to syntax*, 261–280. London/New York: Bloomsbury.
- Labov, William. 1972. Some principles of linguistic methodology, *Language in Society* 1. 97–120. <https://doi.org/10.1017/S0047404500006576>
- Lamiroy, B. 2003. Grammaticalization and external possessor structures. In Martine Coene & Yves D'hulst (eds.), *From NP to DP*, vol. 2, 257–280. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/la.56.15lam>
- Lantes Vitureira, Xesus. 1980. Os pronomes reflexivos en galego. Santiago de Compostela: Universidade de Santiago de Compostela BA thesis.
- Lantes Vitureira, Xesus. 1986. O uso do 'reflexivo' galego. *Concepción Arenal Ciencias y Humanidades* 14, 24–36.
- Lapesa, Rafael. 2000. Morfosintaxis histórica del verbo español. In Rafael Lapesa, *Estudios de morfosintaxis histórica del español*, 730–885. Madrid: Gredos.
- Larochette, Joe. 1939. Les aspects verbaux en espagnol ancien. *Revue des langues romanes* 8. 327–421.
- Linzen, Tal. 2009. Hebrew possessive dative: corpus evidence for the role of affectedness. Tel Aviv: Tel Aviv University MA thesis.
- Luque Castro, Antonio. 2006. Sobre el desarrollo de las predicaciones inacusativas en el español medieval. In José J. Bustos Tovar & José L. Girón Alconchel (coords.), *Actas del VI Congreso Internacional de Historia de la Lengua española*, vol 1, 855–868. Madrid: Arco-Libros.
- Luque Castro, Antonio. 2007. Sobre la sintaxis histórica de las predicaciones reflexivas indirectas. *Interlingüística* 17. 648–657.
- Maldonado, Ricardo. 1988. Energetic reflexives in Spanish. *Berkeley Linguistics Society* 13. 34–52. <https://doi.org/10.3765/bls.v14i0.1772>
- Maldonado, Ricardo. 1989. Se gramaticalizó: A diachronic account of energetic reflexives in Spanish. In Ralph Carlson, Scott DeLancey, Spike Gildea, Doris Payne & Aseem Saxena (eds.), *Proceedings of the Fourth Meeting of the Pacific Linguistics Conference*, 339–360. Eugene: Dept. of Linguistics, University of Oregon.
- Maldonado, Ricardo. 1999. *A media voz*. México DF: Universidad Nacional Autónoma de México.
- Maldonado, Ricardo. 2000. Conceptual distance and transitivity increase in Spanish reflexives. In Zygmunt Frajzyngier & Traci S. Curl (eds.), *Reflexives: Forms and functions*, 153–185. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.40.07mal>
- Martín Cid, Manuel. 2004. El significado *se*. *CAUCE, Revista de Filología, Comunicación y sus Didácticas* 27. 247–263.
- Martín Zorraquino, María Antonia. 1979. *Las construcciones pronominales: Paradigma y desviaciones*. Madrid: Gredos.
- Martín Zorraquino, María Antonia. 1993. *Ir e irse* en el Cantar de Mío Cid. *Aragón en la Edad Media* 10–11. 575–578.
- Martín Zorraquino, María Antonia. 1998. *Venir/venirse* en el Cantar de Mío Cid. In Claudio García Turza, Fabián González Bachiller & José J. Mangado Martínez, *Actas del IV Congreso Internacional de Historia de la Lengua Española*, vol. 1, 581–596. La Rioja: Universidad de La Rioja.
- Martín Zorraquino, María Antonia. 2002. Las construcciones pronominales en los textos del primitivo romance hispánico. In Carmen Saralegui & Manuel Casado (eds.), *Pulchre, bene, recte. Estudios en homenaje al profesor F. González Ollé*, 843–884. Pamplona: EUNSA.

- Martínez, José Antonio. 1981. Sobre el /se/ en español. *Revista de Filología de la Universidad de la Laguna* 0. 87–92.
- Maslova, Elena (2008): “Reflexive encoding of reciprocity: Cross-linguistic and language-internal variation” in E. König / V. Gast (eds.), *Reciprocals and Reflexives. Theoretical and Typological Explorations*. Berlin / New York: Mouton de Gruyter, 225–257.
- McCready, Eric & Chiyo Nishida. 2008. Reflexive intransitives in Spanish and event semantics. In Johannes Dölling, Tatjana Heyde-Zybatow & Martin Schäfer (eds.), *Event Structures in Linguistic Form and Interpretation*, 223–243. Berlin & New York: Walter de Gruyter. <https://doi.org/10.1515/9783110925449.223>
- Melis, Chantal & Silvia Peña-Alfaro. 2007. El desarrollo histórico de la pasiva con *se* en español. *Romance Philology* 61(1). 49–77. <https://doi.org/10.1484/J.RPH.2.305824>
- Mendikoetxea, Amaya. 1999a. Construcciones inacusativas y pasivas. In Ignacio Bosque & Violeta Demonte (dirs.), *Gramática descriptiva del español*, vol. 2, 1575–1630. Madrid: Espasa.
- Mendikoetxea, Amaya. 1999b. Las construcciones con *se*: Medias, pasivas e impersonales. In Ignacio Bosque & Violeta Demonte (dirs.), *Gramática descriptiva del español*, vol. 2, 1631–1722. Madrid: Espasa.
- Mendikoetxea, Amaya. 2000. Relaciones de interficie: Los verbos de cambio de estado. In Ana Bravo, Carlos Luján & Isabel Pérez (eds.), *Cuadernos de Lingüística VII*, 125–144. Madrid: Instituto Universitario Ortega y Gasset.
- Merino Hernández, Laura Margarita. 2019. Condicionamiento pragmático y semántico del marcador de voz media *se* en el verbo *morirse*. *Cuadernos de Lingüística de El Colegio de México* 6(1), e81. <https://doi.org/10.24201/clecm.v6i1.81>
- Monge, Félix. 1955. Las frases pronominales de sentido impersonal en español. *Archivo de Filología Aragonesa* 7. 7–102.
- Morimoto, Yuko. 2008. *Me estuve quieto*: El concepto de estado y el llamado *se* aspectual. In Inés Olza Moreno, Manuel Casado Valverde & Ramón González Ruíz (eds.), *Actas Del XXXVII Simposio Internacional de La Sociedad Española de Lingüística (SEL)*, 591–600. Pamplona: Servicio de Publicaciones de la Universidad de Navarra.
- Mous, Maarten. 2001. The middle in Cushitic Languages. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, 75–86. <https://doi.org/10.3765/bls.v27i2.1081>
- Mous, Maarten. 2007. The middle and passive derivations in Konso. In Azeb Amha, Maarten Mous & Graziano Savà, *Omotiic and Cushitic language studies. Papers from the fourth Cushitic Omotic Conference*, 213–229. Köln: Rüdiger Köppe Verlag.
- Moyse-Faurie, Claire. 2008. Constructions expressing middle, reflexive and reciprocal situations in some Oceanic languages. In Ekkehard König & Volker Gast (eds.), *Reciprocal and reflexives: Theoretical and typological explorations*, 105–168. Berlin/New York: Mouton de Gruyter.
- Nedjalkov, Vladimir P. 2007a. Overview of the research. Definition of terms, framework, and related issues. In Vladimir P. Nedjalkov (ed.), *Reciprocal constructions*, vol 1, 3–114. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.71.06ned>
- Nedjalkov, Vladimir P. 2007b. Encoding of the reciprocal meaning. In Vladimir P. Nedjalkov (ed.), *Reciprocal constructions*, vol 1, 147–207. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.71.08ned>
- Nedjalkov, Vladimir P. 2007c. Polysemy of reciprocal markers. In Vladimir P. Nedjalkov (ed.), *Reciprocal constructions*, vol 1, 231–333. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.71.10ned>

- Newmeyer, Frederick J. 2005. *Possible and probable languages. A generative perspective on linguistic typology*. Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199274338.001.0001>
- NGLE (2009) = RAE & ASALE. 2009. *Nueva gramática de la lengua española*. 2 vols. Madrid: Espasa.
- Nishida, Chiyo. 1994. The Spanish reflexive clitic *se* as an aspectual class marker. *Linguistics* 32. 425–458. <https://doi.org/10.1515/ling.1994.32.3.425>
- Nuzzo, Regina. 2014. Statistical errors: P values, the ‘gold standard’ of statistical validity, are not as reliable as many scientists assume. *Nature* 506(7487). 150–52.
<https://doi.org/10.1038/506150a>
- OLD. 1982 = *Oxford Latin Dictionary*. Oxford: Oxford University Press.
- Otero, Carlos Peregrín. 1999. Pronombres reflexivos y recíprocos. In Ignacio Bosque & Violeta Demonte (dirs.), *Gramática descriptiva de la lengua española*, vol. 1, 1427–1518. Madrid: Espasa.
- Otero, Carlos Peregrín. 2002. Facetas de *se*. In Cristina Sánchez López (ed.), *Las construcciones con se*. Madrid: Visor, 165–206.
- Pato, Enrique. 2003. La sustitución del imperfecto de subjuntivo por el condicional simple y el imperfecto de indicativo en el castellano septentrional peninsular. (Estudio de variación dialectal). Madrid: Universidad Autónoma de Madrid PhD thesis.
- Pato, Enrique. 2012. Variación dialectal y análisis estadístico: Formas indicativas por subjuntivas en español rural. In Enrique Pato y Javier Rodríguez Molina (eds.), *Estudios de filología y lingüística españolas. Nuevas voces en la disciplina*, 93–133. Bern: Peter Lang.
<https://doi.org/10.3726/978-3-0351-0357-1/5>
- Pedersen, Johan. 2005. The Spanish impersonal *se*-construction: Constructional variation and change. *Constructions* 1. 1–49. <https://www.constructions.uni-osnabrueck.de/wp-content/uploads/2014/06/2005-Pedersen-3071-6091-1-PB.pdf>. (9 February, 2022.)
- Pena, Jesús. 1982. La voz en español. Intento de caracterización. *Verba* 9. 215–252.
- Polinsky, Maria. 2013. Antipassive Constructions. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/108>. (9 February, 2022.)
- Portilla, Mario. 2007. Diacronía de las construcciones con pronombres reflexivos en español. *Filología y Lingüística* XXXIII(1). 131–149.
- RAE. 1973. *Esbozo de una nueva gramática de la lengua española*. Madrid: Espasa.
- Rapold, Christian J. 2010. Beneficiary and other roles of the dative in Tashelhiyt. In Fernando Zúñiga & Seppo Kittilä (eds.), *Benefactives and malefactives. Typological perspectives and case studies*, 351–376. Amsterdam/Philadelphia: John Benjamins.
<https://doi.org/10.1075/tsl.92.15rap>
- Richter, Elise. 1909. Zur Entwicklung des reflexiven Ausdrucks in Romanischen. *Zeitschrift für Romanische Philologie* 33(2). 135–42. <https://doi.org/10.1515/zrph.1909.33.2.135>
- Riemer, Nicholas. 2008. Grammaticality as evidence and as prediction in a Galilean linguistics. *Language Sciences* 31(5). 612–633. <https://doi.org/10.1016/j.langsci.2008.04.001>
- Rigau, Gemma. 1994. Les propietats dels verbs pronominals. *Els Marges* 50. 29–39.
- Rivas, Javier. 2011. “Como no me trago el humo... :” A corpus-based approach to aspectual *se*. *Studies in Hispanic and Lusophone Linguistics* 4(2). 379–416.
<https://doi.org/10.1515/shll-2011-1106>
- Rodríguez Arrizabalaga, Beatriz. 2005. Construcciones atributivas con *se*. *Verba* 32. 103–120.

- Rodríguez Molina, Javier. 2010. La gramaticalización de los tiempos compuestos: Cinco cambios diacrónicos. Madrid: Universidad Autónoma de Madrid PhD thesis. <https://repositorio.uam.es/handle/10486/6279>. (9 February, 2022.)
- Rodríguez Ramalle, Teresa María. 2005. *Manual de sintaxis del español*. Madrid: Castalia.
- Romero Pascual, Cristina & Ismael Iván Teomiro García. 2012. La relación entre estructura eventiva y papeles temáticos: El *se* aspectual del español. *Revista de Filología Románica* 29(2). 233–243. https://doi.org/10.5209/rev_RFRM.2012.v29.n2.40156
- Sánchez López, Cristina. 2002. Las construcciones con *se*. Estado de la cuestión. In Cristina Sánchez López (ed.), *Las construcciones con se*, 13–163. Madrid: Visor.
- Sanchis Guarner, Manuel. 1953. *La cartografía lingüística en la actualidad y el Atlas de la Península Ibérica*. Madrid: CSIC.
- Sanz, Montserrat & Itziar Laka. 2002. Oraciones transitivas con *se*. El modo de acción en la sintaxis. In Cristina Sánchez López (ed.), *Las construcciones con se*, 309–391. Madrid: Visor.
- Saussure, Ferdinand de. 1945 [1916]. *Curso de lingüística general*. Buenos Aires: Losada.
- Schäfer, Florian. 2008. *The syntax of (anti-)causatives*. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/la.126>
- Schäfer, Florian & Margot Vivanco. 2015. Reflexively marked anticausatives are not semantically reflexive. In Enoch Anoh, Jeannette Schaeffer & Petra Sleeman (eds.), *Romance Languages and Linguistic Theory 2013. Selected papers from 'Going Romance' Amsterdam 2013*, 203–219. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/rllt.8.11sch>
- Schütze, Carson T. 1996. *The empirical base of linguistics*. Chicago/London: The University of Chicago Press.
- Segura Mungaía, Santiago. 1985. *Diccionario etimológico latino-español*. Madrid: Anaya.
- Signes Codoñer, Juan. 2016. *La quimera de los gramáticos*. Salamanca: Ediciones Universidad de Salamanca. <https://doi.org/10.2307/j.ctt1j2n8kb>
- Silverstein, Michael. 1972. Chinook jargon: Language contact and the problem of multi-level generative systems, II. *Language* 48. 596–625. <https://doi.org/10.2307/412037>
- Smith, Tomoko Yamashita. 2010. Cross-linguistic categorizations of benefactives by event structure. In Fernando Zúñiga & Seppo Kittilä (eds.), *Benefactives and malefactives. Typological perspectives and case studies*, 71–95. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.92.03smi>
- Suárez Fernández, José. 1990. Los bloqueadores funcionales átonos. *Lletres Asturianas* 35. 39–44.
- Teomiro García, Ismael Iván. 2013. Low applicatives and optional *se* in Spanish non-anticausative intransitive verbs. *Revista de Lingüística y Lenguas Aplicadas* 8. 140–153. <https://doi.org/10.4995/rlyla.2013.1354>
- Timberlake, Alan. 1977. Reanalysis and actualization in syntactic change. In Charles N. Li (ed.), *Mechanisms of syntactic change*, 141–177. Austin/London: University of Texas Press. <https://doi.org/10.7560/750357-006>
- Torres Cacoullos, Rena & Scott A. Schwenter. 2008. Constructions and pragmatics: Variable middle marking in Spanish *subir(se)* 'go up' and *bajar(se)* 'go down'. *Journal of Pragmatics* 40(8). 1455–77. <https://doi.org/10.1016/j.pragma.2008.01.005>
- Turley, Jeffrey S. 1998. A prototype analysis of Spanish indeterminate reflexive constructions. *Language Sciences* 20(2). 137–162. [https://doi.org/10.1016/S0388-0001\(97\)00027-2](https://doi.org/10.1016/S0388-0001(97)00027-2)
- Van Gorp, Lise. 2014. El porqué de la decena de verbos pseudo-copulativos de cambio en español. Hacia una aclaración cognitiva y funcional. Leuven: KU Leuven PhD thesis.

- Vera Luján, Agustín. 1996. Esquemas oracionales ergativos reflexivos. *ELUA* 11. 385–409.
<https://doi.org/10.14198/ELUA1996-1997.11.19>
- Vicente Mateu, Juan A. 2001. Entre la reflexividad y la inacusatividad: Algunas consideraciones sobre construcciones con el verbo *romper*. *Revista de investigación lingüística* 2(IV). 83–101.
- Vivanco, Margot. 2016. Causatividad y cambio de estado en español. La alternancia causativo-inacusativa. Madrid: Universidad Complutense de Madrid PhD thesis.
- Vivanco, Margot. 2021. Scalar constraints on anticausative *se*: The aspectual hypothesis revisited. In Grant Armstrong & Jonathan E. MacDonald (eds.), *Unraveling the complexity of SE. Studies in Natural Language and Linguistic Theory* 99, 291–318. Cham: Springer.
https://doi.org/10.1007/978-3-030-57004-0_12
- Waltereit, Richard. 2000. What it means to deceive yourself: The semantic relation of French reflexive verbs and their corresponding transitive verbs. In Zygmunt Frajzyngier & Traci S. Curl (eds.), *Reflexives: Forms and functions*, 257–278. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.40.10wal>
- Wasserstein, Ronald L. & Nicole A. Lazar. 2016. The ASA's Statement on p -Values: Context, process, and purpose. *The American Statistician* 70(2). 129–33.
<https://doi.org/10.1080/00031305.2016.1154108>
- Weinreich, Uriel, William Labov & Marvin I. Herzog. 1968. Empirical Foundations for a Theory of Language Change. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Directions for Historical Linguistics. A symposium*, 95–188. Austin/London: University of Texas Press.
- Willems, Klaas. 2012. Intuition, introspection and observation in linguistic inquiry. *Language Sciences* 34. 665–681. <https://doi.org/10.1016/j.langsci.2012.04.008>
- Willems, Klaas. 2013–4. Is frequency an explanatory causal concept in linguistics? *Energia* 5. 105–112. <https://energeia-online.org/zeitschrift/2013/Klaas-Willems-is-frequency-an-explanatory-causal-concept.html>. (9 February, 2022.)
- Zagona, Karen. 1996. Compositionality of aspect: Evidence from Spanish aspectual *se*. In Claudia Parodi (ed.), *Aspects of Romance linguistics*, 475–488. Washington D.C.: Georgetown University Press.
- Zúñiga, Fernando & Seppo Kittilä. 2019. *Grammatical voice*. Cambridge: Cambridge University Press.
- Zribi-Hertz, Anne. 1982. La construction *se-moyen* du français et son statut dans le triangle: moyen/passive/réfléchi. *Linguisticae Investigationes* 6. 345–401.
<https://doi.org/10.1075/li.6.2.05zri>
- Zyzik, Eve. 2006. Transitivity alternations and sequence learning: Insights from L2 Spanish production data. *Studies in Second Language Acquisition* 28. 449–485.
<https://doi.org/10.1017/S0272263106060190>

APPENDIX 1

Localities interviewed with the corpus COSER

- COSER 103: El Burgo (Álava), 1 hour and 1 minute, Sucorpus NE. Spanish.
COSER 106: Leza (Álava), 1 hour and 21 minutes, Sucorpus NE. Spanish.
COSER 107: Luzuriaga (Álava), 1 hour and 33 minutes, Sucorpus E. Spanish.
COSER 109: Menagaray (Álava), 51 minutes, Sucorpus E. Spanish.
COSER 204: Barrax (Albacete), 2 hours, Sucorpus NE. Spanish.
COSER 211: Higuera (Albacete), 1 hour and 43 minutes, Sucorpus E. Spanish.
COSER 214: Liétor (Albacete), 1 hour and 28 minutes, Sucorpus NE. Spanish.
COSER 222: Povedilla (Albacete), 1 hour and 42 minutes, Sucorpus NE. Spanish.
COSER 304: Busot (Alicante), 54 minutes, Sucorpus NE. Spanish.
COSER 310: Salinas (Alicante), 2 hours and 5 minutes, Sucorpus E. Spanish.
COSER 404: Bares (Almería), 1 hour and 51 minutes, Sucorpus E. Spanish.
COSER 425: Terque (Almería), 1 hour and 15 minutes, Sucorpus NE. Spanish.
COSER 506: Alea – Linares (Ribadesella, Asturias), 1 hour and 7 minutes, Sucorpus E. Spanish.
COSER 509: Fechaladrona – Villoria (Laviana, Asturias), 1 hour and 28 minutes, Sucorpus E. Spanish.
COSER 525: Colinas de Abajo (Sangoñedo, Tineo, Asturias), 2 hours and 22 minutes, Sucorpus E. Spanish.
COSER 528: Grullas (Candamo, Asturias), 56 minutes, Sucorpus E. Spanish.
COSER 604: Burgo (Ávila), 1 hour (2 interviews), Sucorpus NE. Spanish.
COSER 607: El Barco de Ávila (Ávila), 55 minutes, Sucorpus NE. Spanish.
COSER 609: Madrigal de las Altas Torres (Ávila), 54 minutes, Sucorpus E. Spanish.
COSER 614: Narros del Puerto (Ávila), 54 minutes, Sucorpus NE. Spanish.
COSER 716: La Garrovilla (Badajoz), 1 hour, Sucorpus NE. Spanish.
COSER 723: Orellana de la Sierra (Badajoz), 2 hours and 12 minutes, Sucorpus NE. Spanish.
COSER 726: San Francisco de Olivenza (Badajoz), 56 minutes, Sucorpus E. Spanish.
COSER 728: Valencia del Ventoso (Badajoz), 1 hour and 47 minutes, Sucorpus NE. Spanish.
COSER 922: Humada (Burgos), 1 hour and 44 minutes, Sucorpus E. Spanish.
COSER 924: La Horra (Burgos), 1 hour and 14 minutes, Sucorpus NE. Spanish.
COSER 934: Pedruzo (Condado de Treviño, Burgos), 1 hour and 54 minutes, Sucorpus NE. Spanish.
COSER 939: Quintana de los Prados (Burgos), 1 hour and 2 minutes, Sucorpus NE. Spanish.
COSER 959: Villaverde–Mogina (Burgos), 2 hours and 44 minutes, Sucorpus NE. Spanish.
COSER 1012: Jarandilla de la Vera (Cáceres), 1 hour and 1 minute, Sucorpus NE. Spanish.
COSER 1014: Moraleja (Cáceres), 47 minutes, Sucorpus NE. Spanish.
COSER 1015: Naval Moral de la Mata (Cáceres), 2 hours and 43 minutes, Sucorpus NE. Spanish.
COSER 1020: Talaván (Cáceres), 1 hour and 45 minutes, Sucorpus E. Spanish.
COSER 1023: Campo Lugar (Cáceres), 42 minutes, Sucorpus E. Spanish.
COSER 1102: Algar (Cádiz), 55 minutes, Sucorpus NE. Spanish.

- COSER 1107: Espera (Cádiz), 1 hour, Sucorpus E. Spanish.
- COSER 1116: San José de Malcocinado (Cádiz), 59 minutes, Sucorpus E. Spanish.
- COSER 1205: Castrillo de Valdelomar (Cantabria), 1 hour and 24 minutes, Sucorpus E. Spanish.
- COSER 1212: Ledantes (Cantabria), 43 minutes, Sucorpus E. Spanish.
- COSER 1219: Portillo de Solórzano (Cantabria), 59 minutes, Sucorpus E. Spanish.
- COSER 1228: Silió (Cantabria), 41 minutes, Sucorpus E. Spanish.
- COSER 1232: Vega de Pas (Cantabria), 1 hour and 7 minutes, Sucorpus E. Spanish.
- COSER 1307: Jérica (Castellón), 2 hours and 15 minutes, Sucorpus E. Spanish.
- COSER 1318: Lucena del Cid (Castellón), 1 hour and 13 minutes, Sucorpus NE. Spanish.
- COSER 1401: Aldea del Rey (Ciudad Real), 50 minutes, Sucorpus NE. Spanish.
- COSER 1404: Argamasilla de Alba (Ciudad Real), 1 hour, Sucorpus E. Spanish.
- COSER 1414: Malagón (Ciudad Real), 1 hour and 21 minutes, Sucorpus E. Spanish.
- COSER 1417: Porzuna (Ciudad Real), 1 hour and 34 minutes, Sucorpus NE. Spanish.
- COSER 1503: Aguilar de la Frontera (Córdoba), 1 hour and 31 minutes, Sucorpus NE. Spanish.
- COSER 1514: Santaella (Córdoba), 50 minutes, Sucorpus E. Spanish.
- COSER 1603: Belmonte (Cuenca), 43 minutes, Sucorpus NE. Spanish.
- COSER 1607: Cardenete (Cuenca), 1 hour and 6 minutes, Sucorpus E. Spanish.
- COSER 1634: Valeria (Cuenca), 1 hour and 24 minutes, Sucorpus NE. Spanish.
- COSER 1636: Villaconejos de Trabaque (Cuenca), 2 hours and 2 minutes, Sucorpus NE. Spanish.
- COSER 1834: Tablones (Granada), 2 hours and 11 minutes, Sucorpus E. Spanish.
- COSER 1838: Ventorros de San José (Granada), 1 hours and 29 minutes, Sucorpus NE. Spanish.
- COSER 1901: Alboreca (Guadalajara), 54 minutes, Sucorpus NE. Spanish.
- COSER 1902: Canredondo (Guadalajara), 1 hour and 15 minutes, Sucorpus NE. Spanish.
- COSER 1907: Cubillejos del Sitio (Guadalajara), 1 hour and 2 minutes, Sucorpus E. Spanish.
- COSER 1921: Yebra (Guadalajara), 1 hour and 45 minutes, Sucorpus NE. Spanish.
- COSER 2001: Aguinaga (Guipúzcoa), 1 hour and 19 minutes (2 interviews), Sucorpus NE. Spanish.
- COSER 2004: Gabiria (Guipúzcoa), 58 minutes, Sucorpus NE. Spanish.
- COSER 2005: Lizarza (Guipúzcoa), 50 minutes, Sucorpus E. Spanish.
- COSER 2122: Zufre (Huelva), 1 hour and 20 minutes, Sucorpus E. Spanish.
- COSER 2206: Banastón (Huesca), 1 hour and 40 minutes, Sucorpus NE. Spanish.
- COSER 2207: Bandaliés (Huesca), 2 hours and 9 minutes, Sucorpus NE. Spanish.
- COSER 2222: Oliván (Huesca), 1 hour and 11 minutes, Sucorpus NE. Spanish.
- COSER 2229: Tramaced (Huesca), 1 hour and 11 minutes, Sucorpus E. Spanish.
- COSER 2301: Arjona (Jaén), 1 hour and 38 minutes, Sucorpus E. Spanish.
- COSER 2303: Cabra del Santo Cristo (Jaén), 1 hour and 20 minutes, Sucorpus NE. Spanish.
- COSER 2314: Los Pascuales (Jaén), 1 hour, Sucorpus NE. Spanish.
- COSER 2321: Santo Tomé (Jaén), 48 minutes, Sucorpus NE. Spanish.
- COSER 2401: Orellán (La Coruña), 2 hours and 16 minutes, Sucorpus E. Galician.
- COSER 2402: Santiago (La Coruña), 1 hour and 12 minutes, Sucorpus E. Spanish.
- COSER 2403: Zas (La Coruña), 1 hour and 58 minutes, Sucorpus E. Spanish.
- COSER 2501: Ausejo (La Rioja), 3 hour and 16 minutes, Sucorpus NE. Spanish.
- COSER 2506: Huércanos (La Rioja), 55 minutes (2 interviews), Sucorpus E. Spanish.
- COSER 2512: Nestares (La Rioja), 59 minutes, Sucorpus NE. Spanish.
- COSER 2515: Sajazarra (La Rioja), 48 minutes, Sucorpus NE. Spanish.
- COSER 2606: Cifuentes de Rueda (León), 1 hour and 15 minutes, Sucorpus E. Spanish.

- COSER 2614: Lorenzana (León), 1 hour and 2 minutes, Sucorpus E. Spanish.
 COSER 2627: Santa Colomba de Curueño (León), 54 minutes, Sucorpus E. Spanish.
 COSER 2637: Almazcara (León), 1 hour and 5 minutes, Sucorpus E. Spanish.
 COSER 2644: Lucillo (León), 1 hour and 37 minutes, Sucorpus E. Spanish.
 COSER 2801: Guitiriz (Lugo), 1 hour and 3 minutes, Sucorpus E. Spanish.
 COSER 2803: Teixeira (Lugo), 1 hour and 24 minutes, Sucorpus E. Galician.
 COSER 2907: Humanes (Madrid), 2 hours and 35 minutes, Sucorpus E. Spanish.
 COSER 2910: Manzanares el Real (Madrid), 1 hour and 5 minutes, Sucorpus NE. Spanish.
 COSER 2914: Sieteiglesias (Madrid), 1 hour and 32 minutes, Sucorpus NE. Spanish.
 COSER 2915: Valdilecha (Madrid), 1 hour and 16 minutes, Sucorpus NE. Spanish.
 COSER 3001: Antequera (Málaga), 1 hour and 51 minutes, Sucorpus E. Spanish.
 COSER 3002: La Atalaya (Málaga), 1 hour and 20 minutes, Sucorpus NE. Spanish.
 COSER 3107: Fuente del Pino (Murcia), 1 hour and 40 minutes, Sucorpus E. Spanish.
 COSER 3109: Lo Ferro (Murcia), 1 hour and 19 minutes, Sucorpus NE. Spanish.
 COSER 3203: Azcona (Navarra), 2 hours and 2 minutes (2 interviews), Sucorpus NE. Spanish.
 COSER 3209: Etxauri (Navarra), 31 minutes, Sucorpus NE. Spanish.
 COSER 3214: Leitza (Navarra), 1 hour and 8 minutes, Sucorpus NE. Spanish.
 COSER 3222: Miranda de Arga (Navarra), 1 hour and 17 minutes, Sucorpus E. Spanish.
 COSER 3228: Mérida (Navarra), 45 minutes, Sucorpus NE. Spanish.
 COSER 3301: Astariz (Orense), 1 hour and 20 minutes, Sucorpus E. Galician / Spanish.
 COSER 3302: Cádavos (Orense), 1 hour and 21 minutes, Sucorpus E. Spanish.
 COSER 3303: O Malladoiro (Orense), 1 hour and 10 minutes, Sucorpus E. Galician.
 COSER 3402: Astudillo (Palencia), 1 hour and 2 minutes, Sucorpus NE. Spanish.
 COSER 3412: Muñeca (Palencia), 1 hour and 16 minutes, Sucorpus NE. Spanish.
 COSER 3414: Olmos de Ojeda (Palencia), 1 hour and 22 minutes, Sucorpus E. Spanish.
 COSER 3423: Santervás de la Vega (Palencia), 1 hour and 28 minutes, Sucorpus E. Spanish.
 COSER 3426: Valle de Cerrato (Palencia), 1 hour and 16 minutes, Sucorpus E. Spanish.
 COSER 3601: Alaraz (Salamanca), 1 hour and 4 minutes, Sucorpus NE. Spanish.
 COSER 3610: Palencia de Negrilla (Salamanca), 1 hour and 2 minutes, Sucorpus NE. Spanish.
 COSER 3611: Peralejos de Abajo (Salamanca), 45 minutes, Sucorpus NE. Spanish.
 COSER 3614: Puebla de Yeltes (Salamanca), 2 hours and 10 minutes, Sucorpus E. Spanish.
 COSER 3701: Anaya (Segovia), 1 hour and 3 minutes, Sucorpus NE. Spanish.
 COSER 3706: Moraleja de Cuéllar (Segovia), 1 hour and 15 minutes, Sucorpus NE. Spanish.
 COSER 3707: Muñozeros (Segovia), 1 hour and 11 minutes, Sucorpus NE. Spanish.
 COSER 3712: Santiuste de San Juan Bautista (Segovia), 59 minutes, Sucorpus E. Spanish.
 COSER 3806: Almadén de la Plata (Sevilla), 1 hour and 39 minutes, Sucorpus E. Spanish.
 COSER 3814: Constantina (Sevilla), 1 hour and 11 minutes, Sucorpus NE. Spanish.
 COSER 3901: Almajano (Soria), 2 hours and 10 minutes, Sucorpus E. Spanish.
 COSER 3916: Tarancueña (Soria), 1 hour and 20 minutes, Sucorpus E. Spanish.
 COSER 3923: Aguaviva de la Vega (Soria), 1 hour and 8 minutes, Sucorpus NE. Spanish.
 COSER 3924: Beratón (Soria), 1 hour and 32 minutes, Sucorpus NE. Spanish.
 COSER 4102: Alcalá de la Selva (Teruel), 1 hour and 38 minutes, Sucorpus E. Spanish.
 COSER 4108: Bronchales (Teruel), 1 hour and 12 minutes, Sucorpus NE. Spanish.
 COSER 4117: Fuentes Claras (Teruel), 1 hour and 47 minutes, Sucorpus NE. Spanish.
 COSER 4128: Perales de Alfambra (Teruel), 52 minutes (2 interviewss), Sucorpus NE. Spanish.
 COSER 4206: Caleruela (Toledo), 57 minutes, Sucorpus E. Spanish.
 COSER 4214: La Nava de Ricomalillo (Toledo), 1 hour and 8 minutes, Sucorpus NE. Spanish.

- COSER 4218: Los Navalmorales (Toledo), 1 hour and 6 minutes, Sucorpus NE. Spanish.
COSER 4227: Pulgar (Toledo), 1 hour and 12 minutes, Sucorpus NE. Spanish.
COSER 4230: Tembleque (Toledo), 1 hour and 1 minute, Sucorpus NE. Spanish.
COSER 4301: Mas de los Mudos (Valencia), 1 hour, Sucorpus NE. Spanish.
COSER 4310: Enguera (Valencia), 2 hours and 14 minutes, Sucorpus E. Spanish.
COSER 4401: Arrabal del Portillo (Valladolid), 38 minutes, Sucorpus NE. Spanish.
COSER 4403: Casasola de Arión (Valladolid), 45 minutes, Sucorpus NE. Spanish.
COSER 4407: Cigales (Valladolid), 48 minutes, Sucorpus NE. Spanish.
COSER 4417: Rábano (Valladolid), 59 minutes, Sucorpus NE. Spanish.
COSER 4419: Velascálvaro (Valladolid), 41 minutes, Sucorpus NE. Spanish.
COSER 4421: Villalba de la Loma (Valladolid), 1 hour, Sucorpus E. Spanish.
COSER 4501: Aulesti (Vizcaya), 1 hour and 47 minutes, Sucorpus NE. Spanish.
COSER 4503: Ermua (Vizcaya), 1 hour and 12 minutes, Sucorpus NE. Spanish.
COSER 4504: Errigoitia (Vizcaya), 33 minutes, Sucorpus E. Spanish.
COSER 4506: Gordejuela (Vizcaya), 57 minutes, Sucorpus NE. Spanish.
COSER 4508: Lequeitio (Vizcaya), 1 hour and 58 minutes, Sucorpus NE. Spanish.
COSER 4602: Cotanes (Zamora), 1 hour and 24 minutes, Sucorpus NE. Spanish.
COSER 4611: Villalba de Lampreana (Zamora), 54 minutes, Sucorpus NE. Spanish.
COSER 4613: Villamor de los Escuderos (Zamora), 1 hour and 8 minutes, Sucorpus NE. Spanish.
COSER 4617: Mahíde (Zamora), 1 hour, Sucorpus E. Spanish.
COSER 4706: Almonacid de la Cuba (Zaragoza), 47 minutes, Sucorpus E. Spanish.
COSER 4714: Mesones de Isuela (Zaragoza), 1 hour and 1 minute, Sucorpus NE. Spanish.
COSER 4719: San Mateo de Gállego (Zaragoza), 1 hour and 30 minutes, Sucorpus E. Spanish.
COSER 4720: Urriés (Zaragoza), 1 hour, Sucorpus NE. Spanish.

APPENDIX 2

Questionnaire videos

1. Subir la escalera 'to climb the stairs'
2. Bajar la escalera 'to climb down the stairs'
3. Entrar en casa 'to go into the house'
4. Irse de casa 'to leave the house'
5. Quedarse en casa 'to stay at home'
6. Acercarse 'to get closer'
7. Llevarse (la silla en la que está sentada) 'to take (the chair she is sitting on)'
8. Traerse una silla (al llegar y ver que no hay) 'to bring a chair (after seeing there is not one for her)'
9. Resbalar (en el suelo mojado) 'to slip down'
10. Tropezar (con una piedra) 'to stumble'
11. Romperse, partirse (una zanahoria) 'to break'
12. Tropezar con un cubo, caerse el contenido 'to stumble with a bucket, to fall out (the content)'
13. Caer la pelota de la mesa 'to fall down the chair from the table'
14. Caérsele un libro 'to fall down (a book, from someone who is carrying it)'
15. Reírse de alguien 'to laugh at someone'
16. Tregar / subir al árbol, fiarse del árbol 'to climb the tree, to trust the tree'
17. Caerse de la silla 'to fall down from a chair'
18. Apoyarse en la puerta 'to lean against the door'
19. Apoyarse el uno en el otro 'to lean on each other'
20. Llevarse la caja 'to take the box with him'
21. Pudrirse una uva 'to rot (a grape)'
22. Brotar una planta 'to sprout (a plant)'
23. Aparecer una persona 'to appear (someone)'
24. Desaparecer una persona 'to disappear (someone)'
25. Gastarse, acabarse la tinta 'to be none left (ink)'
26. Faltar un vaso 'to lack (a glass)'
27. Multiplicarse 'to multiply'
28. Faltar sal 'to lack (salt)'
29. Llenarse la estantería 'to get filled (the bookcase)'
30. Vacarse la estantería 'to get empty (the bookcase)'
31. Apagarse el ordenador 'to turn off (the computer)'
32. Encenderse el ordenador 'to turn on (the computer)'
33. Abrirse la puerta 'to open (the door)'
34. Cerrarse la puerta 'to close (the door)'
35. Quemarse, arder (un papel) 'to burn (a paper)'
36. Tenerse en pie, caerse 'to stay in place, to fell down'

37. Arreglarse la televisión 'to get fixed'
38. Arder, quemarse (piedra, papel) 'to burn (stone, paper)'
39. Moverse la calabaza 'to move (the pumpkin)'
40. Llevarse la silla 'to take the chair away'
41. Caer agua del tejado 'to fell down (water from the roof)'
42. Caer hojas, frutos 'to fell down (fruits, leaves)'
43. Nublarse 'to get cloudy'
44. Secarse la toalla 'to get dried (the towel)'
45. Soltarse, caerse la pinza 'to get released, to fall down (the clothespin)'
46. Mancharse el pantalón 'to get dirty (the trousers)'
47. Limpiarse el pantalón 'to get clean (the trousers)'
48. Enrollarse la tela 'to roll up'
49. Acostarse, meterse en la cama 'to get in bed'
50. Dormir 'to sleep'
51. Levantarse de la cama 'to get up from bed'
52. Estirarse (una persona) 'to stretch'
53. Cogerse, mirarse el pie 'to grab her feet, to look at her feet'
54. Lavarse la cara 'to wash her face'
55. Mirarse un grano 'to look at pimple in her face'
56. Mirarse en el espejo 'to look into the mirror'
57. Reflejarse 'to be reflected in the water'
58. Tirarse del pelo 'to pull her hair'
59. Cortarse las uñas 'to cut her nails'
60. Calentarse (en la chimenea) 'to get warm (at the fireplace)'
61. Taparse 'to cover with a blanket'
62. Calentarse las manos (en la chimenea) 'to warm her hands (at the fireplace)'
63. Llegar trayendo algo 'to arrive bringing something'
64. Creerse guapa 'to think that she's pretty'
65. Colocarse el sombrero 'to put her hat on'
66. Ponerse las gafas 'to put her glasses on'
67. Quitarse las gafas 'to take her glasses off'
68. Taparse los ojos 'to cover her eyes'
69. Quitarse la chaqueta 'to take her jacket off'
70. Ponerse el gorro 'to put his cap on'
71. Meterse la mano en el bolsillo 'to put her hand into her pocket'
72. Meter la cabeza en el cubo 'to put his head into the bucket'
73. Dormirse en clase 'to fall asleep in class'
74. Concentrarse, despistarse, distraerse 'to get focussed, to lose focus, to get distracted'
75. Apuntarse algo 'to note something down'
76. Olvidarse las llaves 'to forget the keys'
77. Enterarse de algo 'to find out about something'
78. Confesarse 'to confess'
79. Aburrirse 'to get bored'
80. Salirse del dibujo 'to colour out of the line'
81. Pensar 'to think'
82. Equivocarse 'to be wrong, to make a mistake'
83. Dudar 'to doubt'

84. Decidirse 'to decide'
85. Pensárselo 'to think over something'
86. Aprenderse, estudiarse algo 'to learn something, to study something'
87. Merecerse, buscárselo 'to deserve, to look for something'
88. Subirse (a la silla) 'to climb up the chair'
89. Bajarse (de la silla) 'to climb down the chair'
90. Escondarse 'to hide'
91. Asomarse 'to show her head'
92. Sentarse 'to sit down'
93. Levantarse de la silla 'to get up from the chair'
94. Encerrarse 'to lock herself up'
95. Escondarse 'to hide'
96. Asomar la cabeza 'to show her face'
97. Colgarse de la reja 'to hang from the grille'
98. Arrastrarse 'to crawl'
99. Tirarse a la piscina 'to jump into the pool'
100. Meterse en el armario 'to get in the closet'
101. Salir del armario 'to get out of the closet'
102. Aprenderse, arrimarse 'to get closer to each other'
103. Escaparse de la cárcel 'to escape prison'
104. Huir el ladrón 'to flee (the thief)'
105. Escapársele el perro 'to get loose (the dog from someone)'
106. Ponerse de rodillas 'to kneel down'
107. Tumbarse 'to lay down'
108. Acurrucarse, encogerse 'to curl up'
109. Correr 'to run'
110. Acercarse una a otra 'to get closer to each other'
111. Alejarse, separarse una de otra 'to get further from each other'
112. Separarse de la mesa 'to get further from the table'
113. Colocarse en la foto 'to get in the picture'
114. Irse, quitarse de la foto 'to get out of the picture'
115. Darse la vuelta 'to turn around'
116. Bailar 'to dance'
117. Cantar 'to sing'
118. Comer una galleta 'to eat a cookie'
119. Beber agua 'to drink water'
120. Leer y fumar 'to read and to smoke'
121. Pegarse a la pared (con pegamento) 'to glue himself to the wall'
122. Taparse la cara 'to cover her face'
123. Quemarse el dedo 'to burn her finger'
124. Pegarse los dedos (con pegamento) 'to glue his fingers to each other'
125. Mojarse los pies (con las manos) 'to wet her feet (with her hands)'
126. Mojarse (con las manos) 'to get wet (with her hands)'
127. Mojarse (con la manguera) 'to get wet inadvertently (with the hose)'
128. Meter el pie en el agua 'to put her foot in the water'
129. Bañarse (nadar) 'to take a bath (swim)'
130. Ahogarse 'to drown'

131. Secarse 'to get dry'
132. Olerse las manos 'to smell her hands'
133. Darse un golpe (a propósito) 'to hit herself'
134. Arrancarse un pelo 'to pull a hair out'
135. Olerse 'to smell herself'
136. Morderse las uñas 'to bite her nails'
137. Pintarse el brazo 'to paint her arm'
138. Rascarse la cabeza 'to scratch her head'
139. Secarse el pie 'to dry her foot'
140. Rascarse 'to scratch'
141. Hervir el agua 'to boil (the water)'
142. Mezclarse el aceite y el agua 'to get mixed (oil and water)'
143. Salirse el agua, desbordarse el vaso 'to overflow (the glass)'
144. Derramarse el vaso 'to spill (the glass)'
145. Flotar, hundirse 'to float, to sink'
146. Llenarse el vaso 'to fill up (the glass)'
147. Congelarse 'to freeze'
148. Acabarse el yogur 'to finish the yoghourt'
149. Comerse una galleta 'to eat up the cookie'
150. Prepararse un sándwich 'to prepare a sandwich'
151. Beberse un vaso de agua 'to drink up a glass of water'
152. Fumarse un cigarrillo 'to smoke a cigarette'
153. Apuntarse a una competición 'to sign up for a competition'
154. Leerse un libro 'to read a book'
155. Bailar 'to dance'
156. Leer 'to read'
157. Negarse a dar algo 'to refuse to give something'
158. Callarse 'to get quiet'
159. Discutir 'to discuss'
160. Negarse a comer algo 'to refuse to eat something'
161. Presentarse, saludar 'to introduce themselves to each other, to say hello'
162. Despedirse 'to say goodbye'
163. Saludarse a una mismo 'to say hello to herself'
164. Saludarse unos a otros 'to say hello to each other'
165. Medirse una con otra 'to compare each other by measuring themselves'
166. Tocarse una a otra 'to touch each other'
167. Olerse una a otra 'to smell each other'
168. Darse la mano una a otra 'to shake each other's hand'
169. Quitarse la chaqueta una a otra 'to take the jacket off each other'
170. Intercambiarse, darse un cenicero 'to interchange an ashtray, to give an ashtray to each other'
171. Recibir un regalo 'to receive a present'
172. Morderse el dedo una a otra 'to bite each other's finger'
173. Pelearse, luchar 'to fight'
174. Chocarse 'to collide with each other'
175. Darse un golpe una a otra 'to hit each other'
176. Escribirse una a otra 'to write to one another'

177. Secarse una a otra 'to dry each other'
178. Abrazarse 'to hug each other'
179. Morderse una a otra 'to bite each other'
180. Estrangularse una a otra 'to strangle each other'
181. Perseguirse una a otra, correr una tras otra 'to chase one another, to run after each other'
182. Quejarse 'to complain, to moan'
183. Llorar, calmarse 'to cry, to calm down'
184. Sonreír 'to smile'
185. Sonreírse una a otra 'to smile to each other'
186. Reírse 'to laugh'
187. Cansarse 'to get tired'
188. Descansar 'to rest'
189. Sudar 'to sweat'
190. Oír 'to hear'
191. Oler un tomate 'to smell a tomato'
192. Alegrarse 'to get happy'
193. Esperar 'to wait'
194. Asustarse 'to get scared'
195. Temblar 'to tremble'
196. Engordar 'to put on weight'
197. Curarse 'to get cured'
198. Bostezar 'to yawn'
199. Enfriarse 'to get cold'
200. Estornudar 'to sneeze'
201. Toser 'to cough'
202. Encontrarse algo en el suelo 'to find something on the floor'
203. Encontrar algo que se está buscando 'to find something that you are looking for'
204. Iluminarse la casa, encenderse la luz 'to light up (a house)'
205. Imaginarse (Dibujos animados) 'to imagine' (Cartoons)
206. Soñar (Dibujos animados) 'to dream' (Cartoons)
207. Crecer (Dibujos animados) 'to grow up' (Cartoons)
208. Morirse (Dibujos animados) 'to die' (Cartoons)
209. Amanecer (Dibujos animados) 'to dawn' (Cartoons)
210. Anochecer (Dibujos animados) 'to get dark' (Cartoons)
211. Florecer (Dibujos animados) 'to bloom' (Cartoons)
212. Perderse (Dibujos animados) 'to get lost' (Cartoons)
213. Latir (Dibujos animados) 'to beat (a heart)' (Cartoons)
214. Divertirse, pasárselo bien (Dibujos animados) 'to get amused, to have fun' (Cartoons)

APPENDIX 3

Localities interviewed with the questionnaire

- QT001: Antequera (Málaga). Female. Spanish.
- QT002: Las Pinedas (La Carlota, Córdoba). Male. Spanish.
- QT003: Zumaia (Guipúzcoa). Male. Spanish.
- QT004: San Sebastián (Guipúzcoa). Female. Spanish.
- QT006: Candolías (Vega de Pas, Cantabria). Male. Spanish.
- QT007: La Revilla (San Vicente de la Barquera, Cantabria). Female. Spanish.
- QT008: Vibaño (Llanes, Asturias). Female. Spanish (sometimes Asturian).
- QT009: La Borbolla (Asturias). Female. Spanish.
- QT010: Folgueiras de Boiro (Ibias de San Antolín, Asturias). 3 females and 4 males. Galician.
- QT011: El Condado (Laviana, Asturias). Male. Spanish.
- QT012: Santa Olalla del Cala (Huelva). Male. Spanish.
- QT013: Bonares (Huelva). Female. Spanish.
- QT014: Navahermosa (Toledo). Female. Spanish.
- QT015: Jimena de la Frontera (Cádiz). Male. Spanish.
- QT016: Albires (Izagre, León). Male. Spanish.
- QT017: Turcia (León). Female. Spanish.
- QT018: Quintanilla de Rueda (Cubillas de Rueda, León). 1 female and 1 male. Spanish.
- QT019: Cádavos (A Mezquita, Orense). Male. Spanish.
- QT020: Astariz (Castrelo de Miño, Orense). 2 females. Spanish / Galician.
- QT021: Laxoso (Ponte Caldelas, Pontevedra). Female. Galician.
- QT022: Santiago (La Coruña). Male. Galician.
- QT023: Narón (La Coruña). Male. Galician.
- QT024: O Corgo (Lugo). Male. Galician.
- QT025: Teixeira (Paradavella, A Fonsagrada, Lugo). Female. Galician.
- QT026: Villanueva de Omaña (Murias de Paredes, León). Female. Spanish.
- QT027: Gumiel de Mercado (Burgos). Female. Spanish.
- QT028: Suances (Cantabria). Males. Spanish.
- QT030: Labastida (Álava). Female. Spanish.
- QT032: Épila (Zaragoza). 1 female and 1 male. Spanish.
- QT033: San Esteban de Gormaz (Soria). Male. Spanish.
- QT034: Revenga (Segovia). Male. Spanish.
- QT036: Peleas de Arriba (Corrales, Zamora). Female. Spanish.
- QT037: Horcajo Medianero (Salamanca). 2 females. Spanish.
- QT038: Navalморal de la Sierra (Ávila). Female. Spanish.
- QT039: Vélez-Blanco (Almería). Female. Spanish.
- QT040: Ventorros de San José (Loja, Granada). Female. Spanish.
- QT041: Alcolea del Pinar (Guadalajara). Female. Spanish.
- QT042: Quesa (Valencia). Female. Spanish.

- QT043: Pinoso (Alicante). Female. Spanish.
QT044: Arlós / Rodiles (Corvera, Asturias). Female. Spanish (sometimes Asturian).
QT045: Colinas de Abajo (Tineo, Asturias). Male. Spanish (sometimes Asturian).
QT046: Chanuces / Llanuces (Quirós, Asturias). Female. Spanish (sometimes Asturian).
QT047: Santaella (Córdoba). Male. Spanish.

APPENDIX 4

List of analysed verbs and their classification

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
abajar	yes ($N = 2$)	no	no	no	no	no	no	no
abaratar	yes ($N = 1$)	no	no	no	no	no	no	no
abastecer	yes ($N = 1$)	no	no	no	no	no	no	no
ablandar	yes ($N = 1$)	no	no	no	no	no	no	no
ablentar	yes ($N = 1$)	no	no	no	no	no	no	no
abollar	yes ($N = 1$)	no	no	no	no	no	no	no
abrasar	yes ($N = 3$)	no	no	no	no	no	no	no
abrir	yes ($N = 37$)	no	no	no	no	no	no	no
abrochar	yes ($N = 1$)	no	no	no	no	no	no	no
abrujar	yes ($N = 1$)	no	no	no	no	no	no	no
aburrir	no	no	no	yes ($N = 18$)	no	no	no	no
abusar	no	no	no	no	no	yes ($N = 1$)	no	no
acabar	yes ($N = 140$)	no	no	no	no	no	no	no
acarrear	no	no	no	no	yes ($N = 1$)	no	no	no
accidentalr	yes ($N = 1$)	no	no	no	no	no	no	no
acedar	yes ($N = 2$)	no	no	no	no	no	no	no
acercar	yes ($N = 30$)	no	no	no	no	no	no	no
acertar	no	no	no	no	yes ($N = 2$)	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible Transitive
achicar	no	no	no	no	no	no	yes (N = 2)
achichorrar	yes (N = 1)	no	no	no	no	no	no
aclamar	yes (N = 2)	no	yes (N = 1)	no	no	no	no
acobardar	no	no	no	yes (N = 2)	no	no	no
acoplar	yes (N = 1)	no	no	no	no	no	no
acordar	no	no	no	no	yes (N = 67)	no	yes (N = 484)
acortar	yes (N = 1)	no	no	no	no	no	no
acostar	yes (N = 102)	no	no	no	no	no	no
acostumbrar	yes (N = 17)	no	no	no	no	no	no
acudir	no	no	no	no	no	yes (N = 1)	no
acurrucar	no	no	no	no	no	no	yes (N = 1)
adaptar	yes (N = 13)	no	no	no	no	no	no
adelantar	yes (N = 1)	no	no	no	no	no	no
adiestrar	yes (N = 1)	no	no	no	no	no	no
administrar	no	no	yes (N = 3)	no	no	no	no
admirar	no	no	no	yes (N = 2)	no	no	no
adobar	yes (N = 3)	no	no	no	no	no	no
adornecer	yes (N = 1)	no	no	no	no	no	no
afiliar	yes (N = 1)	no	no	no	no	no	no
afincar	yes (N = 1)	no	no	no	no	no	no
aflojar	yes (N = 1)	no	no	no	no	no	no
agachar	yes (N = 20)	no	no	no	no	no	no
agarrar	no	no	no	yes (N = 29)	no	no	no
agotar	yes (N = 2)	no	no	no	no	no	no
agrarar	no	no	no	no	no	no	yes (N = 1)
agriar	yes (N = 1)	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
agrietar	yes (N = 1)	no	no	no	no	no	no	no
aguantar	yes (N = 1)	no	yes (N = 22)	no	no	no	no	yes (N = 1)
aguardar	no	yes (N = 2)	no	no	no	no	no	yes (N = 1)
agusanar	no	no	no	no	no	no	yes (N = 2)	no
ahogar	yes (N = 20)	no	no	no	no	no	no	no
ahorcar	yes (N = 3)	no	no	no	no	no	no	no
ahuecar	yes (N = 6)	no	no	no	no	no	no	no
ahumar	yes (N = 1)	no	no	no	no	no	no	no
airear	yes (N = 3)	no	no	no	no	no	no	no
ajuntar	yes (N = 22)	no	no	no	no	no	no	no
ajustar	yes (N = 4)	no	no	no	no	no	no	no
alargar	yes (N = 1)	no	no	no	no	no	no	no
alegrar	no	no	no	yes (N = 13)	no	no	no	no
alejar	yes (N = 1)	no	no	no	no	no	no	no
alimentar	yes (N = 8)	no	no	no	no	no	no	no
alistar	yes (N = 1)	no	no	no	no	no	no	no
allanar	yes (N = 1)	no	no	no	no	no	no	no
almacenar	yes (N = 1)	no	no	no	no	no	no	no
almorzar	no	no	no	no	no	no	no	yes (N = 1)
alumbrar	yes (N = 2)	no	no	no	no	no	no	no
alunar	no	no	no	no	no	no	yes (N = 3)	no
amedrentar	yes (N = 1)	no	no	no	no	no	no	no
amolar	yes (N = 2)	no	no	no	no	no	no	no
amoldar	yes (N = 5)	no	no	no	no	no	no	no
amontar	yes (N = 1)	no	no	no	no	no	no	no
amontonar	yes (N = 1)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
andar	no	no	no	no	no	yes (N = 1)	no	no
animar	no	no	no	yes (N = 6)	no	no	no	no
anotar	yes (N = 1)	no	no	no	no	no	no	no
anticipar	yes (N = 1)	no	no	no	no	no	no	no
antojar	no	no	no	no	no	no	yes (N = 6)	no
anunciar	yes (N = 1)	no	no	no	no	no	no	no
añejar	yes (N = 3)	no	no	no	no	no	no	no
apagar	yes (N = 9)	no	no	no	no	no	no	no
apañar	no	no	yes (N = 39)	no	no	no	no	yes (N = 3)
apartar	yes (N = 11)	no	no	no	no	no	no	no
apedrear	yes (N = 1)	no	no	no	no	no	no	no
apegar	yes (N = 1)	no	no	no	no	no	yes (N = 1)	no
apelldar	yes (N = 1)	no	no	no	no	no	no	no
apelmazar	yes (N = 3)	no	no	no	no	no	no	no
apercibir	no	no	yes (N = 1)	no	no	no	no	no
aplstar	yes (N = 8)	no	no	no	no	no	no	no
aplicar	no	no	yes (N = 5)	no	no	no	no	no
apoderar	yes (N = 4)	no	no	no	no	no	no	no
apolillar	yes (N = 2)	no	no	no	no	no	no	no
aposentar	yes (N = 1)	no	no	no	no	no	no	no
aprender	no	no	no	no	no	no	no	yes (N = 21)
aprensar	yes (N = 3)	no	no	no	no	no	no	no
apretar	yes (N = 5)	no	no	no	no	no	no	no
aprobar	yes (N = 1)	no	no	no	no	no	no	no
aprovechar	no	no	no	no	yes (N = 13)	no	no	no
aproximar	yes (N = 2)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
apudrir	yes ($N = 1$)	no	no	no	no	no	no	no
apuntar	yes ($N = 20$)	yes ($N = 1$)	no	no	no	no	no	no
apurar	no	no	no	yes ($N = 5$)	no	no	no	no
arder	no	no	no	no	no	yes ($N = 3$)	no	no
arrancar	yes ($N = 1$)	no	no	no	no	no	no	no
arranciar	yes ($N = 5$)	no	no	no	no	no	no	no
arrastrar	yes ($N = 3$)	no	no	no	no	no	no	no
arrebatar	yes ($N = 3$)	no	no	no	no	no	no	no
arrecoger	yes ($N = 1$)	no	no	no	no	no	no	no
arreglar	yes ($N = 9$)	no	yes ($N = 44$)	no	no	no	no	yes ($N = 12$)
arrepentir	no	no	no	no	no	no	yes ($N = 5$)	no
arrestar	no	no	yes ($N = 2$)	no	no	no	no	no
arreunir	yes ($N = 1$)	no	no	no	no	no	no	no
arriesgar	no	no	yes ($N = 1$)	no	no	no	no	no
arrimar	yes ($N = 35$)	no	no	no	no	no	no	no
arrodear	yes ($N = 1$)	no	no	yes ($N = 1$)	no	no	no	no
arrodillar	yes ($N = 15$)	no	no	no	no	no	no	no
arronper	yes ($N = 1$)	no	no	no	no	no	no	no
arropar	yes ($N = 1$)	no	no	no	no	no	no	no
arruinar	yes ($N = 2$)	no	no	no	no	no	no	no
asalar	yes ($N = 1$)	no	no	no	no	no	no	no
asar	yes ($N = 3$)	no	no	no	no	no	no	no
asentar	yes ($N = 9$)	no	no	no	no	no	no	no
asfixiar	yes ($N = 4$)	no	no	no	no	no	no	no
asimilar	no	no	yes ($N = 1$)	no	no	no	no	no
asolear	yes ($N = 2$)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
asomar	no	no	yes (N = 54)	no	no	no	no	no
asombrar	no	no	no	yes (N = 1)	no	no	no	no
asustar	no	no	no	yes (N = 29)	no	no	no	no
atar	yes (N = 2)	no	no	no	no	no	no	no
atascar	yes (N = 1)	no	no	no	no	no	no	no
atontar	yes (N = 1)	no	no	no	no	no	no	no
atrancar	yes (N = 2)	no	no	no	no	no	no	no
atrasar	yes (N = 2)	no	no	no	no	no	no	no
atravesar	no	no	no	no	yes (N = 2)	no	no	no
atrever	no	no	no	no	no	no	yes (N = 29)	no
aturullar	yes (N = 1)	no	no	no	no	no	no	no
aumentar	yes (N = 3)	no	no	no	no	no	no	no
aunir	yes (N = 1)	no	no	no	no	no	no	no
avanzar	yes (N = 3)	no	no	no	no	no	no	no
avénir	yes (N = 5)	no	no	no	no	no	no	no
averiar	yes (N = 1)	no	no	no	no	no	no	no
avinagrar	yes (N = 2)	no	no	no	no	no	no	no
bajar	yes (N = 312)	no	no	no	no	yes (N = 62)	no	yes (N = 2)
beber	no	no	no	no	no	no	no	yes (N = 19)
beneficiar	no	no	no	yes (N = 1)	no	no	no	no
blanquear	yes (N = 1)	no	no	no	no	no	no	no
blanquear	yes (N = 7)	no	no	no	no	no	no	no
borrar	yes (N = 5)	no	no	no	no	no	no	no
brincar	no	no	no	no	no	yes (N = 3)	no	no
burlar	no	no	no	no	no	no	yes (N = 2)	no
buscar	yes (N = 2)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
cabrear	no	no	no	yes (N = 4)	no	no	no	no
cachondear	no	no	no	no	no	no	yes (N = 1)	no
caer	no	no	no	no	no	yes (N = 150)	no	no
cagar	no	no	no	no	no	yes (N = 16)	no	no
calar	yes (N = 4)	no	no	no	no	no	no	no
calcular	no	no	no	no	no	no	no	yes (N = 1)
caldear	yes (N = 6)	no	no	no	no	no	no	no
calentar	yes (N = 45)	no	no	no	no	no	no	no
callar	no	no	yes (N = 113)	no	no	no	no	yes (N = 1)
calzar	no	no	no	no	no	no	no	yes (N = 1)
cambiar	yes (N = 134)	no	no	no	yes (N = 11)	no	no	no
canecer	no	no	no	no	no	yes (N = 2)	no	no
cansar	no	no	no	yes (N = 64)	no	no	no	no
cargar	yes (N = 3)	no	no	no	yes (N = 3)	no	no	yes (N = 6)
carrear	no	no	no	no	no	no	no	yes (N = 1)
casar	yes (N = 647)	no	no	no	no	no	no	no
cascar	yes (N = 2)	no	no	no	no	no	no	no
cebar	no	no	no	no	no	no	yes (N = 1)	no
cenar	no	no	no	no	no	yes (N = 1)	no	no
centrar	no	no	yes (N = 1)	no	no	no	no	no
cerrar	yes (N = 2)	no	no	no	no	no	no	no
changar	yes (N = 2)	no	no	no	no	no	no	no
chivar	no	no	no	no	no	no	no	yes (N = 1)
chupar	yes (N = 1)	no	no	no	no	no	no	yes (N = 1)
clamar	no	no	yes (N = 1)	no	no	no	no	no
clarear	yes (N = 4)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
clavar	yes (N = 4)	no	no	no	no	no	no	no
coagular	yes (N = 16)	no	no	no	no	no	no	no
coagulizar	yes (N = 1)	no	no	no	no	no	no	no
cobijar	yes (N = 1)	no	no	no	no	no	no	no
cobrar	no	no	no	no	no	no	no	yes (N = 1)
cocer	yes (N = 98)	no	no	no	no	no	no	no
cocinar	yes (N = 1)	no	no	no	no	no	no	no
coger	yes (N = 1)	yes (N = 6)	no	no	yes (N = 8)	no	no	yes (N = 5)
colar	yes (N = 11)	no	no	no	no	no	no	no
colegiar	yes (N = 1)	no	no	no	no	no	no	no
colgar	yes (N = 4)	no	no	no	no	no	no	no
colocar	yes (N = 31)	no	no	no	no	no	no	no
combinar	yes (N = 1)	no	no	no	no	no	no	no
comer	no	no	no	no	no	no	no	yes (N = 132)
complicar	yes (N = 2)	no	no	no	no	no	no	no
componer	yes (N = 2)	no	no	no	no	no	no	no
comportar	no	no	no	no	no	no	yes (N = 5)	no
comunicar	yes (N = 6)	no	no	no	no	no	no	no
concentrar	yes (N = 3)	no	no	no	no	no	no	no
concienciar	yes (N = 1)	no	no	no	no	no	no	no
concretar	no	no	yes (N = 1)	no	no	no	no	no
confesar	no	no	yes (N = 45)	no	no	no	no	no
confiar	yes (N = 1)	no	no	no	no	no	no	no
conformar	no	no	no	no	no	yes (N = 1)	yes (N = 10)	no
confrontar	yes (N = 3)	no	no	no	no	no	no	no
confundir	no	no	no	no	yes (N = 3)	no	no	yes (N = 1)

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
congelar	yes (N = 3)	no	no	no	no	no	no	no
conocer	no	no	no	no	no	no	no	yes (N = 1)
conservar	yes (N = 68)	no	no	no	no	no	no	no
constipar	no	no	no	no	no	no	yes (N = 1)	no
consumir	yes (N = 6)	no	no	no	no	no	no	no
contaminar	yes (N = 1)	no	no	no	no	no	no	no
contemplar	yes (N = 1)	no	no	no	no	no	no	no
contener	yes (N = 1)	no	no	no	yes (N = 1)	no	no	no
contentar	no	no	no	yes (N = 1)	no	no	no	no
convertir	yes (N = 5)	no	no	no	no	no	no	no
correr	yes (N = 6)	no	no	no	no	yes (N = 2)	no	yes (N = 1)
corromper	yes (N = 3)	no	no	no	no	no	no	no
cortar	yes (N = 30)	no	no	no	no	no	no	no
crecer	no	no	no	no	no	yes (N = 1)	no	no
crear	no	no	no	no	no	no	no	yes (N = 187)
criar	yes (N = 162)	no	no	no	no	no	no	no
cruzar	yes (N = 1)	no	no	no	no	no	no	no
cuajar	yes (N = 181)	no	no	no	no	no	no	yes (N = 1)
cubrir	yes (N = 1)	no	no	no	no	no	no	no
cuidar	yes (N = 4)	no	no	no	yes (N = 1)	no	no	yes (N = 1)
curar	yes (N = 108)	no	no	no	no	no	no	no
dar	yes (N = 37)	no	no	no	no	no	no	yes (N = 11)
dar_cuenta	no	no	no	no	no	no	no	yes (N = 109)
dar_las_vueltas	no	no	no	no	no	no	no	yes (N = 1)
dar_una_vuelta	no	no	no	no	no	no	no	yes (N = 5)
decidir	no	no	no	no	yes (N = 4)	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
decir	no	no	no	no	no	no	no	yes (N = 1)
declarar	no	no	yes (N = 7)	no	no	no	no	no
dedicar	no	no	yes (N = 108)	no	no	no	no	yes (N = 1)
defender	yes (N = 17)	no	no	no	no	no	no	no
deitar	yes (N = 1)	no	no	no	no	no	no	no
dejar	no	yes (N = 1)	no	no	yes (N = 12)	yes (N = 1)	no	yes (N = 44)
denominar	yes (N = 1)	no	no	no	no	no	no	no
derramar	yes (N = 1)	no	no	no	no	no	no	no
derretir	yes (N = 9)	no	no	no	no	no	no	no
derribar	yes (N = 1)	no	no	no	no	no	no	no
derrochar	no	no	no	no	no	no	no	yes (N = 1)
desahogar	no	no	yes (N = 1)	no	no	no	no	no
desalar	yes (N = 4)	no	no	no	no	no	no	no
desangrar	yes (N = 4)	no	no	no	no	no	no	no
desaparecer	yes (N = 45)	no	no	no	no	no	no	no
despartar	yes (N = 3)	no	no	no	no	no	no	no
desarmar	yes (N = 1)	no	no	no	no	no	no	no
desarrollar	yes (N = 8)	no	no	no	no	no	no	no
desbaratar	yes (N = 3)	no	no	no	no	no	no	no
desbocar	no	no	no	no	no	no	yes (N = 1)	no
descargar	no	no	no	no	yes (N = 1)	no	no	no
descender	no	no	no	no	no	yes (N = 1)	no	no
descompensar	yes (N = 1)	no	no	no	no	no	no	no
descomponer	no	no	yes (N = 1)	no	no	no	no	no
desconectar	yes (N = 1)	no	no	no	no	no	no	no
descongelar	yes (N = 2)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
descuidar	no	no	yes ($N = 22$)	no	no	no	no	no
desear	no	no	no	no	yes ($N = 1$)	no	no	no
desemburruñar	yes ($N = 2$)	no	no	no	no	no	no	no
desenganchar	yes ($N = 1$)	no	no	no	no	no	no	no
desenvolver	no	no	no	no	no	yes ($N = 1$)	no	no
desertar	no	no	no	no	yes ($N = 1$)	no	no	no
desesperar	no	no	yes ($N = 2$)	no	no	no	no	no
desgañar	no	no	no	no	yes ($N = 2$)	no	no	no
desgranar	yes ($N = 2$)	no	no	no	no	no	no	no
deshacer	yes ($N = 25$)	no	no	no	yes ($N = 3$)	no	no	no
deshidratar	yes ($N = 1$)	no	no	no	no	no	no	no
desjuntar	yes ($N = 1$)	no	no	no	no	no	no	no
desliar	no	no	no	no	yes ($N = 1$)	no	no	no
deslizar	yes ($N = 1$)	no	no	no	no	no	no	no
desmoralizar	yes ($N = 1$)	no	no	no	no	no	no	no
desobedecer	no	no	no	no	no	no	yes ($N = 1$)	no
despedir	yes ($N = 10$)	no	no	no	no	no	no	yes ($N = 1$)
despegar	yes ($N = 3$)	no	no	no	no	no	no	no
despellejar	yes ($N = 1$)	no	no	no	no	no	no	no
despertar	yes ($N = 16$)	no	no	no	no	no	no	no
despistar	no	no	no	yes ($N = 1$)	no	no	no	no
desplazar	yes ($N = 10$)	no	no	no	no	no	no	no
desplomar	yes ($N = 2$)	no	no	no	no	no	no	no
desprender	yes ($N = 2$)	no	no	yes ($N = 1$)	no	no	no	no
destacar	yes ($N = 1$)	no	no	no	no	no	no	no
destilar	yes ($N = 3$)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
destrozar	yes (N = 1)	no	no	no	no	no	no	no
detener	yes (N = 1)	no	no	no	no	no	no	no
devorar	no	no	no	no	no	no	no	yes (N = 1)
diferenciar	yes (N = 1)	no	no	no	no	no	no	no
diluir	yes (N = 1)	no	no	no	no	no	no	no
dirigir	no	no	yes (N = 4)	no	no	no	no	no
disculpar	yes (N = 1)	no	no	no	no	no	no	no
discurrir	no	no	no	no	no	no	no	yes (N = 1)
discutir	no	no	no	no	no	yes (N = 2)	no	no
disfrutar	no	no	no	no	yes (N = 6)	no	no	no
disgustar	no	no	no	yes (N = 1)	no	no	no	no
disparar	yes (N = 1)	no	no	no	no	no	no	no
dispersar	yes (N = 1)	no	no	no	no	no	no	no
distinguir	yes (N = 1)	no	no	no	no	no	no	no
distraer	no	no	no	yes (N = 10)	no	no	no	no
distribuir	yes (N = 1)	no	no	no	no	no	no	no
disuadir	yes (N = 1)	no	no	no	no	no	no	no
divertir	no	no	no	yes (N = 59)	no	no	no	yes (N = 1)
divorciar	yes (N = 11)	no	no	no	no	no	no	no
doblar	yes (N = 4)	no	no	no	no	no	no	no
dorar	yes (N = 1)	no	no	no	no	no	no	no
dormir	no	no	no	no	no	yes (N = 17)	no	no
drogar	yes (N = 3)	no	no	no	no	no	no	no
echar	yes (N = 30)	no	no	no	no	no	no	yes (N = 37)
echar a perder	yes (N = 36)	no	no	no	no	no	no	no
echar cuenta	no	no	no	no	no	no	no	yes (N = 2)

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
educar	yes (N = 1)	no	no	no	no	no	no	no
embadurnar	yes (N = 1)	no	no	no	no	no	no	no
emborrachar	yes (N = 11)	no	no	no	no	no	no	no
emocionar	no	no	no	yes (N = 2)	no	no	no	no
empadronar	yes (N = 1)	no	no	no	no	no	no	no
empanar	yes (N = 1)	no	no	no	no	no	no	no
empeñar	no	no	yes (N = 8)	no	no	no	no	no
empezar	yes (N = 28)	no	no	no	yes (N = 3)	no	no	no
empinar	yes (N = 6)	no	no	no	no	no	no	no
emplear	yes (N = 1)	no	yes (N = 2)	no	no	no	no	no
empollar	yes (N = 1)	no	no	no	no	no	no	no
emporcar	yes (N = 2)	no	no	no	no	no	no	no
emprefiar	yes (N = 8)	no	no	no	no	no	no	no
enamorar	yes (N = 5)	no	no	yes (N = 16)	no	no	no	no
encalar	yes (N = 2)	no	no	no	no	no	no	no
encanecer	no	no	no	no	no	yes (N = 2)	no	no
encantar	yes (N = 1)	no	no	no	no	no	no	no
encaprichar	no	no	no	no	no	no	yes (N = 2)	no
encarecer	yes (N = 1)	no	no	no	no	no	no	no
encargar	no	no	no	no	yes (N = 49)	no	no	no
encender	yes (N = 4)	no	no	no	no	no	no	no
encerrar	yes (N = 1)	no	no	no	no	no	no	yes (N = 1)
enchargar	yes (N = 1)	no	no	no	no	no	no	no
encher	yes (N = 3)	no	no	no	no	no	no	no
enchufar	yes (N = 4)	no	no	no	no	no	no	no
encoger	yes (N = 2)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
enconar	yes (N = 1)	no	no	no	no	no	no	no
encontrar	yes (N = 20)	no	no	no	yes (N = 5)	no	no	yes (N = 28)
encorvar	yes (N = 2)	no	no	no	no	no	no	no
endemoniar	no	no	no	yes (N = 1)	no	no	no	no
endurecer	yes (N = 5)	no	no	no	no	no	no	no
enemistar	yes (N = 1)	no	no	no	no	no	no	no
enfadar	yes (N = 3)	no	no	yes (N = 46)	no	no	no	no
enfermar	yes (N = 5)	no	no	no	no	no	no	no
enfriar	yes (N = 71)	no	no	no	no	no	no	no
enganchar	yes (N = 10)	no	no	no	no	no	no	no
engañar	yes (N = 4)	no	no	no	no	no	no	yes (N = 1)
engarañar	yes (N = 1)	no	no	no	no	no	no	no
engarranchar	no	no	no	no	no	no	yes (N = 2)	no
englobar	yes (N = 1)	no	no	no	no	no	no	no
engordar	yes (N = 22)	no	no	no	no	no	no	no
engrandecer	yes (N = 1)	no	no	no	no	no	no	no
enjugar	yes (N = 4)	no	no	no	no	no	no	no
enllenar	yes (N = 1)	no	no	no	no	no	no	no
enmendar	yes (N = 3)	no	no	no	no	no	no	no
enojar	yes (N = 1)	no	no	no	no	no	no	no
enorgullecer	yes (N = 1)	no	no	no	no	no	no	no
enranciar	yes (N = 6)	no	no	no	no	no	no	no
enredar	yes (N = 2)	no	no	no	no	no	no	no
enrollar	yes (N = 1)	no	no	no	no	no	yes (N = 2)	no
enroscar	yes (N = 1)	no	no	no	no	no	no	no
ensayar	yes (N = 1)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
enseñar	yes (<i>N</i> = 17)	no	no	no	no	no	no	no
ensuciar	yes (<i>N</i> = 5)	no	no	no	no	no	no	no
entender	no	no	no	yes (<i>N</i> = 13)	no	no	no	yes (<i>N</i> = 2)
enterar	yes (<i>N</i> = 105)	no	no	no	no	no	no	no
enterrar	yes (<i>N</i> = 1)	no	no	no	no	no	no	no
entrar	no	no	no	no	no	yes (<i>N</i> = 15)	no	no
entregar	yes (<i>N</i> = 6)	no	no	no	no	no	no	no
entretener	no	no	no	yes (<i>N</i> = 31)	no	no	no	no
entumir	no	no	no	no	no	no	yes (<i>N</i> = 5)	no
envejecer	yes (<i>N</i> = 1)	no	no	no	no	no	no	no
envolver	yes (<i>N</i> = 6)	no	no	no	no	no	no	no
equipar	no	no	no	no	no	no	no	yes (<i>N</i> = 1)
equivocar	yes (<i>N</i> = 30)	no	no	no	no	no	no	no
esbarar	no	no	no	no	no	yes (<i>N</i> = 1)	no	no
esbaratar	yes (<i>N</i> = 1)	no	no	no	no	no	no	no
escabultar	no	no	no	no	no	no	yes (<i>N</i> = 1)	no
escaldar	yes (<i>N</i> = 2)	no	no	no	no	no	no	no
escapar	no	no	no	no	no	yes (<i>N</i> = 59)	no	no
escarriar	yes (<i>N</i> = 1)	no	no	no	no	no	no	no
esconder	yes (<i>N</i> = 36)	no	no	no	no	no	no	no
escribir	no	no	no	no	no	no	no	yes (<i>N</i> = 1)
escullar	no	no	no	no	no	yes (<i>N</i> = 1)	no	no
escurrir	yes (<i>N</i> = 42)	no	no	no	no	no	no	no
esforzar	no	no	no	no	no	no	yes (<i>N</i> = 1)	no
esmenorar	yes (<i>N</i> = 2)	no	no	no	no	no	no	no
esmerar	no	no	no	no	no	no	yes (<i>N</i> = 2)	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
esmoronar	yes (N = 1)	no	no	no	no	no	no	no
espabilar	yes (N = 5)	no	no	no	no	no	no	no
espartar	no	no	no	yes (N = 11)	no	no	no	no
espanzar	no	no	no	no	no	no	yes (N = 1)	no
esparcir	yes (N = 2)	no	no	no	no	no	no	no
espatarrar	no	no	no	no	no	no	yes (N = 1)	no
especializar	yes (N = 1)	no	no	no	no	no	no	no
esperar	no	no	no	no	no	yes (N = 55)	no	yes (N = 1)
esperar	yes (N = 8)	no	no	no	no	no	no	no
espizar	yes (N = 1)	no	no	no	no	no	no	no
esponjar	yes (N = 5)	no	no	no	no	no	no	no
esquebrajar	yes (N = 1)	no	no	no	no	no	no	no
establecer	no	no	yes (N = 2)	no	no	no	no	no
estampar	yes (N = 1)	no	no	no	no	no	no	no
estar	no	no	no	no	no	yes (N = 88)	no	no
estilar	yes (N = 1)	no	no	no	no	no	no	no
estirar	yes (N = 7)	no	no	no	no	no	no	no
estropear	yes (N = 79)	no	no	no	no	no	no	no
estrozar	yes (N = 1)	no	no	no	no	no	no	no
estudiar	no	no	no	no	no	no	no	yes (N = 7)
esvinagrar	yes (N = 1)	no	no	no	no	no	no	no
esvolar	no	no	no	no	no	no	yes (N = 1)	no
eternecer	no	no	no	no	no	no	yes (N = 1)	no
existir	no	no	no	no	no	yes (N = 1)	no	no
expatriar	yes (N = 1)	no	no	no	no	no	no	no
explicar	no	no	yes (N = 22)	no	no	no	no	yes (N = 4)

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
explicotear	no	no	yes (N = 1)	no	no	no	no	no
explotar	yes (N = 9)	no	no	no	no	no	no	no
exponer	yes (N = 5)	no	no	no	no	no	no	no
expresar	no	no	yes (N = 2)	no	no	no	no	no
exprimir	yes (N = 1)	no	no	no	no	no	no	no
extender	yes (N = 5)	no	no	no	no	no	no	no
extrañar	no	no	no	yes (N = 1)	no	no	no	no
fastidiar	yes (N = 6)	no	no	yes (N = 1)	no	no	no	no
fermentar	yes (N = 47)	no	no	no	no	no	no	no
fiar	no	no	no	no	no	no	yes (N = 18)	no
figurar	no	no	no	no	no	no	yes (N = 7)	no
fijar	no	no	yes (N = 282)	no	no	no	no	no
filtrar	yes (N = 2)	no	no	no	no	no	no	no
fisurar	yes (N = 1)	no	no	no	no	no	no	no
fomentar	yes (N = 4)	no	no	no	no	no	no	no
formar	yes (N = 9)	no	no	no	no	no	no	no
forrar	yes (N = 1)	no	no	no	no	no	no	no
fragmentar	yes (N = 1)	no	no	no	no	no	no	no
freir	yes (N = 7)	no	no	no	no	no	no	no
frenar	yes (N = 1)	no	no	no	no	no	no	no
fugar	no	no	no	no	no	no	yes (N = 1)	no
fumar	no	no	no	no	no	no	no	yes (N = 7)
fusionar	yes (N = 2)	no	no	no	no	no	no	no
ganar	no	no	no	no	no	no	no	yes (N = 25)
gastar	yes (N = 7)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
gitar	yes (N = 4)	no	no	no	no	no	no	no
gobernar	no	no	yes (N = 5)	no	no	no	no	no
guardar	yes (N = 1)	no	no	no	no	no	no	no
guasear	no	no	no	no	no	no	yes (N = 2)	no
guiar	yes (N = 3)	no	no	no	no	no	no	no
gustar	no	no	no	no	no	yes (N = 1)	no	no
haber	no	no	no	no	no	no	no	yes (N = 7)
habituár	yes (N = 1)	no	no	no	no	no	no	no
hacer	yes (N = 257)	no	no	no	no	no	yes (N = 19)	yes (N = 28)
hartar	no	no	no	yes (N = 14)	no	no	no	no
helar	yes (N = 19)	no	no	no	no	no	no	no
herniar	no	no	no	no	no	no	yes (N = 1)	no
hervir	yes (N = 106)	no	no	no	no	no	no	no
hinchar	yes (N = 19)	no	no	no	no	no	no	no
hospedar	yes (N = 1)	no	no	no	no	no	no	no
hundir	yes (N = 22)	no	no	no	no	no	no	no
igualar	yes (N = 6)	no	no	no	no	no	no	no
imaginar	no	no	no	no	no	no	no	no
imponer	yes (N = 2)	no	no	no	no	no	no	yes (N = 49)
incendiar	yes (N = 1)	no	no	no	no	no	no	no
incomodar	no	no	no	yes (N = 3)	no	no	no	no
incorporar	yes (N = 6)	no	no	no	no	no	no	no
inculpar	yes (N = 2)	no	no	no	no	no	no	no
infectar	yes (N = 5)	no	no	no	no	no	no	no
inflamar	yes (N = 1)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
inflar	yes ($N = 2$)	no	no	no	no	no	no	no
informar	yes ($N = 5$)	no	no	no	no	no	no	no
inmigrar	no	no	no	no	no	yes ($N = 1$)	no	no
instalar	no	no	yes ($N = 3$)	no	no	no	no	no
intentar	no	no	no	no	no	no	no	yes ($N = 1$)
interesar	no	no	no	yes ($N = 6$)	no	no	no	no
intoxicar	yes ($N = 2$)	no	no	no	no	no	no	no
inventar	no	no	no	no	no	no	no	yes ($N = 2$)
ir	no	no	no	no	no	yes ($N = 880$)	no	no
joder	yes ($N = 12$)	no	no	no	no	no	no	no
jorobar	yes ($N = 1$)	no	no	no	no	no	no	no
jubilar	yes ($N = 56$)	no	no	no	no	no	no	no
jugar	no	no	no	no	no	yes ($N = 1$)	no	yes ($N = 14$)
juntar	yes ($N = 183$)	no	no	no	no	no	no	no
ladear	yes ($N = 1$)	no	no	no	no	no	no	no
lanzar	yes ($N = 2$)	no	no	no	no	no	no	no
largar	yes ($N = 5$)	no	no	no	no	no	no	no
lavar	yes ($N = 2$)	no	no	no	no	no	no	no
leer	no	no	no	no	no	no	no	yes ($N = 3$)
levantar	yes ($N = 211$)	no	no	no	no	no	no	no
liar	yes ($N = 10$)	no	no	yes ($N = 11$)	no	no	no	no
librar	yes ($N = 16$)	no	no	no	no	no	no	no
licenciar	yes ($N = 3$)	no	no	no	no	no	no	no
licuar	yes ($N = 1$)	no	no	no	no	no	no	no
ligar	yes ($N = 1$)	no	no	no	no	no	no	yes ($N = 2$)

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
limpiar	yes (N = 3)	no	no	no	no	no	no	no
llamar	yes (N = 480)	no	no	no	no	no	no	no
llegar	no	no	no	no	no	yes (N = 6)	no	no
llenar	yes (N = 36)	no	no	no	no	no	no	no
llevar	no	no	no	no	no	no	no	yes (N = 232)
luchar	no	no	no	no	no	yes (N = 3)	no	no
lucir	no	no	yes (N = 7)	no	no	no	no	no
lucrar	no	no	no	no	no	no	yes (N = 3)	no
machacar	yes (N = 1)	no	no	no	no	no	no	no
madurar	yes (N = 9)	no	no	no	no	no	no	no
mamar	no	yes (N = 1)	no	no	no	no	no	no
mancar	yes (N = 3)	no	no	no	no	no	no	yes (N = 1)
manchar	yes (N = 10)	no	no	no	no	no	no	no
manejar	no	no	yes (N = 1)	no	no	no	no	no
mantener	yes (N = 19)	no	no	no	no	no	no	no
marchar	no	no	no	no	no	yes (N = 145)	no	no
marchitar	yes (N = 1)	no	no	no	no	no	no	no
marear	yes (N = 9)	no	no	no	no	no	no	no
masagar	yes (N = 1)	no	no	no	no	no	no	no
matar	yes (N = 29)	no	no	no	no	no	no	no
mear	no	no	no	no	no	yes (N = 10)	no	no
mejorar	yes (N = 17)	no	no	no	no	no	no	no
menear	yes (N = 7)	no	no	no	no	no	no	no
mentalizar	yes (N = 2)	no	no	no	no	no	no	no
merecer	no	no	no	no	no	no	no	yes (N = 6)

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
merendar	no	no	no	no	no	no	no	yes (N = 3)
mermar	yes (N = 9)	no	no	no	no	no	no	no
meter	yes (N = 314)	no	no	no	no	no	no	no
mezclar	yes (N = 4)	no	no	no	no	no	no	no
mirar	no	yes (N = 4)	no	no	no	no	no	yes (N = 3)
mocear	no	no	no	no	no	yes (N = 1)	no	no
modernizar	yes (N = 4)	no	no	no	no	no	no	no
modificar	yes (N = 1)	no	no	no	no	no	no	no
mojar	yes (N = 31)	no	no	no	no	no	no	no
moler	yes (N = 3)	no	no	no	no	no	no	no
molestar	yes (N = 5)	no	no	yes (N = 2)	no	no	no	no
molturar	no	no	no	no	no	no	no	yes (N = 1)
montar	no	no	no	no	no	yes (N = 35)	no	no
morir	no	no	no	no	no	yes (N = 284)	no	no
mosquear	no	no	no	yes (N = 1)	no	yes (N = 1)	no	no
mostear	no	no	no	no	no	no	no	no
mover	yes (N = 54)	no	no	no	no	no	no	no
mudar	yes (N = 1)	no	no	no	yes (N = 5)	no	yes (N = 1)	no
mullar	no	no	no	no	no	no	no	no
multiplicar	yes (N = 1)	no	no	no	no	no	yes (N = 1)	no
musir	no	no	no	no	no	no	yes (N = 1)	no
nacer	no	no	no	no	no	yes (N = 2)	no	no
negar	no	no	no	no	yes (N = 2)	no	no	no
nublar	yes (N = 1)	no	no	no	no	no	no	no
ocupar	no	no	no	yes (N = 13)	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
ocurrir	no	no	no	no	no	no	yes (N = 19)	no
ofrecer	yes (N = 5)	no	no	no	no	no	no	no
oir	no	yes (N = 1)	no	no	no	no	no	yes (N = 1)
olvidar	yes (N = 56)	no	no	no	yes (N = 12)	no	no	yes (N = 1)
oponer	yes (N = 1)	no	no	no	no	no	no	no
orear	yes (N = 20)	no	no	no	no	no	no	no
organizar	yes (N = 1)	no	yes (N = 1)	no	no	no	no	no
orientar	yes (N = 1)	no	no	no	no	no	no	no
orinar	no	no	no	no	no	yes (N = 1)	no	no
oxidar	yes (N = 2)	no	no	no	no	no	no	no
panelar	yes (N = 1)	no	no	no	no	no	no	no
parar	yes (N = 91)	no	no	no	no	no	no	no
partir	yes (N = 10)	no	no	no	no	no	no	no
pasar	yes (N = 404)	no	no	no	no	yes (N = 95)	no	yes (N = 69)
pasar	yes (N = 43)	no	no	no	no	yes (N = 12)	no	no
pavorar	no	no	no	yes (N = 1)	no	no	no	no
pegar	yes (N = 39)	no	no	no	no	no	no	yes (N = 6)
pelar	yes (N = 3)	no	no	no	no	no	no	yes (N = 2)
pelear	no	no	no	no	no	yes (N = 27)	no	no
pensar	no	no	no	no	no	no	no	yes (N = 16)
perder	yes (N = 55)	no	no	no	no	no	no	yes (N = 7)
perforar	yes (N = 1)	no	no	no	no	no	no	no
persignar	yes (N = 3)	no	no	no	no	no	no	no
pesar	no	no	no	no	no	yes (N = 1)	no	no
pescar	no	no	no	no	no	no	no	yes (N = 2)

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
pícar	yes (N = 4)	no	no	yes (N = 6)	no	no	no	yes (N = 1)
picardear	yes (N = 3)	no	no	no	no	no	no	no
pillar	yes (N = 1)	no	no	no	no	no	no	no
pinchar	yes (N = 4)	no	no	no	no	no	no	no
pirar	no	no	no	no	no	no	yes (N = 1)	no
pisar	no	no	no	no	no	yes (N = 1)	no	no
pitorrear	no	no	no	no	no	no	yes (N = 1)	no
plantar	yes (N = 4)	no	no	no	no	no	no	no
poner	yes (N = 716)	yes (N = 3)	no	no	no	no	no	no
porcar	yes (N = 1)	no	no	no	no	no	no	no
portar	no	no	no	no	no	no	no	no
posar	yes (N = 3)	no	no	no	no	no	yes (N = 49)	no
preciar	no	no	no	no	no	no	no	no
prender	yes (N = 13)	no	no	no	no	no	yes (N = 2)	no
prensar	yes (N = 8)	no	no	no	no	no	no	no
preñar	yes (N = 3)	no	no	no	no	no	no	no
preocupar	no	no	no	yes (N = 56)	no	no	no	no
preparar	yes (N = 5)	no	yes (N = 8)	no	no	no	no	no
presentar	yes (N = 54)	no	no	no	no	no	no	no
prestar	no	no	no	no	no	no	yes (N = 2)	no
privar	yes (N = 1)	no	no	no	no	no	no	no
propagar	yes (N = 1)	no	no	no	no	no	no	no
proposar	no	no	yes (N = 2)	no	no	no	no	no
puerir	yes (N = 23)	no	no	no	no	no	no	no
purgar	yes (N = 1)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
quebrar	yes (N = 1)	no	no	no	no	no	no	no
quedar	no	no	no	no	no	yes (N = 580)	no	yes (N = 9)
quejar	no	no	no	no	no	no	yes (N = 46)	no
quemar	yes (N = 69)	no	no	no	no	no	no	no
querer	no	no	no	no	no	no	no	yes (N = 2)
quitar	yes (N = 58)	no	no	no	no	no	no	no
rajar	yes (N = 2)	no	no	no	no	no	no	no
rasgar	yes (N = 1)	no	no	no	no	no	no	no
rebajar	yes (N = 2)	no	no	no	no	no	no	no
recalar	yes (N = 1)	no	no	no	no	no	no	no
recoger	yes (N = 8)	no	no	no	no	no	no	no
reconcentrar	yes (N = 1)	no	no	no	no	no	no	no
recordar	no	no	no	no	yes (N = 54)	no	no	yes (N = 8)
recorgajar	no	no	no	no	no	no	yes (N = 1)	no
recorrer	no	no	no	no	no	no	no	yes (N = 3)
recrear	no	no	yes (N = 2)	no	no	no	no	no
recuperar	no	no	yes (N = 1)	no	no	no	no	no
reducir	yes (N = 1)	no	no	no	no	no	no	no
referir	no	no	no	no	yes (N = 15)	no	no	no
refrescar	yes (N = 1)	no	no	no	no	no	no	no
refugiar	yes (N = 3)	no	no	no	no	no	no	no
regular	no	no	yes (N = 1)	no	no	no	no	no
regustar	no	no	no	no	no	no	yes (N = 1)	no
rehogar	yes (N = 1)	no	no	no	no	no	no	no
reír	no	no	no	no	no	yes (N = 80)	no	yes (N = 1)

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
relacionar	yes (N = 4)	no	no	no	no	no	no	no
relajar	yes (N = 3)	no	no	no	no	no	no	no
rellenar	yes (N = 2)	no	no	no	no	no	no	no
remediar	no	no	yes (N = 2)	no	no	no	no	no
remojar	yes (N = 9)	no	no	no	no	no	no	no
remontar	no	no	no	no	no	no	yes (N = 1)	no
remover	yes (N = 1)	no	no	no	no	no	no	no
rendir	yes (N = 2)	no	no	no	no	no	no	no
reñir	no	no	no	no	no	yes (N = 3)	no	no
repartir	yes (N = 4)	no	no	no	no	no	no	no
reparar	no	no	no	no	no	no	no	yes (N = 1)
repetir	yes (N = 1)	no	no	no	no	no	no	no
reponer	no	no	yes (N = 3)	no	no	no	no	no
reposar	no	no	no	no	no	yes (N = 3)	no	no
reproducir	yes (N = 1)	no	no	no	no	no	no	no
repudrir	yes (N = 2)	no	no	no	no	no	no	no
resbalar	no	no	no	no	no	yes (N = 5)	no	no
resecar	yes (N = 1)	no	no	no	no	no	no	no
resfriar	yes (N = 1)	no	no	no	no	no	no	no
resguardar	yes (N = 2)	no	no	no	no	no	no	no
resignar	no	no	no	no	no	no	yes (N = 1)	no
resistir	no	no	yes (N = 2)	no	no	no	no	no
restringir	no	no	yes (N = 1)	no	no	no	no	no
resucitar	yes (N = 1)	no	no	no	no	no	no	no
retener	yes (N = 2)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
retirar	yes (N = 18)	no	no	no	no	no	no	no
retrasar	yes (N = 1)	no	no	no	no	no	no	no
retumbar	no	no	no	no	no	yes (N = 2)	no	no
reunir	yes (N = 53)	no	no	no	no	no	no	yes (N = 2)
revalorizar	yes (N = 4)	no	no	no	no	no	no	no
revenir	no	no	no	no	no	no	yes (N = 3)	no
reventar	yes (N = 26)	no	no	no	no	no	no	no
revestir	no	no	no	no	no	no	yes (N = 2)	no
revolucionar	yes (N = 1)	no	no	no	no	no	no	no
revolver	yes (N = 3)	no	no	no	no	no	no	no
rodear	no	no	no	yes (N = 2)	no	no	no	no
rodillar	yes (N = 1)	no	no	no	no	no	no	no
romper	yes (N = 65)	no	no	no	no	no	no	no
saber	no	no	no	no	no	no	no	yes (N = 33)
sacar	yes (N = 3)	no	no	no	no	no	no	yes (N = 19)
saciar	yes (N = 1)	no	no	no	no	no	no	no
sacrificar	yes (N = 8)	no	no	no	no	no	no	no
salar	yes (N = 5)	no	no	no	no	no	no	no
salgar	yes (N = 1)	no	no	no	no	no	no	no
salir	no	no	no	no	no	yes (N = 71)	no	no
saltar	no	no	no	no	no	yes (N = 5)	no	no
salvar	yes (N = 9)	no	no	no	no	no	no	no
sanar	yes (N = 4)	no	no	no	no	no	no	no
sangrar	yes (N = 31)	no	no	no	no	no	no	no
sazonar	yes (N = 4)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
secar	yes (N = 273)	no	no	no	no	no	no	no
seguir	no	no	no	no	no	yes (N = 1)	no	no
sentar	yes (N = 184)	no	no	no	no	no	no	no
sentir	yes (N = 11)	no	no	no	no	no	no	no
separar	yes (N = 18)	no	no	no	no	no	no	no
serenar	yes (N = 2)	no	no	no	no	no	no	no
servir	yes (N = 1)	no	no	yes (N = 2)	no	yes (N = 1)	no	no
situar	no	no	yes (N = 1)	no	no	no	no	no
sobrepasar	no	no	yes (N = 1)	no	no	no	no	no
socarrar	yes (N = 6)	no	no	no	no	no	no	no
sofocar	yes (N = 1)	no	no	no	no	no	no	no
soltar	yes (N = 17)	no	no	no	no	no	no	no
sonreir	no	no	no	no	no	yes (N = 1)	no	no
sostener	yes (N = 3)	no	no	no	no	no	no	no
suavizar	yes (N = 1)	no	no	no	no	no	no	no
subir	yes (N = 331)	no	no	no	no	yes (N = 86)	no	yes (N = 4)
sublevar	yes (N = 2)	no	no	no	no	no	no	no
sujetar	yes (N = 3)	no	no	no	yes (N = 1)	no	no	no
sumir	yes (N = 1)	no	no	no	no	no	no	no
suponer	no	no	no	no	no	no	no	yes (N = 8)
suscribir	no	no	no	no	no	yes (N = 2)	no	no
tapar	yes (N = 2)	no	no	no	no	no	no	no
tardar	no	no	no	no	no	yes (N = 1)	no	no
temer	no	no	no	no	no	no	no	yes (N = 1)
tender	yes (N = 3)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
tener	yes (N = 4)	no	no	no	no	no	no	yes (N = 2)
terciar	no	no	no	no	no	no	yes (N = 4)	no
terminar	yes (N = 129)	yes (N = 1)	no	no	no	no	no	yes (N = 1)
tirar	yes (N = 40)	no	no	no	no	no	no	yes (N = 62)
titular	yes (N = 2)	no	no	no	no	no	no	no
tiznar	yes (N = 1)	no	no	no	no	no	no	no
tolerar	yes (N = 2)	no	no	no	no	no	no	no
tomar	no	no	no	no	no	no	no	yes (N = 62)
torcer	yes (N = 4)	no	no	no	no	no	no	no
tornar	yes (N = 3)	no	no	no	no	no	no	no
tostar	yes (N = 1)	no	no	no	no	no	no	no
trabajar	no	no	no	no	no	yes (N = 1)	no	yes (N = 1)
trabarr	yes (N = 3)	no	no	no	no	no	no	no
traer	no	no	no	no	no	no	no	yes (N = 98)
tragar	no	no	no	no	no	no	no	yes (N = 4)
trasladar	yes (N = 1)	no	no	no	no	no	no	no
trasponer	yes (N = 1)	no	no	no	no	no	no	no
tratar	no	no	no	no	yes (N = 1)	no	no	no
triar	no	no	no	no	no	no	yes (N = 1)	no
trillar	yes (N = 1)	no	no	no	no	no	no	no
triplicar	yes (N = 1)	no	no	no	no	no	no	no
triscar	yes (N = 1)	no	no	no	no	no	no	no
tronchar	yes (N = 1)	no	no	no	no	no	no	no
tropezar	no	no	no	no	no	no	no	yes (N = 1)
tumbar	yes (N = 13)	no	no	no	no	no	no	no

Verb	Anticausative	Absolute	Deobjective	Conversive	Antipassive	Intransitive	Non-reversible	Transitive
ubicar	no	no	yes (N = 2)	no	no	no	no	no
unir	yes (N = 7)	no	no	no	no	no	no	no
usar	no	no	no	no	no	no	no	yes (N = 1)
valer	yes (N = 10)	no	no	yes (N = 9)	no	yes (N = 2)	no	yes (N = 3)
vaporizar	yes (N = 1)	no	no	no	no	no	no	no
variar	yes (N = 16)	no	no	no	no	no	no	no
vencer	yes (N = 1)	no	no	no	no	no	no	no
vender	no	no	no	no	no	no	no	yes (N = 4)
venir	no	no	no	no	no	yes (N = 175)	no	no
ventear	yes (N = 2)	no	no	no	no	no	no	no
ventilar	yes (N = 1)	no	no	no	no	no	no	no
ver	yes (N = 31)	yes (N = 1)	no	no	no	no	no	yes (N = 7)
verter	yes (N = 2)	no	no	no	no	no	no	no
vivir	no	no	no	no	no	yes (N = 2)	no	no
volar	no	no	no	no	no	yes (N = 1)	no	no
volcar	yes (N = 4)	no	no	no	no	no	no	no
volver	yes (N = 21)	no	no	no	no	yes (N = 26)	no	no
zarandear	yes (N = 1)	no	no	no	no	no	no	no

Verb	Grooming	Body process	Posture	Non translational	Translational	Emotion
abajar	no	no	no	no	yes (<i>N</i> = 2)	no
abastecer	no	no	no	no	no	no
ablandar	no	no	no	no	no	no
ablentar	no	no	no	no	no	no
abrasar	no	no	no	no	no	no
abrir	no	no	no	no	no	no
aburrir	no	no	no	no	no	yes (<i>N</i> = 14)
abusar	no	no	no	no	no	no
acabar	no	no	no	no	no	no
acarrear	no	no	no	no	no	no
acercar	no	no	no	no	yes (<i>N</i> = 17)	no
acertar	no	no	no	no	no	no
achicar	no	no	no	no	no	yes (<i>N</i> = 2)
achichorrar	no	no	no	no	no	no
aclarar	yes (<i>N</i> = 2)	no	no	no	no	no
acordar	no	no	no	no	no	no
acortar	no	no	no	no	no	no
acostar	no	no	yes (<i>N</i> = 40)	no	no	no
acostumbrar	no	no	no	no	no	no
acudir	no	no	no	no	yes (<i>N</i> = 10)	no
acurrucar	no	no	no	yes (<i>N</i> = 1)	no	no
adaptar	no	no	no	no	no	no
adelantar	no	no	no	no	yes (<i>N</i> = 1)	no
adiestrar	no	no	no	no	no	no
administrar	no	no	no	no	no	no
adobar	no	no	no	no	no	no
afiliar	no	no	no	no	no	no
agachar	no	no	yes (<i>N</i> = 15)	no	no	no
agarrar	no	no	no	no	no	no
agotar	no	no	no	no	no	no
agravar	no	no	no	no	no	no
agriar	no	no	no	no	no	no
aguantar	no	no	no	no	no	yes (<i>N</i> = 14)
aguardar	no	no	no	no	no	no
agusanar	no	no	no	no	no	no
ahogar	no	no	no	no	no	no
ahuecar	no	no	no	no	no	no
ahumar	no	no	no	no	no	no
ajuntar	no	no	no	no	no	no
ajustar	no	no	no	yes (<i>N</i> = 1)	no	no
alargar	no	no	no	no	no	no
alegrar	no	no	no	no	no	yes (<i>N</i> = 4)
alimentar	no	no	no	no	no	no

Cognition	Perception	Spontaneous	Symmetric	No semantic ascription	Pseudo-copulative	Auxiliary
no	no	no	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	yes (N = 2)	no	no	no	no
no	no	yes (N = 15)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	yes (N = 60)	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	no	no	no
yes (N = 2)	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	no	no	no	no	no
yes (N = 491)	no	no	no	no	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 17)	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 13)	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	yes (N = 3)	no	no
no	no	yes (N = 3)	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	no	no	no
no	no	yes (N = 3)	no	yes (N = 26)	no	yes (N = 15)
no	no	yes (N = 2)	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	yes (N = 2)	no	no
no	no	yes (N = 2)	no	no	no	no
no	no	yes (N = 6)	no	no	no	no
no	no	yes (N = 2)	no	no	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	no	yes (N = 21)	no	no	no
no	no	no	no	yes (N = 3)	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 8)	no	no

Verb	Grooming	Body process	Posture	Non translational	Translational	Emotion
alistar	no	no	no	no	no	no
almacenar	no	no	no	no	no	no
alumbrar	no	no	no	no	no	no
amedrentar	no	no	no	no	no	yes (<i>N</i> = 1)
amolar	no	no	no	no	no	no
amontar	no	no	no	no	no	no
amontonar	no	no	no	no	no	no
andar	no	no	no	no	yes (<i>N</i> = 178)	no
animar	no	no	no	no	no	yes (<i>N</i> = 4)
añejar	no	no	no	no	no	no
anotar	no	no	no	no	no	no
anticipar	no	no	no	no	yes (<i>N</i> = 1)	no
antojar	no	no	no	no	no	no
anunciar	no	no	no	no	no	no
apagar	no	no	no	no	no	no
apañar	no	no	no	no	no	no
apartar	no	no	no	no	yes (<i>N</i> = 3)	no
apellidar	no	no	no	no	no	no
apelmazar	no	no	no	no	no	no
apercibir	no	no	no	no	no	no
aplastar	no	no	no	no	no	no
aplicar	no	no	no	no	no	no
apoderar	no	no	no	no	no	no
apolillar	no	no	no	no	no	no
aprensar	no	no	no	no	no	no
apretar	no	no	no	no	yes (<i>N</i> = 1)	no
aprobar	no	no	no	no	no	no
aprovechar	no	no	no	no	no	no
aproximar	no	no	no	no	yes (<i>N</i> = 1)	no
apudrir	no	no	no	no	no	no
apuntar	no	no	no	no	no	no
apurar	no	no	no	no	no	yes (<i>N</i> = 5)
arder	no	no	no	no	no	no
arrancar	no	no	no	no	no	no
arrastrar	no	no	no	no	yes (<i>N</i> = 2)	no
arrebatar	no	no	no	no	no	no
arrecoger	no	no	no	no	no	no
arreglar	no	no	no	no	no	no
arreñir	no	no	no	no	no	no
arrepentir	no	no	no	no	no	yes (<i>N</i> = 1)
arrestar	no	no	no	no	no	no
arriesgar	no	no	no	no	no	no
arrimar	no	no	no	no	yes (<i>N</i> = 3)	no

Cognition	Perception	Spontaneous	Symmetric	No semantic ascription	Pseudo-copulative	Auxiliary
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	yes ($N = 2$)	no	no
no	no	no	no	no	no	no
no	no	yes ($N = 2$)	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	no	no	no
yes ($N = 1$)	no	no	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 7$)	no	no	no	no
no	no	no	no	yes ($N = 39$)	no	no
no	no	no	yes ($N = 4$)	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	yes ($N = 1$)	no	no	no	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	yes ($N = 5$)	no	no
no	no	no	no	yes ($N = 4$)	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	yes ($N = 2$)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	yes ($N = 11$)	no	no
no	no	no	no	no	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	yes ($N = 21$)	no	no
no	no	no	no	no	no	no
no	no	yes ($N = 12$)	no	no	no	no
no	no	no	no	no	no	yes ($N = 2$)
no	no	no	no	no	no	no
no	no	yes ($N = 2$)	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 4$)	yes ($N = 3$)	yes ($N = 43$)	no	no
no	no	no	yes ($N = 1$)	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes ($N = 2$)	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	no	no	no

Verb	Grooming	Body process	Posture	Non translational	Translational	Emotion
arrodear	no	no	no	yes (N = 2)	no	no
arrodiillar	no	no	yes (N = 2)	no	no	no
arropar	yes (N = 1)	no	no	no	no	no
arruinar	no	no	no	no	no	no
asalar	no	no	no	no	no	no
asentar	no	no	yes (N = 3)	no	no	no
asfixiar	no	no	no	no	no	no
asimilar	no	no	no	no	no	no
asomar	no	no	no	yes (N = 34)	no	no
asombrar	no	no	no	no	no	yes (N = 1)
asustar	no	no	no	no	no	yes (N = 20)
atar	no	no	no	no	no	no
atontar	no	no	no	no	no	no
atrancar	no	no	no	no	no	no
atrasar	no	no	no	no	yes (N = 1)	no
atravesar	no	no	no	no	yes (N = 2)	no
atrever	no	no	no	no	no	yes (N = 10)
aumentar	no	no	no	no	no	no
aunir	no	no	no	no	no	no
avanzar	no	no	no	no	yes (N = 3)	no
avenir	no	no	no	no	no	no
bajar	no	no	no	no	yes (N = 173)	no
beneficiar	no	no	no	no	no	no
blandear	no	no	no	no	no	no
blanquear	no	no	no	no	no	no
borrar	no	no	no	no	no	no
brincar	no	no	no	yes (N = 6)	no	no
burlar	no	no	no	no	no	yes (N = 1)
buscar	no	no	no	no	no	no
cabrear	no	no	no	no	no	yes (N = 1)
caer	no	no	no	no	no	no
cagar	no	yes (N = 16)	no	no	no	no
calar	no	no	no	no	no	no
caldear	no	no	no	no	no	no
calentar	yes (N = 5)	no	no	no	no	no
callar	no	no	no	no	no	no
cambiar	no	no	no	no	no	no
canecer	no	no	no	no	no	no
cansar	no	no	no	no	no	yes (N = 29)
cargar	no	no	no	no	no	no
casar	no	no	no	no	no	no
cascar	no	no	no	no	no	no
cebar	no	no	no	no	no	yes (N = 1)

Cognition	Perception	Spontaneous	Symmetric	No semantic ascription	Pseudo-copulative	Auxiliary
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 2)	no	no
yes (<i>N</i> = 1)	no	no	no	no	no	no
no	no	yes (<i>N</i> = 2)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 2)	no	no	no	no
no	no	no	yes (<i>N</i> = 1)	no	no	no
no	no	no	no	no	no	no
no	no	no	yes (<i>N</i> = 2)	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	yes (<i>N</i> = 5)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 2)	no	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 220)	no	no	no	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 2)	no	no	no	no
no	no	yes (<i>N</i> = 2)	no	no	no	no
no	no	yes (<i>N</i> = 13)	no	no	no	no
yes (<i>N</i> = 61)	no	no	no	no	no	no
no	no	yes (<i>N</i> = 61)	no	yes (<i>N</i> = 7)	no	no
no	no	yes (<i>N</i> = 2)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 6)	no	no
no	no	no	yes (<i>N</i> = 541)	no	no	no
no	no	yes (<i>N</i> = 2)	no	no	no	no
no	no	no	no	no	no	no

Verb	Grooming	Body process	Posture	Non translational	Translational	Emotion
cenar	no	no	no	no	no	no
cerrar	no	no	no	no	no	no
changar	no	no	no	no	no	no
clamar	no	no	no	no	no	no
clarear	no	no	no	no	no	no
clavar	no	no	yes (<i>N</i> = 4)	no	no	no
coagular	no	no	no	no	no	no
cobijar	no	no	no	no	no	no
cocer	no	no	no	no	no	no
cocinar	no	no	no	no	no	no
coger	no	no	no	no	no	no
coincidir	no	no	no	no	no	no
colaborar	no	no	no	no	no	no
colar	no	no	no	no	yes (<i>N</i> = 4)	no
colegiar	no	no	no	no	no	no
colgar	no	no	yes (<i>N</i> = 2)	no	no	no
colocar	no	no	yes (<i>N</i> = 12)	no	no	no
combinar	no	no	no	no	no	no
complicar	no	no	no	no	no	no
componer	no	no	no	no	no	no
comportar	no	no	no	no	no	no
comunicar	no	no	no	no	no	no
concentrar	no	no	no	no	no	no
concretar	no	no	no	no	no	no
confesar	no	no	no	no	no	no
confiar	no	no	no	no	no	no
conformar	no	no	no	no	no	no
confundir	no	no	no	no	no	no
congelar	no	no	no	no	no	no
conservar	no	no	no	no	no	no
consumir	no	no	no	no	no	no
contaminar	no	no	no	no	no	no
contemplar	no	no	no	no	no	no
contener	no	no	no	no	no	no
correr	no	no	no	no	yes (<i>N</i> = 89)	no
cortar	no	no	no	no	no	no
crecer	no	no	no	no	no	no
criar	no	no	no	no	no	no
cuajar	no	no	no	no	no	no
cuidar	yes (<i>N</i> = 1)	no	no	no	no	no
curar	no	no	no	no	no	no
dar	no	no	no	no	no	no
decidir	no	no	no	no	no	no

Cognition	Perception	Spontaneous	Symmetric	No semantic ascription	Pseudo-copulative	Auxiliary
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 2$)	no	no	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 3$)	no	no	no	no
no	no	no	no	no	no	no
no	no	yes ($N = 5$)	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 54$)	no	no	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	yes ($N = 15$)	no	yes ($N = 7$)
no	no	no	yes ($N = 1$)	no	no	no
no	no	no	yes ($N = 4$)	no	no	no
yes ($N = 2$)	no	no	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	yes ($N = 1$)	no	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	yes ($N = 2$)	no	no
no	no	no	no	yes ($N = 5$)	no	no
no	no	no	yes ($N = 5$)	no	no	no
no	no	no	yes ($N = 1$)	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
yes ($N = 17$)	no	no	no	no	no	no
yes ($N = 1$)	no	no	no	no	no	no
yes ($N = 2$)	no	no	yes ($N = 1$)	no	no	no
yes ($N = 1$)	no	no	no	no	no	no
no	no	yes ($N = 2$)	no	no	no	no
no	no	yes ($N = 31$)	no	no	no	no
no	no	yes ($N = 5$)	no	no	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	yes ($N = 2$)	no	no
no	no	no	no	no	no	no
no	no	yes ($N = 11$)	no	no	no	no
no	no	yes ($N = 28$)	no	no	no	no
no	no	no	no	yes ($N = 160$)	no	no
no	no	yes ($N = 99$)	no	no	no	no
no	no	no	no	yes ($N = 3$)	no	no
no	no	yes ($N = 64$)	no	no	no	no
no	no	yes ($N = 8$)	no	yes ($N = 2$)	yes ($N = 2$)	yes ($N = 1$)
yes ($N = 3$)	no	no	no	no	no	no

Verb	Grooming	Body process	Posture	Non translational	Translational	Emotion
dedicar	no	no	no	no	no	no
defender	no	no	no	no	no	no
deitar	no	no	yes (<i>N</i> = 1)	no	no	no
dejar	no	no	no	no	no	no
denominar	no	no	no	no	no	no
derramar	no	no	no	no	no	no
derretir	no	no	no	no	no	no
derribar	no	no	no	no	no	no
desalar	no	no	no	no	no	no
desangrar	no	no	no	no	no	no
desaparecer	no	no	no	no	no	no
desapartar	no	no	no	no	no	no
desarmar	no	no	no	no	no	no
desarrollar	no	no	no	no	no	no
desbaratar	no	no	no	no	no	no
descargar	no	no	no	no	no	no
descender	no	no	no	no	yes (<i>N</i> = 1)	no
descomponer	no	yes (<i>N</i> = 1)	no	no	no	no
desconectar	no	no	no	no	no	no
descongelar	no	no	no	no	no	no
descuidar	no	no	no	no	no	no
desemburruñar	no	no	no	yes (<i>N</i> = 2)	no	no
desenganchar	no	no	no	no	no	no
desenvolver	no	no	no	no	no	no
desertar	no	no	no	no	no	no
desesperar	no	no	no	no	no	yes (<i>N</i> = 1)
deshacer	no	no	no	no	no	no
desjuntar	no	no	no	no	no	no
desliar	no	no	no	no	no	no
desmoralizar	no	no	no	no	no	yes (<i>N</i> = 1)
despedir	no	no	no	no	no	no
despegar	no	no	no	no	no	no
despellejar	no	no	no	no	no	no
despertar	no	no	no	no	no	no
despistar	no	no	no	no	no	no
desplazar	no	no	no	no	yes (<i>N</i> = 4)	no
desprender	no	no	no	no	no	no
destacar	no	no	no	no	no	no
destilar	no	no	no	no	no	no
diferenciar	no	no	no	no	no	no
dirigir	no	no	no	no	no	no
discutir	no	no	no	no	no	no
disfrutar	no	no	no	no	no	yes (<i>N</i> = 5)

Cognition	Perception	Spontaneous	Symmetric	No semantic ascription	Pseudo-copulative	Auxiliary
no	no	no	no	yes (N = 106)	no	no
no	no	no	no	yes (N = 17)	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 14)	no	yes (N = 1)
no	no	no	no	yes (N = 1)	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	yes (N = 2)	no	no	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	yes (N = 2)	no	no	no	no
no	no	yes (N = 2)	no	no	no	no
no	no	yes (N = 21)	no	no	no	no
no	no	no	yes (N = 2)	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	yes (N = 5)	no	no	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	yes (N = 1)	no	no	no
no	no	yes (N = 1)	no	no	no	no
yes (N = 6)	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	no	no	no
no	no	yes (N = 7)	no	yes (N = 3)	no	no
no	no	no	yes (N = 1)	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	no	no	no
no	no	no	yes (N = 4)	no	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	yes (N = 6)	no	no	no	no
yes (N = 1)	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	yes (N = 3)	no	no	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	no	no	yes (N = 4)	no	no
no	no	no	yes (N = 18)	no	no	no
no	no	no	no	no	no	no

Verb	Grooming	Body process	Posture	Non translational	Translational	Emotion
disparar	no	no	no	no	no	no
dispersar	no	no	no	no	no	no
distraer	no	no	no	no	no	no
distribuir	no	no	no	no	no	no
disuadir	no	no	no	no	no	no
divertir	no	no	no	no	no	yes (<i>N</i> = 31)
divorciar	no	no	no	no	no	no
doblar	no	no	no	yes (<i>N</i> = 2)	no	no
dormir	no	no	no	no	no	no
drogar	no	no	no	no	no	no
echar	no	no	yes (<i>N</i> = 5)	no	yes (<i>N</i> = 5)	no
echar a perder	no	no	no	no	no	no
educar	no	no	no	no	no	no
emborrachar	no	no	no	no	no	no
empadronar	no	no	no	no	no	no
empanar	no	no	no	no	no	no
empeñar	no	no	no	no	no	no
empezar	no	no	no	no	no	no
empinar	no	no	no	yes (<i>N</i> = 1)	no	no
emplear	no	no	no	no	no	no
empollar	no	no	no	no	no	no
emporcar	yes (<i>N</i> = 1)	no	no	no	no	no
empreñar	no	no	no	no	no	no
enamorar	no	no	no	no	no	yes (<i>N</i> = 10)
encalar	no	no	no	no	no	no
encanecer	no	no	no	no	no	no
encarecer	no	no	no	no	no	no
encargar	no	no	no	no	no	no
encender	no	no	no	no	no	no
encerrar	no	no	no	no	no	no
encher	no	no	no	no	no	no
enchufar	no	no	no	no	no	no
encoger	no	no	no	yes (<i>N</i> = 1)	no	no
enconar	no	no	no	no	no	no
encontrar	no	no	no	no	no	no
endurecer	no	no	no	no	no	no
enemistar	no	no	no	no	no	no
enfadar	no	no	no	no	no	yes (<i>N</i> = 20)
enfermar	no	no	no	no	no	no
enfriar	no	no	no	no	no	no
enganchar	no	no	no	no	no	no
engañar	no	no	no	no	no	no
engarañar	no	no	no	no	no	no

Cognition	Perception	Spontaneous	Symmetric	No semantic ascription	Pseudo-copulative	Auxiliary
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
yes (<i>N</i> = 6)	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	no	no	no	no	no
no	no	no	yes (<i>N</i> = 7)	no	no	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 106)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 3)	no	no
no	no	no	no	yes (<i>N</i> = 1)	yes (<i>N</i> = 1)	yes (<i>N</i> = 7)
no	no	yes (<i>N</i> = 11)	no	no	no	no
yes (<i>N</i> = 1)	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 11)	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
yes (<i>N</i> = 3)	no	no	no	yes (<i>N</i> = 2)	no	no
no	no	yes (<i>N</i> = 18)	no	yes (<i>N</i> = 1)	no	yes (<i>N</i> = 2)
no	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 3)	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 8)	no	no	no	no
no	no	no	yes (<i>N</i> = 2)	no	no	no
no	no	yes (<i>N</i> = 2)	no	no	no	no
no	no	yes (<i>N</i> = 2)	no	no	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 48)	no	no
no	no	yes (<i>N</i> = 2)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	yes (<i>N</i> = 3)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 4)	no	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 25)	yes (<i>N</i> = 8)	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	no	yes (<i>N</i> = 1)	no	no	no
no	no	no	yes (<i>N</i> = 1)	no	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	yes (<i>N</i> = 22)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 9)	no	yes (<i>N</i> = 7)
yes (<i>N</i> = 1)	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no

Verb	Grooming	Body process	Posture	Non translational	Translational	Emotion
englobar	no	no	no	no	no	no
engordar	no	no	no	no	no	no
enjuagar	no	no	no	no	no	no
enmendar	no	no	no	no	no	yes (N = 2)
enranciar	no	no	no	no	no	no
enredar	no	no	no	no	no	yes (N = 1)
enrollar	no	no	no	no	no	no
ensayar	no	no	no	no	no	no
enseñar	no	no	no	no	no	no
entender	no	no	no	no	no	no
enterar	no	no	no	no	no	no
enterrar	no	no	no	no	no	no
entrar	no	no	no	no	yes (N = 271)	no
entregar	no	no	no	no	no	no
entretener	no	no	no	no	no	yes (N = 13)
envolver	no	no	no	yes (N = 4)	no	no
equivocar	no	no	no	no	no	no
escabultar	no	no	no	no	no	no
escaldar	no	no	no	no	no	no
escapar	no	no	no	no	yes (N = 41)	no
escarriar	no	no	no	no	no	no
esconder	no	no	no	no	no	no
escullar	no	no	no	no	no	no
escurrir	no	no	no	no	no	no
esforzar	no	no	no	no	no	no
esmerar	no	no	no	no	no	no
espabilar	no	no	no	no	no	no
espantar	no	no	no	no	no	yes (N = 11)
especializar	no	no	no	no	no	no
esperar	no	no	no	no	no	no
espesar	no	no	no	no	no	no
espizar	no	no	no	no	no	no
esponjar	no	no	no	no	no	no
establecer	no	no	no	no	no	no
estampar	no	no	no	no	no	no
estar	no	no	no	no	no	no
estilar	no	no	no	no	no	no
estirar	no	no	no	yes (N = 3)	no	no
estropear	no	no	no	no	no	no
esvolar	no	no	no	no	yes (N = 1)	no
eternecer	no	no	no	no	no	no
existir	no	no	no	no	no	no
expatriar	no	no	no	no	no	no

Cognition	Perception	Spontaneous	Symmetric	No semantic ascription	Pseudo-copulative	Auxiliary
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 13$)	no	no	no	no
no	no	yes ($N = 2$)	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	no	no	yes ($N = 3$)
no	no	no	yes ($N = 1$)	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
yes ($N = 1$)	no	no	no	no	no	no
yes ($N = 11$)	no	no	no	no	no	no
yes ($N = 46$)	no	no	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	no	no	no
no	no	no	no	yes ($N = 6$)	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
yes ($N = 8$)	no	no	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 2$)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	yes ($N = 36$)	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	yes ($N = 9$)	no	no
yes ($N = 1$)	no	no	no	no	no	no
no	no	no	no	yes ($N = 2$)	no	no
yes ($N = 1$)	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	yes ($N = 55$)	no	no
no	no	yes ($N = 2$)	no	no	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	yes ($N = 2$)	no	no	no	no
no	no	no	no	yes ($N = 2$)	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	yes ($N = 88$)	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	no	no	no
no	no	yes ($N = 40$)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	yes ($N = 1$)	no	no

Verb	Grooming	Body process	Posture	Non translational	Translational	Emotion
explicar	no	no	no	no	no	no
explotar	no	no	no	no	no	no
exponer	no	no	no	no	no	no
expresar	no	no	no	no	no	no
exprimir	no	no	no	no	no	no
extender	no	no	no	no	no	no
extrañar	no	no	no	no	no	no
fermentar	no	no	no	no	no	no
fiar	no	no	no	no	no	no
figurar	no	no	no	no	no	no
fijar	no	no	no	no	no	no
filtrar	no	no	no	no	no	no
fomentar	no	no	no	no	no	no
formar	no	no	no	no	no	no
forrar	no	no	no	no	no	no
freir	no	no	no	no	no	no
frenar	no	no	no	no	yes (<i>N</i> = 1)	no
fugar	no	no	no	no	yes (<i>N</i> = 1)	no
fusionar	no	no	no	no	no	no
girar	no	no	no	yes (<i>N</i> = 4)	no	no
gobernar	no	no	no	no	no	no
guardar	no	no	no	no	no	no
guiar	no	no	no	no	no	no
habituarse	no	no	no	no	no	no
hacer	no	no	no	no	no	no
hartar	no	no	no	no	no	yes (<i>N</i> = 8)
helar	no	no	no	no	no	no
herniar	no	yes (<i>N</i> = 1)	no	no	no	no
hervir	no	no	no	no	no	no
hincharse	no	no	no	no	no	no
hospedar	no	no	no	no	no	no
hundir	no	no	no	no	no	no
imponer	no	no	no	no	no	no
incendiar	no	no	no	no	no	no
incomodar	no	no	no	no	no	yes (<i>N</i> = 2)
incorporar	no	no	no	no	no	no
inculpar	no	no	no	no	no	no
infectar	no	yes (<i>N</i> = 3)	no	no	no	no
inflamar	no	no	no	no	no	no
inflar	no	no	no	no	no	no
instalar	no	no	no	no	no	no
ir	yes (<i>N</i> = 72)	no	no	no	yes (<i>N</i> = 3522)	no
joder	no	no	no	no	no	no

Cognition	Perception	Spontaneous	Symmetric	No semantic ascription	Pseudo-copulative	Auxiliary
yes (<i>N</i> = 15)	no	no	no	no	no	no
no	no	yes (<i>N</i> = 4)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 5)	no	no
yes (<i>N</i> = 1)	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
yes (<i>N</i> = 1)	no	no	no	no	no	no
no	no	yes (<i>N</i> = 27)	no	no	no	no
yes (<i>N</i> = 5)	no	no	no	no	no	no
yes (<i>N</i> = 1)	no	no	no	no	no	no
no	yes (<i>N</i> = 173)	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 2)	no	no
no	no	yes (<i>N</i> = 2)	no	no	no	no
no	no	yes (<i>N</i> = 2)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	yes (<i>N</i> = 2)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	yes (<i>N</i> = 2)	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 5)	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	no	no	yes (<i>N</i> = 3)	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
yes (<i>N</i> = 4)	no	yes (<i>N</i> = 19)	yes (<i>N</i> = 17)	yes (<i>N</i> = 158)	yes (<i>N</i> = 82)	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 8)	no	no	no	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 53)	no	no	no	no
no	no	yes (<i>N</i> = 11)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	yes (<i>N</i> = 8)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 2)	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 6)	no	no
no	no	no	no	yes (<i>N</i> = 2)	no	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 3)	no	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 5)	no	no	no	no

Verb	Grooming	Body process	Posture	Non translational	Translational	Emotion
jubilar	no	no	no	no	no	no
jugar	no	no	no	no	no	no
juntar	no	no	no	no	no	no
lanzar	no	no	no	no	yes (N = 2)	no
levantar	no	no	yes (N = 94)	no	no	no
liar	no	no	no	no	no	no
librar	no	no	no	no	no	no
licenciar	no	no	no	no	no	no
ligar	no	no	no	no	no	no
llamar	no	no	no	no	no	no
llegar	no	no	no	no	yes (N = 378)	no
llenar	no	no	no	no	no	no
luchar	no	no	no	no	no	no
lucir	no	no	no	no	no	no
lucrar	no	no	no	no	no	no
madurar	no	no	no	no	no	no
mamar	no	no	no	no	no	no
mancar	no	yes (N = 1)	no	no	no	no
manchar	yes (N = 4)	no	no	no	no	no
manejar	no	no	no	no	no	no
mantener	no	no	no	no	no	no
marchar	no	no	no	no	yes (N = 211)	no
marear	no	no	no	no	no	no
matar	no	no	no	no	no	no
mear	no	yes (N = 11)	no	no	no	no
mejorar	no	no	no	no	no	no
menear	no	no	no	no	yes (N = 4)	no
mentalizar	no	no	no	no	no	no
meter	no	no	no	no	yes (N = 155)	no
mezclar	no	no	no	no	no	no
mirar	no	no	no	no	no	no
mocear	no	no	no	no	no	no
modernizar	no	no	no	no	no	no
mojar	no	no	no	no	no	no
moler	no	no	no	no	no	no
molestar	no	no	no	no	no	yes (N = 3)
montar	no	no	no	no	yes (N = 18)	no
morir	no	no	no	no	no	no
mover	no	no	no	no	yes (N = 23)	no
mudar	no	no	no	no	no	no
multiplicar	no	no	no	no	no	no
musir	no	no	no	no	no	no
nacer	no	no	no	no	no	no

Cognition	Perception	Spontaneous	Symmetric	No semantic ascription	Pseudo-copulative	Auxiliary
no	no	no	no	yes (N = 56)	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	yes (N = 140)	no	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
yes (N = 3)	no	yes (N = 2)	no	yes (N = 10)	no	yes (N = 10)
no	no	no	no	yes (N = 15)	no	no
no	no	no	no	yes (N = 3)	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	no	no	yes (N = 478)	no	no
no	no	no	no	no	no	no
no	no	yes (N = 13)	no	no	no	no
no	no	no	yes (N = 1)	no	no	no
no	no	no	no	yes (N = 6)	no	no
no	no	no	no	yes (N = 3)	no	no
no	no	yes (N = 3)	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	yes (N = 19)	no	no
no	no	no	no	no	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	yes (N = 11)	no	yes (N = 1)	no	no
no	no	no	no	no	no	no
no	no	yes (N = 6)	no	no	no	no
no	no	no	no	no	no	no
yes (N = 1)	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	yes (N = 1)	no	no	no
no	yes (N = 63)	no	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	yes (N = 14)	no	no	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	yes (N = 424)	no	no	no	no
no	no	no	no	no	no	no
no	no	yes (N = 1)	no	yes (N = 3)	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	yes (N = 185)	no	no	no	no

Verb	Grooming	Body process	Posture	Non translational	Translational	Emotion
negar	no	no	no	no	no	no
nublar	no	no	no	no	no	no
ocupar	no	no	no	no	no	no
ocurrir	no	no	no	no	no	no
ofrecer	no	no	no	no	no	no
oir	no	no	no	no	no	no
olvidar	no	no	no	no	no	no
oponer	no	no	no	no	no	no
orear	no	no	no	no	no	no
organizar	no	no	no	no	no	no
orientar	no	no	no	no	no	no
orinar	no	yes (N = 2)	no	no	no	no
panelar	no	no	no	no	no	no
parar	no	no	no	no	yes (N = 52)	no
partir	no	no	no	no	no	no
pasar	no	no	no	no	yes (N = 206)	no
pasear	no	no	no	no	yes (N = 17)	no
pavorar	no	no	no	no	no	yes (N = 1)
pegar	no	no	no	no	no	no
pelar	no	no	no	no	no	no
pelear	no	no	no	no	no	no
perder	no	no	no	no	no	no
persignar	no	no	no	no	no	no
pesar	no	no	no	no	no	no
picar	no	no	no	no	no	no
pisar	no	yes (N = 1)	no	no	no	no
pitorrear	no	no	no	no	no	yes (N = 1)
plantar	no	no	no	no	no	no
poner	no	no	yes (N = 103)	no	no	no
portar	no	no	no	no	no	no
posar	no	no	yes (N = 2)	no	no	no
preciar	no	no	no	no	no	no
prender	no	no	no	no	no	no
preñar	no	no	no	no	no	no
prensar	no	no	no	no	no	no
preocupar	no	no	no	no	no	yes (N = 17)
preparar	no	no	no	no	no	no
presentar	no	no	no	no	no	no
prestar	no	no	no	no	no	no
privar	no	no	no	no	no	no
propagar	no	no	no	no	no	no
propasar	no	no	no	no	no	no
pudrir	no	no	no	no	no	no

Cognition	Perception	Spontaneous	Symmetric	No semantic ascription	Pseudo-copulative	Auxiliary
yes (<i>N</i> = 1)	no	no	no	no	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 13)	no	no
yes (<i>N</i> = 7)	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 5)	no	no
no	yes (<i>N</i> = 2)	no	no	no	no	no
yes (<i>N</i> = 31)	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	yes (<i>N</i> = 4)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 2)	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 1)	no	yes (<i>N</i> = 1)	no	no
no	no	yes (<i>N</i> = 216)	no	yes (<i>N</i> = 1)	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	yes (<i>N</i> = 3)	yes (<i>N</i> = 36)	no	no
no	no	yes (<i>N</i> = 1)	no	yes (<i>N</i> = 2)	no	no
no	no	no	yes (<i>N</i> = 16)	no	no	no
no	no	yes (<i>N</i> = 17)	no	yes (<i>N</i> = 17)	no	no
no	no	no	no	yes (<i>N</i> = 3)	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	no	no	yes (<i>N</i> = 2)	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 4)	no	no
no	no	yes (<i>N</i> = 2)	yes (<i>N</i> = 1)	yes (<i>N</i> = 452)	yes (<i>N</i> = 194)	yes (<i>N</i> = 108)
no	no	no	no	yes (<i>N</i> = 49)	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (<i>N</i> = 2)	no	no
no	no	yes (<i>N</i> = 6)	no	no	no	no
no	no	yes (<i>N</i> = 3)	no	no	no	no
no	no	yes (<i>N</i> = 5)	no	no	no	no
no	no	no	no	no	no	no
no	no	yes (<i>N</i> = 3)	no	yes (<i>N</i> = 8)	no	no
no	no	no	no	yes (<i>N</i> = 54)	no	no
no	no	no	no	yes (<i>N</i> = 2)	no	no
no	no	no	no	yes (<i>N</i> = 1)	no	no
no	no	yes (<i>N</i> = 1)	no	no	no	no
no	no	no	no	yes (<i>N</i> = 2)	no	no
no	no	yes (<i>N</i> = 8)	no	no	no	no

Verb	Grooming	Body process	Posture	Non translational	Translational	Emotion
purgar	no	no	no	no	no	no
quedar	no	no	no	no	no	no
quejar	no	no	no	no	no	yes (N = 24)
quemar	no	no	no	no	no	no
quitar	no	no	no	no	yes (N = 16)	no
rajar	no	no	no	no	no	no
rasgar	no	no	no	no	no	no
recalar	no	no	no	no	no	no
recoger	no	no	no	no	no	no
reconcentrar	no	no	no	no	no	no
recordar	no	no	no	no	no	no
recuperar	no	no	no	no	no	no
referir	no	no	no	no	no	no
refrescar	no	no	no	no	no	no
refugiar	no	no	no	no	no	no
regular	no	no	no	no	no	no
rehogar	no	no	no	no	no	no
reir	no	no	no	no	no	yes (N = 48)
relacionar	no	no	no	no	no	no
remediar	no	no	no	no	no	no
remontar	no	no	no	no	no	no
remover	no	no	no	no	no	no
rendir	no	no	no	no	no	no
reñir	no	no	no	no	no	no
repartir	no	no	no	no	no	no
repetir	no	no	no	no	no	no
reponer	no	no	no	no	no	yes (N = 2)
reposar	no	no	no	no	no	no
reproducir	no	no	no	no	no	no
resbalar	no	no	no	no	no	no
resguardar	no	no	no	no	no	no
resignar	no	no	no	no	no	no
resistir	no	no	no	no	no	no
restringir	no	no	no	no	no	no
resucitar	no	no	no	no	no	no
retener	no	no	no	no	no	no
retirar	no	no	no	no	yes (N = 12)	no
retrasar	no	no	no	no	yes (N = 1)	no
retumbar	no	no	no	no	no	no
reunir	no	no	no	no	no	no
revalorizar	no	no	no	no	no	no
revenir	no	no	no	no	no	no
reventar	no	no	no	no	no	no

Cognition	Perception	Spontaneous	Symmetric	No semantic ascription	Pseudo-copulative	Auxiliary
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	yes ($N = 580$)	yes ($N = 937$)	no
no	no	no	no	no	no	no
no	no	yes ($N = 20$)	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	yes ($N = 8$)	no	no
no	no	no	yes ($N = 1$)	no	no	no
yes ($N = 23$)	no	no	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
yes ($N = 3$)	no	no	no	no	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	yes ($N = 3$)	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	yes ($N = 1$)	no	no	no
no	no	no	no	yes ($N = 2$)	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	yes ($N = 2$)	no	no
no	no	no	yes ($N = 7$)	no	no	no
no	no	no	no	yes ($N = 4$)	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	no	no	no	no	no
no	no	no	no	yes ($N = 3$)	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 2$)	no	no	no	no
no	no	no	no	yes ($N = 2$)	no	no
yes ($N = 1$)	no	no	no	no	no	no
no	no	no	no	yes ($N = 2$)	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	yes ($N = 2$)	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes ($N = 2$)	no	no
no	no	no	yes ($N = 14$)	no	no	no
no	no	yes ($N = 4$)	no	no	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	yes ($N = 3$)	no	no	no	no

Verb	Grooming	Body process	Posture	Non translational	Translational	Emotion
revolucionar	no	no	no	no	no	no
revolver	no	no	no	no	no	no
rodear	no	no	no	yes (N = 2)	no	no
rodillar	no	no	yes (N = 1)	no	no	no
romper	no	no	no	no	no	no
sacar	no	no	no	no	yes (N = 2)	no
saciar	no	no	no	no	no	no
sacrificar	no	no	no	no	no	no
salar	no	no	no	no	no	no
salgar	no	no	no	no	no	no
salir	no	no	no	no	yes (N = 710)	no
saltar	no	no	no	yes (N = 28)	no	no
salvar	no	no	no	no	no	no
sanar	no	no	no	no	no	no
sangrar	no	no	no	no	no	no
secar	no	no	no	no	no	no
seguir	no	no	no	no	no	no
sentar	no	no	yes (N = 90)	no	no	no
sentir	no	no	no	no	no	no
separar	no	no	no	no	no	no
servir	no	no	no	no	no	no
situar	no	no	no	no	no	no
sobrepasar	no	no	no	no	no	no
socarrar	no	no	no	no	no	no
sofocar	no	no	no	no	no	yes (N = 1)
soltar	no	no	no	no	yes (N = 10)	no
sostener	no	no	no	no	no	no
suavizar	no	no	no	no	no	no
subir	no	no	no	no	yes (N = 181)	no
sublevar	no	no	no	no	no	no
sujetar	no	no	no	no	no	no
sumir	no	no	no	no	yes (N = 1)	no
suscribir	no	no	no	no	no	no
tapar	yes (N = 1)	no	no	no	no	no
tardar	no	no	no	no	no	no
tender	no	no	yes (N = 1)	no	no	no
tener	no	no	no	yes (N = 1)	no	no
terciar	no	no	no	no	no	no
terminar	no	no	no	no	no	no
tirar	no	no	no	no	yes (N = 24)	no
titular	no	no	no	no	no	no
tolear	no	no	no	no	no	yes (N = 2)
torcer	no	no	no	yes (N = 2)	no	no

Cognition	Perception	Spontaneous	Symmetric	No semantic ascription	Pseudo-copulative	Auxiliary
no	no	no	no	yes (N = 1)	no	no
no	no	no	yes (N = 1)	no	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no
no	no	yes (N = 26)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	yes (N = 8)	no	no
no	no	yes (N = 3)	no	no	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	yes (N = 208)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 9)	no	no
no	no	yes (N = 4)	no	no	no	no
no	no	yes (N = 18)	no	no	no	no
no	no	yes (N = 131)	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 10)	yes (N = 2)	no
no	no	no	yes (N = 12)	no	no	no
no	no	no	no	yes (N = 4)	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	yes (N = 1)
no	no	no	no	yes (N = 3)	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 2)	no	no
no	no	no	no	yes (N = 4)	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 2)	no	no
no	no	yes (N = 1)	no	no	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	yes (N = 1)	no	no
no	no	no	no	no	no	no
no	no	no	no	yes (N = 4)	no	no
no	no	yes (N = 71)	no	yes (N = 1)	no	no
no	no	no	no	no	no	yes (N = 3)
no	no	no	no	yes (N = 2)	no	no
no	no	no	no	no	no	no
no	no	no	no	no	no	no

Verb	Grooming	Body process	Posture	Non translational	Translational	Emotion
tornar	no	no	no	no	no	no
trabajar	no	no	no	no	no	no
trabar	no	no	no	no	no	no
trasponer	no	no	no	no	yes (<i>N</i> = 1)	no
tratar	no	no	no	no	no	no
tronchar	no	no	no	no	no	no
tumbar	no	no	yes (<i>N</i> = 7)	no	no	no
ubicar	no	no	no	no	no	no
unir	no	no	no	no	no	no
valer	no	no	no	no	no	no
variar	no	no	no	no	no	no
venir	no	no	no	no	yes (<i>N</i> = 2148)	no
ver	no	no	no	no	no	no
vivir	no	no	no	no	no	no
volar	no	no	no	no	yes (<i>N</i> = 7)	no
volcar	no	no	no	no	no	no
volver	no	no	no	yes (<i>N</i> = 4)	yes (<i>N</i> = 93)	no

Cognition	Perception	Spontaneous	Symmetric	No semantic ascription	Pseudo-copulative	Auxiliary
no	no	no	no	no	yes ($N = 2$)	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 2$)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes ($N = 1$)	no	no
no	no	yes ($N = 1$)	no	no	no	no
no	no	no	no	no	no	no
no	no	no	no	yes ($N = 2$)	no	no
no	no	no	yes ($N = 3$)	no	no	no
no	no	no	no	yes ($N = 21$)	no	no
no	no	yes ($N = 5$)	no	no	no	no
no	no	no	no	no	no	no
no	yes ($N = 36$)	no	no	yes ($N = 31$)	yes ($N = 12$)	no
no	no	no	no	yes ($N = 2$)	no	no
no	no	no	no	no	no	no
no	no	yes ($N = 2$)	no	no	no	no
no	no	no	no	yes ($N = 16$)	yes ($N = 6$)	no

APPENDIX 5

List of verbs analysed in corpus eseuTenTen11

Verb	Transitive	Intransitive	Total (tr. + intr.)	Prob. trans.	No RM	RM	Total (no RM + RM)	Prob. RM
abrir	75	20	95	0.79	1	19	20	0.95
acabar	16	54	70	0.23	28	9	37	0.24
acercar	41	106	147	0.28	0	106	106	1
acostar	12	116	128	0.09	0	116	116	1
acostumbrar	13	43	56	0.23	0	43	43	1
adaptar	49	62	111	0.44	0	62	62	1
agachar	33	83	116	0.28	1	82	83	0.99
ahogar	60	58	118	0.51	0	58	58	1
apuntar	84	56	140	0.6	0	31	31	1
arrimar	93	44	137	0.68	0	44	44	1
arrodillar	4	107	111	0.04	0	107	107	1
calentar	90	52	142	0.63	12	40	52	0.77
cambiar	77	65	142	0.54	61	4	65	0.06
coagular	20	69	89	0.22	27	42	69	0.61
cocer	47	44	91	0.52	29	15	44	0.34
colar	42	95	137	0.31	14	81	95	0.85
colocar	116	11	127	0.91	0	11	11	1
conservar	128	3	131	0.98	0	3	3	1
cortar	97	8	105	0.92	0	8	8	1
criar	76	36	112	0.68	0	36	36	1
cuajar	34	82	116	0.29	75	7	82	0.09
curar	83	31	114	0.73	3	28	31	0.9
dar	122	5	127	0.96	1	4	5	0.8
defender	124	16	140	0.89	0	16	16	1
desaparecer	3	120	123	0.02	120	0	120	0
deshacer	42	97	139	0.3	1	96	97	0.99
despertar	81	52	133	0.61	31	21	52	0.4
desplazar	28	99	127	0.22	0	99	99	1
echar	115	17	132	0.87	3	14	17	0.82
echar a perder	76	57	133	0.57	1	56	57	0.98
emborrachar	30	112	142	0.21	0	112	112	1
empezar	34	28	62	0.55	28	0	28	0
encontrar	76	69	145	0.52	0	69	69	1

Verb	Transitive	Intransitive	Total (tr. + intr.)	Prob. trans.	No RM	RM	Total (no RM + RM)	Prob. RM
enfriar	39	103	142	0.27	52	51	103	0.5
enganchar	67	35	102	0.66	1	34	35	0.97
engordar	80	65	145	0.55	64	1	65	0.02
enterar	0	140	140	0	0	140	140	1
equivocar	1	105	106	0.01	0	105	105	1
esconder	53	51	104	0.51	0	51	51	1
escurrir	91	29	120	0.76	12	17	29	0.59
estropear	67	63	130	0.52	0	63	63	1
fermentar	11	46	57	0.19	45	1	46	0.02
hacer	106	11	117	0.91	0	11	11	1
helar	9	14	23	0.39	3	11	14	0.79
hervir	40	80	120	0.33	79	1	80	0.01
hinchar	31	47	78	0.4	1	46	47	0.98
hundir	46	74	120	0.38	0	74	74	1
joder	53	12	65	0.82	0	12	12	1
jubilar	13	59	72	0.18	0	59	59	1
levantar	66	60	126	0.52	0	60	60	1
librar	41	91	132	0.31	6	85	91	0.93
llamar	72	34	106	0.68	0	34	34	1
llenar	110	39	149	0.74	1	38	39	0.97
manchar	51	28	79	0.65	0	28	28	1
mantener	111	33	144	0.77	0	33	33	1
matar	138	2	140	0.99	0	2	2	1
mejorar	125	20	145	0.86	18	2	20	0.1
meter	86	46	132	0.65	0	46	46	1
mojar	41	40	81	0.51	0	40	40	1
mover	54	87	141	0.38	0	87	87	1
olvidar	84	48	132	0.64	0	48	48	1
orear	21	27	48	0.44	9	18	27	0.67
parar	28	64	92	0.3	41	23	64	0.36
pasar	25	62	87	0.29	57	3	60	0.05
pasear	20	126	146	0.14	97	27	124	0.22
pegar	97	15	112	0.87	0	15	15	1
perder	114	13	127	0.9	0	13	13	1
poner	112	20	132	0.85	0	20	20	1
prender	72	15	87	0.83	11	4	15	0.27
presentar	118	16	134	0.88	0	16	16	1
pudrir	10	49	59	0.17	2	47	49	0.96
quemar	88	22	110	0.8	0	22	22	1
quitar	141	3	144	0.98	0	3	3	1
retirar	92	45	137	0.67	0	45	45	1

Verb	Transitive	Intransitive	Total (tr. + intr.)	Prob. trans.	No RM	RM	Total (no RM + RM)	Prob. RM
reventar	71	62	133	0.53	57	5	62	0.08
romper	95	20	115	0.83	5	15	20	0.75
sangrar	14	100	114	0.12	99	0	99	0
secar	83	50	133	0.62	2	48	50	0.96
sentar	11	24	35	0.31	7	17	24	0.71
soltar	131	14	145	0.9	0	14	14	1
terminar	35	52	87	0.4	49	3	52	0.06
tirar	124	10	134	0.93	1	9	10	0.9
tumbar	40	35	75	0.53	1	34	35	0.97

Index

A

absolute construction 29,
66–68, 78, 102, 107, 199–200,
238, 252
accidentality 177, 179, 184
actor 24, 62
adjunct 51, 235
adjuncts 51, 136, 217, 236, 248,
250
affected participant 17, 22, 27,
29, 61–62, 64, 99, 103–104,
142–145, 148–149, 157, 172–173,
175–176, 185, 190, 206, 211,
220, 224–227, 230–231, 241,
268–269, 283
affectedness 59, 80, 91, 95,
99–100, 102, 104, 126, 142–
143, 145, 149, 172–173, 190,
206, 222, 224, 245, 286
agent 5, 11, 13, 27, 29, 32, 34–36,
62–63, 75–76, 80, 87, 103–104,
116, 179, 185, 187, 196, 216,
220–221, 225–226, 236, 241–
243, 253, 262, 283, 285, 287
agentive 36–37, 62–63, 125, 133,
168–170, 193–194, 196, 203,
225–226, 243–244, 248, 250,
256, 262–263, 283, 287, 289
agentivity 168–170, 193,
195–196, 201, 203, 240, 242,
248, 253, 261
agents 36, 62, 64, 129, 168, 179,
211, 248
allobenefactive(s) 276–277
alternation
antipassive 12, 80–81
causative-anticausative 11,
60–62, 64–65, 75, 83,
97–100, 103–104, 109,
124–127, 129–136, 144, 158,
179–180, 180, 208, 225, 227
conversive 12, 76, 107

diathesis 10, 28, 86, 107,
106, 113
transitive-intransitive *see*
alternation (causative-
anticausative)
Amharic 253
analogy 3, 25, 40–41, 149,
157–160, 162, 164, 168, 170,
176, 182, 198, 203, 206–207,
209–212, 215, 229, 233–234,
244, 251, 255–256, 259–260,
265, 278–279, 282, 285–288
paradigmatic 3, 159, 168,
182, 256, 285
syntagmatic 3, 40–41, 164,
168, 256, 285
animacy 16–18, 24, 50, 61–63,
97, 149–155, 162–164, 168–171,
180–183, 201–204, 211–212,
232, 239, 249, 266–267, 284
anticausative 60–65, 76, 81, 97–
98, 104, 119–156, 162–164, 187,
198, 211–212, 225–228, 283, 285
antipassive 12, 78–82, 106,
119, 150
argument
external 62, 64, 104, 116,
129–130, 172, 179, 185, 190,
206, 211
internal 209
aspectual *se* 125, 137–142, 144,
148, 157, 160, 165, 168, 176–
178, 185–186, 203, 206, 209,
215, 237, 241, 248, 251, 253–257,
262, 278, 283–284, 287, 292
Asturian 4, 8–15, 18–19, 45, 216,
218, 224, 284
attributive verbs 108–109, 192,
195, 197–198, 201
autobenefactive
construction 32–33, 220–223,
241–244, 256, 259, 261, 287

meaning 230–231, 234–236,
242, 277, 285
autocausative verbs 61–62, 170,
193, 203
auxiliary verbs 111

B

benefaction 220–221, 242
benefactive
participants 33, 215–218,
220–221, 234, 241, 245,
276, 277
construction 221, 241
beneficiary *see* benefactive
participants
body
part 22, 217, 224, 226
process 35, 88–89, 93–96,
116, 160, 185–188, 211, 216,
260

C

causation 24, 119, 129–132, 134,
143–144, 154, 185, 211, 225, 283
external 24, 64, 119, 125,
129–134, 148, 152, 154–155,
157, 175–176, 185, 211,
225–226, 283, 285
internal 119, 129–134, 154,
185, 283
cause argument 12, 24, 62–64,
76, 125, 129–133, 136, 142–146,
148–149, 152, 154–155, 157,
172–173, 175–176, 178–179,
182–183, 185, 211, 225–226,
268, 283, 285
unintentional 129, 131,
142–145, 172, 211
causer *see* cause argument
cognition verbs 35, 88–89,
98–102, 108–109, 219, 226,
229–231, 251–261, 265, 278

colloquial uses 190–191, 244–245, 275–277
 communicative immediacy *see* colloquial uses
 consumption verbs 208, 230–231, 237–245, 260, 286
 control 24, 36, 80, 87, 99, 104, 187, 189, 194–195, 201, 211–212, 226, 260, 285
 conversive verbs 75–77, 81, 100, 107, 187, 211–212, 225–226, 283, 285
 coreferentiality 8–9, 32, 34, 91, 93, 216, 218, 220, 225, 242, 273, 283
 corporal *see* body process
 counter-expectation 175, 284
 Cushitic 31, 103, 158

D

dative 119, 145–148, 153–156, 172–173, 180–184, 218–219, 227, 268–271, 273, 283
 affected 17, 142–145, 148, 172–173, 190, 206, 211, 268–269
 benefactive 15, 217–218, 220
 causer 131, 142–149
 ethical 13, 25
 experiencer 99, 145, 212
 locative 145, 173, 182, 190, 268–269
 possessive 71, 142–145, 149, 173, 182, 190, 215, 217–218, 222–226, 233, 268, 288
 recipient 8, 215–220, 226, 251, 269
 reflexive *see* indirect reflexive construction
 decausative verb 20–21, 61
 demoted 13, 32, 75, 78–79, 116, 282–283
 de-objective verb 32, 66, 70–74, 81, 106, 119, 149, 200
 detransitivisation *see* indirect detransitivised construction
 diffusion
 spatial 19, 38–39, 41–42
 lexical 128, 207, 209–210
 direct reflexive construction 4, 7, 9–10, 21–22, 35, 60,

92–95, 154, 215–216, 259, 283–285, 87
 distinguishability
 of participants 33–36, 91, 96, 99, 109, 132, 216
 of events 91, 132

E

elaboration
 of events 33–34, 59, 114
 of participants 103, 132
 emotion verbs 35, 75–76, 81, 87–89, 100–102, 158, 187, 211–212, 226
 emphasis 9, 24, 126, 221, 242–245, 248, 256, 274–275, 285
 emphatic prepositional phrase 6–7, 10, 34, 93, 98, 227
 energetic *se* 25, 29, 59, 87, 103–104, 168, 177–178
 English 28, 61, 70, 72, 92–94, 97, 195, 223, 256–257
 experiencer 11–13, 18, 25, 36, 75–76, 87, 98–102, 145, 196, 212, 227, 248, 283

F

final result state 108, 137–140, 143–144, 165, 187, 203, 205
 French 31, 104, 223

G

Galician 1–2, 4–6, 8–15, 18–19, 45, 51, 54, 64, 86, 90, 108, 113, 131, 173, 216, 218, 224, 251, 275, 281, 284
 German 102, 104, 129–130, 134, 223
 goal role 25, 145, 167, 173–175, 182, 268
 grammaticality 52
 grooming verb 35, 93–95, 160

H

human external participant 185, 206

I

imperative 68, 95, 169–170, 176–177, 196, 201, 203, 250, 253, 266, 271, 277
 imperfective tense 94, 186–187, 190, 260
 implicature 172, 241–242, 255–258, 277, 287
 impressiveness *see* emphasis
 inceptive aspect 16, 187, 189
 incipient reading 187, 190, 260
 indirect
 detransitivised construction 225–226, 228, 259, 265, 285
 reflexive construction 4, 7–10, 13, 18, 21–22, 26, 33, 35, 60, 71, 215–219, 221, 225–229, 233, 245, 251, 270, 273, 279, 283, 287
 infinitive 77, 90–91, 139, 141
 innovation 39–42, 283–284, 288
 intentionality 62, 145, 177, 179, 189, 226, 261–264, 283
 intransitive reflexive verb 10–11, 13–16, 21, 24–25, 82–85, 104, 107, 115, 157–215, 226, 230, 234, 247, 260, 266, 268–272, 274, 279, 281, 283–284, 286–287
 intransitivisation 7, 11–12, 16, 20–21, 29, 31, 37, 61, 64, 109, 115, 126–127, 149, 193, 211, 283–284
 involuntary *see* intentionality

L

lability 28, 61, 65, 120, 124, 133–134, 142–143
 language change 19, 38–41
 lexical causative 179, 187, 207, 209–210, 212, 282–283, 295
 lexicalisation 21, 70, 109, 212, 279, 285–287
 light verb 226, 252, 259–260, 275
 location (change of) 24, 124, 178
 locative phrases 25, 174, 176, 178, 182, 190, 196, 224, 248, 250, 268–269, 286

M

- malefactive 245
- mental event 24, 35, 71, 73, 87–88, 98–100, 109, 219
- metaphorical use 25, 73, 92
- motion event 14, 16, 24–26, 35, 83–84, 93, 96–98, 116–117, 124, 145, 158, 165–178, 207–208, 211, 227, 230, 266–276
- movement *see* motion event

N

- non-oppositional *see* non-reversible
- non-paradigmatic reflexive 5, 17–18, 31, 33, 60
- non-reversible 37, 64, 81, 86–90, 106, 158, 175, 251
- non-translational event 35, 88, 93, 96–97, 116, 158

O

- object
 - delimited 237–240, 242–245, 248, 254, 262, 286
 - non-delimited 237–240, 248, 262
 - null 66, 69, 240, 257
 - oblique 12, 15, 32, 66, 75–76, 78–80, 86, 89–90, 135, 192, 199, 208, 223, 234
 - partitive 70–71, 238, 286
 - prepositional *see* object (oblique)
 - temporal 246–250
- one-form language 34, 92

P

- paradigmatic reflexive 1, 4–5, 10, 15, 18, 26, 40, 60
- participation (degree of) 27, 33, 36, 62, 93, 99
- passive
 - reflexive 5, 11, 17, 20–21, 24, 31, 37, 60, 103, 135
 - voice 27, 29, 32, 75, 78
 - experiencer 78, 87, 99–100
- perception 35, 66–67, 75, 90, 99–100, 102–103, 109, 116, 219
- perfective tense 31
- periphrases 111, 139–142

- periphrastic expression 228, 260
- physical process 103, 185–192
- Portuguese 18, 104
- possession
 - external 71, 217, 223–224
 - possession 71, 224
- posture (change of body) 16, 35, 93, 95–96, 116, 167
- predicative *see* secondary predicate
- pseudocopulative verb 108–110, 201, 235
- psychological 16, 25, 71, 185–192

Q

- questionnaire 52–56, 126, 183–184, 191–192, 207, 210, 262–264, 276

R

- reanalysis 22, 25, 39–41, 283–284, 286
- reciprocal 4, 6–7, 9–10, 16, 20–23, 26, 35, 99, 105–106, 158, 160–161, 261, 287

S

- secondary predicate 108, 259
- self-benefaction *see* autobenefactive
- source 99, 167, 171–176, 182–183, 185, 212, 244, 268–269
- space (delimited) 166, 185
- speech verb 8, 24, 73, 80, 83, 101, 109, 219–220, 276
- spontaneous event 24, 35, 89, 103–105, 108, 117, 158–159, 185, 203–207, 211–212
- state (change of) 16, 24, 29, 35–36, 62, 79, 87, 89–90, 93, 103–104, 108, 124, 131–133, 137, 139, 141, 145, 159, 177–178, 198, 201, 206, 254
- stimulus 75–76, 98–100, 102
- subject
 - activity of 25
 - human 12, 16–17, 149, 179, 187, 193–195, 197, 207, 211–212

- inanimate 11, 16–17, 21, 24, 97, 103, 149–153, 155, 162–163, 168, 172, 180, 190, 198–199, 201–204, 211, 232, 240, 248, 262, 266, 283
- involvement of 80, 93, 221, 236, 240, 242–245, 248, 253, 272, 285
- responsibility of 47, 87, 104, 175, 193, 201
- temporal 205
- subjunctive 138–139
- symmetric predicate *see* reciprocal

T

- telicity 24, 126, 137, 141, 144, 160, 209, 243, 284
- temporal phrase 196, 199, 248
- tense
 - imperfective 31, 66, 186–187, 190, 260
 - perfective 31
- transitive reflexive verb 4, 10, 13, 15, 21, 24–25, 92, 115, 155, 162, 191–192, 208, 212, 215–279, 283–285, 287
- transitivity 10–11, 13, 29–31, 59, 79–80, 99, 126–127, 285
- translational verb 35, 83, 88, 93, 95, 97–98, 116–117, 124, 158, 165–177, 207, 211, 231, 266–276
- two-form language 34, 93

U

- unaccusative 11, 18, 24, 62, 131, 223, 226
- unergative 24, 131
- unexpected event 25, 34, 177, 184–185, 211, 262
- unintentional
 - event 177, 261–265
- unplanned event *see* unintentional event

V

- volitional 95, 99–100, 102–104, 168, 187, 196
- voluntary *see* intentionality

The reflexive constructions that are the focus of this book are the constructions broadly described with the term “middle”: i.e., those that can appear in all persons, and in which the reflexive marker (RM) cannot be understood as a full referential pronoun. One goal of this study is to provide a corpus-based typology of middle and related uses that allow us to compare the behaviour of the RM in these constructions with previous typological accounts, where competing models (based either on changes of diathesis or on the semantics of the verbal event) can be found. A second goal is to shed light on the evolution of the different functions of the RM, by exploring the factors that affect its productivity, with a specific focus on those verbs where reflexive marking is most variable, that is, anticausative verbs and verbs with no change of valency. These reflexive constructions show a notable difference in productivity in Spanish and Galician, although the languages are closely related and contiguous. The languages are thus good candidates for a contrastive and variationist analysis serving these two goals. The semantic class of the predicate, its aspectual properties and the animacy of the subject are some of the most relevant factors that are taken into account to understand the motivations behind the presence (or absence) of the RM. By relying on a corpus of interviews from rural communities across peninsular Spain (except Catalonia), space as a relevant extra-linguistic variable is taken into account, helping uncover previously unknown geographical patterns.

ISBN 978 90 272 1142 2



9 789027 211422

John Benjamins Publishing Company