Differential Object Marking and Role Semantics in Romance

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Philosophischen Fakultät der Universität zu Köln
im Fach Romanische Philologie

vorgelegt von
Sophie Mürmann

geb. am 10. Dezember 1989
in Werne

Köln, 18. März 2021

This publication is a slightly revised version of my doctoral dissertation that was accepted by the Faculty of Arts and Humanities of the University of Cologne in July 2021. The project has been funded by the doctoral programme of the a.r.t.e.s. Integrated Track (2016-2019), the *a.r.t.e.s. international – for all* travel grant and the *Katharina Niemeyer-Grant for Language Editing*.

First supervisor: Prof. Dr. Marco García García

Second supervisor: Prof. Dr. Martin Becker

Third supervisor: Prof. Dr. Nikolaus P. Himmelmann

Chairperson of the committee: Prof. Dr. Daniel Bunčić

Date of the defensio: 7th July 2021

Abstract

This study provides a fine-grained approach to the impact of agentivity on Differential Object Marking (DOM) in Romance languages. While verbal factors have been considered relevant for the understanding of DOM since decades, this dissertation is the first to systematically motivate and investigate communicated agentivity for the direct object as an influencing factor. As a first step, clear role-semantic criteria are elaborated to analyze DOM from a verb-based perspective, independently of animacy. For this purpose, the proto-role model of Blume (1998, 2000) is adopted, which is a modified version of Dowty's (1991) and Primus' (1999) proto-role model. Blume distinguishes between entailed, presupposed and conversationally implicated proto-agent properties for the object. This differentiation is adopted, underpinned by semantic tests allowing to isolate the type of proto-agent properties assigned to the object for a given predicate. Together with a fourth option that is defined by the absence of any proto-agent property, the different types of communicated agentivity for the object are arranged on a four-point scale (*entailed agentivity > presupposed agentivity > potential agentivity > unspecified for agentivity*). As a second step, experimental evidence for the impact of agentivity on DOM is presented for a Western Sicilian variety (Alcamo) and two Central Catalan varieties (Barcelona and Girona). Four verb classes are defined in accordance with the agentivity scale, while the direct object's animacy is kept constant. Both acceptability judgement studies provide further evidence for treating agentivity as a co-influential factor on DOM. As a third step, the findings for Sicilian and Catalan are related to synchronic and diachronic evidence of further Romance languages. It is sketched how role-semantic factors interact with referentiality-based and information-structural parameters (e.g. definiteness and topicality) in early stages of grammaticalization. In sum, the dissertation makes both a theoretical and an empirical contribution to the impact of agentivity on DOM in Romance languages.

Acknowledgements

There are many people who supported me on my way to finish this dissertation and to whom I am very grateful. First and foremost, I thank Marco García García for being the best supervisor I could have wished for. He brought to me the idea of investigating DOM and role semantics across Romance languages and awakened my fascination for verbs in general and role semantics in particular. Throughout all phases of my dissertation, he was present and approachable, being at all times open for inspiring discussion (sometimes even quite spontaneously). His valuable feedback and encouragement always took me a step further. Thanks to his great motivation and trust, especially during the final phase, this project came to a good end. I am grateful to Martin Becker for accompanying the different stages of the project as my second supervisor. He not only helped me shape my project in the beginning but at any occasion gave me insightful remarks that made me rethink parts of my work, in particular with regard to my studies on Sicilian and Catalan. I thank my third supervisor Nikolaus Himmelmann, who, after my Masters, encouraged me to pursue my PhD in Romance linguistics and who remained approachable for questions and discussion. Furthermore, I am thankful to Daniel Bunčić for acting as the chairperson of the committee at my defensio.

I am highly grateful for the three-year scholarship I got from the a.r.t.e.s. *Graduate School for the Humanities Cologne*, as well as for additional fundings I gained, namely the *a.r.t.e.s. international – for all* travel grant and the *Katharina Niemeyer-Grant for Language Editing*. My special thanks go to the SFB1252 *Prominence in Language* of DFG (Deutsche Forschungsgemeinschaft) for providing me the opportunity to publish this dissertation in their series *Papers on Prominence*. During my PhD, I was an associated member of the SFB project B04 *Interaction of nominal and verbal features for Differential Object Marking*, led by Klaus von Heusinger and Marco García García. I am very grateful for an inspiring and productive time within the project team.

I thank Riccardo Ginevra for introducing me to Antonino Rubino. He and his whole family supported my study in Alcamo. They gave me a warm welcome and brought me in contact with locals and Sicilian native speakers, such as Stefano Milotta, with whom I recorded my test sentences. I am deeply grateful for their helpfulness and hospitality. Just as much I thank Anna Pineda who took the time to go through my Catalan test sentences with me during her stay in Cologne. Thanks to Artur Müller-Nübling and Nerea Contreras Benito, these sentences finally got recorded for the varieties of Barcelona and Girona, respectively. Not to forget, I thank Maximilian Hörl for his support with the statistical analysis of my data.

There were many people I met at the *Romanisches Seminar* of the University of Cologne and the SFB *Prominence in Language* who contributed to the inspiring working environment it was and who I enjoyed sharing coffee and ideas with. I particularly thank Diego Romero Heredero, Maria Katarzyna Prenner, Elyesa Seidel and Eric Engel. I am very grateful to Javier Caro Reina for his advice, his honest and critical

feedback as well as discussions on the exciting topic of proper names. My project also greatly benefited from the exchange with Alessia Cassarà which started on Sicilian and ended up in an article on DOM in Italian. I am very happy that Semra Kizilkaya and Clara Stumm crossed my path at the beginning of my PhD. Both were invaluable discussion partners not only on my dissertation, but also on various non-linguistic topics. I am thankful to Pedro Alejandro Martínez Pacheco and Julio Manero González who always lent me an open ear and cheered me up when needed.

I am deeply grateful to my proofreaders Tim Könenberg, Semra Kizilkaya and Clara Stumm (once again). You did a tough job reading through my chapters and gave me very helpful feedback – I am happy I could count on you. As a last-minute proofreader, Faith Ann Gibson carefully lectured the entire manuscript and polished up my English. I thank her for her corrections as well as for her flexibility. All remaining mistakes are mine.

My greatest thanks go to my family. My sisters, Clara and Theresa, luckily never stopped asking about my dissertation, visited me in the library for coffee breaks and took my mind off linguistic theory. My parents, Marianne and Ludger, have supported me throughout my studies and always encouraged me to follow my interests. I am deeply grateful for their patience and especially for overcoming the last months of writing with me. This is for you.

Table of contents

Abstract ..... **III**

Acknowledgements ..... **IV**

Tables and Figures ..... **IX**

Glosses ..... **XIII**

1. Introduction **1**

1.1 Research questions and goals 2

1.2 Theoretical framework and empirical methodology 4

1.3 Overview of the chapters 4

2. DOM variation in Romance **7**

2.1 Nominal factors (and topicality) 7

2.2 Verbal factors 13

2.2.1 Spanish 13

2.2.2 Sicilian 16

2.2.3 Catalan 17

2.3 Summary 19

3. (Proto-)Agentivity, agent-patient asymmetry and (morpho)syntactic argument realization **20**

3.1 Syntactic argument realization in Dowty's (1991) proto-role model 20

3.2 Application to morphosyntactic argument realization 31

3.2.1 Primus (1999a, b, 2006) 32

3.2.2 Blume (1998, 2000) 43

3.3 Application to differential morphosyntactic argument realization: DOM in Spanish (García García 2007, 2014) 57

3.4 Summary 67

4. Communicated agentivity for the object and DOM: Definition of verb classes **69**

4.1 Agentivity and *a*-marking in Spanish 69

4.2 Types of communicated agentivity for the object 73

4.2.1 Empirical diagnostics 74

a. Volition entailed for O 74

b. Sentience entailed for O 77

c. No co-argument dependency 82

d. Autonomous activity presupposed for O 84

e. Sentience presupposed for O 86

f. Implicature of sentience for O 90
g. Subject potential of O 96
h. Summary 98
4.2.2 Proposal of an agentivity scale 98
4.3 Agentivity scale and DOM in Romance languages: A survey of verb classes 100
4.3.1 Entailed agentivity for the object 100
4.3.2 Presupposed agentivity for the object 109
4.3.3 Potential agentivity for the object 113
4.3.4 Object unspecified for agentivity 116
4.4 Working hypothesis 121
4.5 Summary 122
5. DOM and communicated agentivity in Sicilian and Catalan 124
5.1 Acceptability judgement study on Sicilian 124
5.1.1 Variety of Alcamo (province of Trapani, Western Sicilian) 125
5.1.2 Hypotheses 129
5.1.3 Study design 132
5.1.4 Participants 137
5.1.5 Procedure and data analysis 139
5.1.6 Results 140
5.1.6.1 Hypothesis 1 (Agentivity & DOM) 140
5.1.6.2 Hypothesis 2 (Referentiality & DOM) 143
5.1.6.3 Hypothesis 3 (Agentivity & Referentiality & DOM) 146
5.1.7 Discussion 152
5.2 Acceptability judgement study on Catalan 157
5.2.1 Varieties of Barcelona and Girona (Central Catalan) 158
5.2.2 Hypotheses 159
5.2.3 Study design 160
5.2.4 Participants 166
5.2.5 Procedure and data analysis 167
5.2.6 Results 169
5.2.6.1 Hypothesis 1 (DOM vs. no-DOM) 169
5.2.6.2 Hypothesis 2 (Agentivity & DOM) 172
5.2.7 Discussion 181
5.3 General discussion of the results 185

5.4 Summary ... 187
6. Role of agentivity for DOM in Romance ... 190
6.1 Agentivity as a factor for DOM in Romance ... 190
6.1.1 Sicilian and Catalan ... 191
6.1.2 Spanish ... 194
6.1.3 Other Romance languages ... 204
6.2 Interaction of agentivity with nominal factors (and topicality) ... 212
6.2.1 Agentivity and animacy ... 212
6.2.2 Agentivity and definiteness ... 221
6.2.3 Interaction of factors in grammaticalization ... 223
6.3 Summary ... 229
7. Conclusion ... 231
References ... 244
Appendix ... 261

Tables and Figures

List of tables

Chapter 2

Table 1. Cut-off points for DOM in Romance languages and varieties 12
Table 2. References on synchronic cut-off points for DOM in Romance languages and varieties (as displayed in Table 1) 13

Chapter 3

Table 1. Dowty's (1991: 579f.) analysis of SE- and OE-psych verbs 27
Table 2. Revised version of Dowty's (1991: 579f.) analysis of SE- and OE-psych verbs 28
Table 3. Semantic transitivity and case frames for a selection of German verbs 53

Chapter 4

Table 1. Sub-properties of *sentience* 80
Table 2. Applicability of diagnostics to the four levels of communicated agentivity for the object 98
Table 3. Cross-classification of agentivity scale (horizontal) and agent-patient asymmetry (vertical) 99
Table 4. Role-semantic configurations for *entailed agentivity* 101
Table 5. Role-semantic configurations for *presupposed agentivity* 110
Table 6. Role-semantic configurations for *potential agentivity* 113
Table 7. Role-semantic configurations for *unspecified for agentivity* 117

Chapter 5

Table 1. Forms of the definite article in the dialects of Palermo and Alcamo (Western Sicilian) 126
Table 2. Participants of the study 138
Table 3. Estimated marginal means (EMM) of acceptability values in condition 1a (Agentivity & DOM) and 1b (Agentivity & no-DOM), degrees-of-freedom method: kenward-roger, confidence level used: 0.95 142
Table 4. Parameter estimates for the fixed effects in condition 1a (Agentivity & DOM) 143
Table 5. Parameter estimates for the fixed effects in condition 1b (Agentivity & no-DOM) 143
Table 6. Estimated marginal means (EMM) of acceptability values in condition 2a (Referentiality & DOM) and 2b (Referentiality & no-DOM), degrees-of-freedom method: kenward-roger, confidence level used: 0.95 145

Table 7. Parameter estimates for the fixed effects in condition 2a (Referentiality & DOM) and 2b (Referentiality & no-DOM) 146
Table 8. Estimated marginal means (EMM) of acceptability values in condition 3a (Agentivity & definite NP & DOM) and 3b (Agentivity & definite NP & no-DOM), degrees-of-freedom method: kenward-roger, confidence level used: 0.95 149
Table 9. Estimated marginal means (EMM) of acceptability values in condition 3c (Agentivity & indefinite NP & DOM) and 3d (Agentivity & indefinite NP & no-DOM), degrees-of-freedom method: kenward-roger, confidence level used: 0.95 150
Table 10. Parameter estimates for the fixed effects in condition 3a (Agentivity & definite NP & DOM) 150
Table 11. Parameter estimates for the fixed effects in condition 3b (Agentivity & definite NP & no-DOM) 151
Table 12. Parameter estimates for the fixed effects in condition 3c (Agentivity & indefinite NP & DOM) 151
Table 13. Parameter estimates for the fixed effects in condition 3d (Agentivity & indefinite NP & no-DOM) 152
Table 14. Adverbials for *entailed agentivity* 156
Table 15. Adverbials for *unspecified for agentivity* 157
Table 16. Cut-off points for DOM in different Catalan varieties 158
Table 17. Use of the onymic marker and definite article with personal names in Catalan varieties (Caro Reina 2014: 183) 159
Table 18. Use of the onymic marker and definite article with personal names in the two Central Catalan varieties of Barcelona and Girona 159
Table 19. Participants of the study (Barcelona) 167
Table 20. Participants of the study (Girona) 167
Table 21. Estimated marginal means (EMM) of acceptability values for DOM vs. no-DOM (H1) with definite human NPs for Barcelona and Girona, degrees-of-freedom method: kenward-roger, confidence level used: 0.95 172
Table 22. Parameter estimates for DOM vs. no-DOM (H1) with definite human NPs for Barcelona and Girona, degrees-of-freedom method: kenward-roger, t-test 172
Table 23. Estimated marginal means (EMM) of acceptability values in condition 2a (Agentivity & DOM) and 2b (Agentivity & no-DOM) for Barcelona, degrees-of-freedom method: kenward-roger, confidence level used: 0.95 175
Table 24. Parameter estimates for the fixed effects in condition 2a (Agentivity & DOM) for Barcelona 176
Table 25. Parameter estimates for the fixed effects in condition 2b (Agentivity & no-DOM) for Barcelona 176
Table 26. Estimated marginal means (EMM) of acceptability values in condition 2a (Agentivity & DOM) and 2b (Agentivity & no-DOM) for Girona, degrees-of-freedom method: kenward-roger, confidence level used: 0.95 178
Table 27. Parameter estimates for the fixed effects in condition 2a (Agentivity & DOM) for Girona 179

Table 28. Parameter estimates for the fixed effects in condition 2b (Agentivity & no-DOM) for Girona 180

Chapter 6

Table 1. A collection of verbs that are recurrently reported to obligatorily take DOM with human indefinite direct objects in Spanish 198
Table 2. Individual verbs that are recurrently reported to obligatorily take DOM with human indefinite direct objects in Spanish, categorized along the agentivity hierarchy (in square brackets verbs that optionally take the marker order reject it) 199
Table 3. Verbs attested for different Romance languages and varieties showing a diachronic preference for DOM 206
Table 4. Characterization of the four types of communicated agentivity for the object 214
Table 5. Animacy and potential agentivity (García García 2018b: 60) 215

Chapter 7

Table 1. Cross-classification of agentivity scale (horizontal) and agent-patient asymmetry (vertical) 233

List of figures

Chapter 5

Figure 1. Boxplots of acceptability values in condition 1a (Agentivity & DOM) and 1b (Agentivity & no-DOM) with median values and outliers 141
Figure 2. Estimated marginal means (EMM) of acceptability values in condition 1a (Agentivity & DOM) and 1b (Agentivity & no-DOM), confidence level used: 0.95 142
Figure 3. Boxplots of acceptability values in condition 2a (Referentiality & DOM) and 2b (Referentiality & no-DOM) with median values and outliers 144
Figure 4. Estimated marginal means (EMM) of acceptability values in condition 2a (Referentiality & DOM) and 2b (Referentiality & no-DOM), confidence level used: 0.95 145
Figure 5. Boxplots of acceptability values in condition 3a (Agentivity & definite NP & DOM) and 3b (Agentivity & definite NP & no-DOM) with median values and outliers 147
Figure 6. Estimated marginal means (EMM) of acceptability values in condition 3a (Agentivity & definite NP & DOM) and 3b (Agentivity & definite NP & no-DOM), confidence level used: 0.95 148

Figure 7. Boxplots of acceptability values in condition 3c (Agentivity & indefinite NP & DOM) and 3d (Agentivity & indefinite NP & no-DOM) with median values and outliers 148
Figure 8. Estimated marginal means (EMM) of acceptability values in condition 3c (Agentivity & indefinite NP & DOM) and 3d (Agentivity & indefinite NP & no-DOM), confidence level used: 0.95 149
Figure 9. Boxplots of acceptability values for DOM vs. no-DOM (H1) with definite human NPs with median values and outliers for Barcelona..... 170
Figure 10. Boxplots of acceptability values for DOM vs. no-DOM (H1) with definite human NPs with median values and outliers for Girona..... 170
Figure 11. Estimated marginal means (EMM) for DOM vs. no-DOM (H1) with definite human NPs for Barcelona, confidence level used: 0.95..... 171
Figure 12. Estimated marginal means (EMM) for DOM vs. no-DOM (H1) with definite human NPs for Girona, confidence level used: 0.95..... 171
Figure 13. Boxplots of acceptability values in condition 2a (Agentivity & DOM) and 2b (Agentivity & no-DOM) with median values and outliers for Barcelona..... 174
Figure 14. Estimated marginal means (EMM) in condition 2a (Agentivity & DOM) and 2b (Agentivity & no-DOM) for Barcelona, confidence level used: 0.95..... 174
Figure 15. Boxplots of acceptability values in condition 2a (Agentivity & DOM) and 2b (Agentivity & no-DOM) with median values and outliers for Girona..... 178
Figure 16. Estimated marginal means (EMM) in condition 2a (Agentivity & DOM) and 2b (Agentivity & no-DOM) for Girona, confidence level used: 0.95..... 178

Chapter 7

Figure 1. Haspelmath's (2015) data in Malchukov's (2005) two-dimensional hierarchy 242

Glosses

1	first person
2	second person
3	third person
ACC	accusative
CL	clitic
DAT	dative
DEF	definite
DEM	demonstrative
DIM	diminutive
DOM	differential object marker
ERG	ergative
F	feminine
GEN	genitive
IMP	imperative
INF	infinitive
IO	indirect object marker
IPRF	imperfect
M	masculine
NEG	negative particle
NOM	nominative
OM	onymic marker
PASS	passive
PL	plural
PN	personal name
POSS	possessive
PRF	perfect
PRS	present
PST	past
PTCP	participle
REFL	reflexive pronoun
SG	singular

1. Introduction

Among argument realization phenomena, the differential morphosyntactic marking of direct objects has attracted special attention in the research of recent decades. This phenomenon, dubbed *Differential Object Marking* (DOM) by Bossong (1982, 1985), is present in over 300 genetically related and unrelated languages. In these languages, some direct objects are morphologically marked, while others remain unmarked. In Romance, Spanish is a case in point for DOM, illustrated by the examples given in (1):

(1) a. *Pepe vio* ****ø/a*** *la actriz.*
PN see.PST.3SG ø/DOM the actress
'Pepe saw the actress.'

b. *Pepe vio* ***ø/a*** *una actriz.*
PN see.PST.3SG ø/DOM a actress
'Pepe saw an actress.'

c. *Pepe vio* ***ø/*a*** *la casa.*
PN see.PST.3SG ø/DOM the house
'Pepe saw the house.'

The above examples show the typical pattern of DOM in European Spanish: while the *a*-marker is obligatory with human definite NPs (1a), it is usually optional with human indefinite NPs (1b) and ungrammatical with inanimate NPs (1c). Accordingly, both inherent and referential properties of the object NP are argued to determine the presence of the marker, and in a similar way but to different degrees, in other Romance languages such as Catalan, Corsican, Sardinian or Sicilian (cf. Bossong 1998 for an overview). Animacy is thereby seen as the dominant factor. From a cross-linguistic perspective, too, the phenomenon of DOM is mainly described as expanding along the animacy and/or the definiteness scale (Aissen 2003: 437, 471f.; Bossong 1991: 158-162).

Yet for Spanish, an animacy- and definiteness-based approach on DOM is questioned by systematic occurrences of the marker with inanimate objects (cf. Weissenrieder 1985, 1991; García García 2007, 2014; Primus 2012a; García García/Primus/Himmelmann 2018). These instances of the *a*-marker (e.g., *Un artículo precede *ø/a un sustantivo* 'An article precedes a noun'), attested for certain predicate classes, have been instead accounted for within a role-semantic model based on generalized semantic roles (Dowty 1991; Primus 1999a, b, 2006). More precisely, DOM is argued to occur with inanimate objects in Spanish if the subject does not outrank the direct object in terms of agentivity, and as a consequence, the latter functions as a weak *proto-agent* in the given event (García García 2007: 71, 2014: 145). In this view,

DOM can be captured as a linking phenomenon just as non-differential case marking since the morphosyntactic realization of the direct object can be predicted by the relative distribution of role-semantic properties among the co-arguments. This suggests that common linking theories might be a suitable framework for the investigation of DOM in general. One major challenge for the general implementation of a role-semantic approach on DOM is the dissociation of the verbal factor agentivity from the object-inherent property animacy. In a generalization formulated by Primus (2012a: 81), and further elaborated by García García/Primus/Himmelmann (2018: 27), it is proposed that DOM with animate objects can be reasonably subsumed under a role-semantic approach:

(2) Role-dependent DOM: A differential object marker is licensed by an object qualifying as a minimal or potential proto-agent in a given event. The proto-agent properties may either be **subcategorized by the verb** or **assigned according to intrinsic properties of the object referent**. (García García/Primus/Himmelmann 2018: 27; emphasis: S.M.)

Hence, while for example the direct object *la actriz* 'the actress' in (1a) above (*Pepe vio *ø/a la actriz* 'Pepe saw the actress') is not assigned any proto-agent property by the verb, it can potentially function as a proto-agent in the very same event and occupy the subject role due to its intrinsic proto-agent properties derived from its (cognitive-) ontological status as a human being. So, on the one hand, the licensing condition in (2) elegantly combines the verb-based factor agentivity and the nominal-factor animacy in one unified explanation on DOM. On the other hand, however, from an empirical perspective that takes verb meanings as the starting point, it is difficult to derive clear role-semantic criteria from (2) allowing to determine the potential agentivity of an object participant in a given event and make detailed predictions on its morphosyntactic realization.

1.1 Research questions and goals

This study wants to step in here and further develop the idea of applying common linking theories to DOM in Romance, putting special emphasis on animate direct objects. The empirical focus will be on two Romance languages in which DOM is in an advanced though not highly grammaticalized stage: Sicilian and Catalan. Given this stage of development, a dissociation of agentivity from the object-inherent factor animacy, which has not been tackled in previous studies so far, appears especially feasible. I formulated three research questions (RQs) to be answered in the course of this study:

1. DOM from a linking perspective (RQ 1)

How can we include DOM in a broader theory on linking? In particular, how can we capture proto-agent properties of the direct object in more precise terms in order to clearly disentangle agentivity from animacy?

2. DOM and acceptability in Sicilian and Catalan (RQ 2)

Why are acceptability judgement tasks of particular relevance for a synchronic view of DOM? Is there a difference in acceptability of DOM between different degrees of communicated agentivity in the two languages?

3. Implications for a model on DOM in Romance (RQ 3)

What do the results for Sicilian and Catalan tell us about the impact on the object's agentivity on DOM and its interaction with nominal factors? Can they be related to findings for DOM in Spanish? As for DOM in Romance in general, how can we model the impact of role-semantic factors and the interaction with other factors?

RQ 1 aims at elaborating clear role-semantic criteria that allow us to analyze the phenomenon from a verb-based perspective, independently of animacy. For this purpose, I will adopt the proto-role model of Blume (1998, 2000) which is a modified version of Dowty's (1991) and Primus' (1999a, b) proto-role model. Blume distinguishes between *entailed*, *presupposed* and *conversationally implicated* proto-agent properties for the object. This differentiation is adopted, underpinned by semantic tests that allow us to isolate the type of proto-agent properties assigned to the object for a given predicate. Together with a fourth option that is defined by the absence of any proto-agent property, the different types of communicated agentivity for the object are arranged on a four-point scale (*entailed agentivity* > *presupposed agentivity* > *potential agentivity* > *unspecified for agentivity*).

Concerning **RQ 2**, acceptability judgement tests will be conducted for Sicilian and Catalan in order to obtain a detailed picture of the perceived grammaticality of predicates with different types of communicated agentivity for the object in combination with and without DOM. Four verb classes will be defined in accordance with the scale previously introduced, while the object's animacy will be kept constant. Since the two Romance languages have not fully grammaticalized DOM, effects are expected to be clearer than for a language like Spanish that has grammaticalized DOM for nearly every human definite NP.

As for **RQ 3**, I will attempt to establish a generalization on the impact of agentivity on DOM in both languages. In a further step, the findings for Sicilian and Catalan are related to what has been reported on the impact of agentivity on DOM with inanimate objects in Spanish. Finally, taking into account evidence of further Romance languages, I attempt to model the impact of role-semantic factors on DOM in relation to referentiality-based and information-structural parameters, such as definiteness and topicality.

This study wants to make a more general contribution to the understanding of differential morphosyntactic argument realization phenomena. It will show that role-semantic properties that are of minor relevance for linking in a narrow sense, such as the proto-agent property *sentience*, may have a systematic impact on the morphosyntactic expression of an argument. In addition, it provides further empirical evidence on the synchronic impact of verbal factors on DOM in Romance languages. In contrast to nominal factors such as animacy and definiteness as well as the information-structural parameter topicality, verbal factors are still less well studied for DOM.

Although the empirical basis is limited to Sicilian and Catalan, points of reference to other Romance languages and varieties are made at various points. Moreover, it is clearly indicated throughout this work that DOM is a multifactorial phenomenon that cannot be regarded as triggered by one parameter in isolation.

1.2 Theoretical framework and empirical methodology

My approach is theoretically anchored within the framework of generalized semantic roles. Dowty (1991) introduced a model operating with two cluster concepts of semantic roles, a proto-agent and a proto-patient. Both proto-roles are characterized by a list of verbal entailments in the strictly logical sense. By the relative assignment of proto-agent and proto-patient properties for a given transitive predicate, syntactic argument selection, i.e. the selection of subject and direct object, can be predicted. Further research by Primus (1999a, b, 2006) modified Dowty's (1991) model in order to make predictions on universal preferences in morphosyntactic case selection. García García (2007, 2014) made use of this model for his role-semantic account on differentially marked inanimate objects in Spanish. In order to also apply a role-semantic model to DOM with animate objects, further adjustments are needed. For this purpose, I will argue for the adoption of Blume's (1998, 2000) modification of Dowty's and Primus' proto-role models. The theoretical innovation of this study is the development of a four-point agentivity scale that dissociates different types of communicated agentivity for the object and ranks them with each other.

Based on the theoretical preliminaries, a distinction between different verbs and verb classes can be made for the two languages under investigation, Sicilian and Catalan. Importantly, we can clearly disentangle role-semantic information from animacy as an inherent object property. In order to empirically test whether different types of agentivity for the object have an effect on DOM, I conducted an acceptability judgement study for both languages, exposing native speakers to spoken sentences containing the predicates in question. Acceptability judgement tests have the advantage that they allow the verbs in question to be tested in a very controlled way since each verb can be tested both with and without DOM.

1.3 Overview of the chapters

Chapter 2 gives an overview of the phenomenon of DOM in Romance languages. It describes the spread and the main triggering factors put forward in the literature. A special focus lies on verb classes that are reported to show a preference for the marker in Spanish, Sicilian and Catalan. Spanish is the point of departure since it is the Romance language for which a comprehensive account in terms of role semantics has been proposed. The choice of Sicilian and Catalan has been made for the following reasons: first, both languages lack a systematic analysis of verbal factors that are reported to influence DOM in general. Second, as DOM is not grammaticalized to the same extent in Catalan and Sicilian as it is in Spanish, influencing factors such as agentivity are expected to be more easily detectable.

Chapter 3 develops the idea further of classifying DOM as a linking phenomenon, proposed by García García (2007, 2014) as an explanation for the systematic occurrences of *a*-marked inanimate direct objects for certain predicate classes in Spanish. The role-semantic view on DOM adopted in the present work is based on Dowty's (1991) proto-role model, Primus' modification of it (1999a, b, 2006) and Blume's linking model (1998, 2000). As argued in these models, the distribution of proto-agent and proto-patient properties for co-arguments of a given predicate, here termed *agent-patient asymmetry* or *thematic distinctness*, serves as a basis for predicting argument and case selection. The architecture of Blume's proto-role model, which aims at formulating a linking principle for marked case frames, will turn out to be especially helpful for the investigation of DOM with animate objects.

Chapter 4 develops semantic diagnostics in order to identify the communicated agentivity for the object. It proposes that, depending on the general type of conveyed agentivity (*verbal entailment*, *presupposition*, *generalized conversational implicature* or, in the lack of all three, *unspecified for agentivity*), an agentivity scale can be established. Based on this scale, which can be partially cross-classified with the agent-patient asymmetry between subject and object, a distinction of DOM-sensitive verbs and verb classes attested for Romance languages can be made. I put forward the hypothesis that the more a direct object is set on the top of the agentivity scale, the more likely it is to be differentially *a*-marked.

Chapter 5 tests the hypothesis for Sicilian and Catalan. According to the agentivity scale, four verb classes have been formulated that show different types of communicated agentivity for the object (*entailed agentivity*, *presupposed agentivity*, *potential agentivity*, *unspecified for agentivity*). These verb classes have been tested in dependence of DOM in an acceptability judgement task carried out for one variety of Western Sicilian and two subvarieties of Central Catalan. For Sicilian, the nominal factor definiteness has been included as a further variable. The results of both studies reveal clear differences of acceptability between the four verb classes that partially confirm the predictions based on the agentivity scale. In particular verbs that encode agentivity as a verbal entailment in their lexical semantics suggest an impact of agentivity on DOM in both Sicilian and Catalan.

Chapter 6 aims at discussing the role of agentivity for DOM in Romance on the basis of the results obtained from the acceptability judgement studies for Sicilian and Catalan. The findings are compared to previous research on role-semantic factors for DOM in Spanish. A further point of reference concerns other Romance languages exhibiting DOM, where diachronic and synchronic evidence from verb classes suggests that a study comparable to the one conducted here could be fruitful. In addition, the interaction of agentivity with nominal factors of the object NP such as animacy and definiteness is discussed. Finally, I outline how agentivity could be integrated into a model on DOM in Romance languages.

Chapter 7 summarizes the main empirical findings and theoretical implications of this study and discusses them in light of the general motivations provided for the

phenomenon of DOM in the literature. Furthermore, my proposal is related to prototypical semantic transitivity and compared to other scalar verb-based approaches proposed in the typological literature.

2. DOM variation in Romance

In Romance languages, the phenomenon of DOM is widespread and well documented (cf. Bossong 1998 for an overview). In addition to the prime examples of Spanish and Romanian, DOM is also found in Catalan, Corsican, Sardinian, Sicilian and a range of other Southern Italian dialects, just to mention a few (cf. Bossong 2008: 286ff.). In almost all of these languages and varieties, the differential object marker is *a*, which has evolved from the Latin spatial preposition *ad* 'to'. An exception is Romanian, where DOM is expressed by the marker *pe*, which can be traced back to the Latin preposition *per* 'through' (Bossong 1991: 157). While Romance languages show great divergence with regard to the extent to which DOM has grammaticalized, they are broadly consistent in the set of semantic, pragmatic and information-structural factors that favour or trigger the marker: These are, first and foremost, the nominal factors of animacy and definiteness as well as the information-structural factor topicality, with the latter typically correlating with a peripheral syntactic position of the direct object. Section 2.1 delineates the mentioned nominal factors and takes a brief view on topicality. Section 2.2 presents an overview of verbal factors for the relevant languages of this study, that is Spanish, Sicilian and Catalan.

2.1 Nominal factors (and topicality)

The two main nominal factors of the object that can be shown to influence DOM cross-linguistically can be assigned to two dimensions: one is the *domain of inherence* including the object's animacy, the other is the *domain of reference* pointing to the object's individuation or discourse-related definiteness (cf. Bossong 1991: 158ff.). Both factors are typically scalarly captured. As for inherent factors, the first formulation of an animacy hierarchy goes back to Silverstein (1976) who found a connection between split ergativity in Australian languages and the degree of animacy of the subject participant. Accordingly, animate subjects in these languages more often receive nominative/accusative case marking, whereas inanimate ones are more likely to receive ergative case marking. Concerning definiteness, it is typically the NP types that are ranked, ranging from (deictic) pronouns and proper names to discourse-related definite and indefinite NPs. A further distinction is mostly made between specific and non-specific indefinite NPs. Here, I will adopt versions of the animacy and definiteness scale as used in Aissen's (2003: 437) model on DOM, displayed in (1) and (2). Note that both scales are closely connected and often cross-classified for the description of DOM patterns (cf. von Heusinger/Kaiser 2005: 40).

(1) **Animacy scale**

Human > Animate > Inanimate

(2) **Definiteness scale**

Personal pronoun > Proper name > Definite NP > Indefinite specific NP > Non-specific NP

The scales are commonly used to describe the cut-off point for DOM in a language. This is illustrated for Sicilian and Catalan, which will be of particular interest in the course of this study. In Sicilian, strong personal pronouns (3a) and personal names (3b) are obligatorily marked (Guardiano 2000; Iemmolo 2010b, in preparation). As for kinship terms, the marker is rather optional, as displayed by (3c) from the variety of Southeast Metafonetica (province of Ragusa) (Guardiano 2000: 23).[1]

(3) **Sicilian**

a. *nun vitti* ****ø/a*** *ttia*

NEG see.PST.1SG ø/DOM 2SG

'I did not see you.'

b. *vitti* ****ø/a*** *Giovanni*

see.PST.1SG ø/DOM PN

'I saw Giovanni.'

c. *ajeri 'ncuntrai* ***ø/'o*** *tò cucinu*

yesterday meet.PST.1SG ø/DOM.the POSS.2SG cousin

'Yesterday, I met your cousin.' (Guardiano 2000: 20-23)

In Standard Catalan, the *a*-marker is compulsory only for strong pronouns (GIEC 2016). This is illustrated by (4a) where the third person pronoun *ell* 'he' requires *a* whereas the personal name *la Maria* stays unmarked. Definite human NPs, here accompanied by the possessive article (4b), do not get the marker either as far as prescriptive grammar is concerned.

(4) **Standard Catalan**

a. *Només va invitar* ***ø*** *la Maria i* ****ø/a*** *ell.*

only have.PST.3SG invite.INF ø the PN and ø/DOM 3SG

'S/he only invited Mary and him.' (Escandell-Vidal 2009: 838)

b. *He encontrat* ***ø*** *el teu germà.*

have.PRS.1SG meet.PTCP ø the POSS.2SG brother

'I met your brother.' (Bossong 1998: 225)

There are two main hypotheses put forward in the literature to explain DOM systems. According to the first, DOM has a *distinguishing function* serving to disambiguate the direct object from the subject of the sentence (cf. e.g. Bossong 1985; Aissen

[1] Based on data from Western Sicilian (provinces of Agrigento and Palermo), Iemmolo (2010b) by contrast claims DOM to be obligatory with singular and optional with plural kinship terms.

2003). This is particularly necessary in configurations of "unnatural" transitivity, that is when the object bears typical properties of a subject such as animacy and definiteness (cf. Comrie 1989: 128). The second one is the *indexing function* assuming that languages formally mark such direct objects that appear in sentences with a high degree of transitivity (cf. e.g. Hopper/Thompson 1980; Næss 2004). Hopper/Thompson (1980: 252) cite various semantic parameters that – clustered together – indicate a high degree of transitivity of a sentence. On the one hand, these can be attributed to the verb (e.g. agency and affectedness) and on the other hand, they concern the properties of the direct object such as its degree of individuation. The term *individuation* can be understood to include the semantic properties of animacy, definiteness and specificity. Thus, although the two explanations are contrary to each other, they basically make the same predictions, namely the morphosyntactic marking of highly animate and individuated direct objects.[2]

In addition to the two nominal factors mentioned, the information-structural factor of topicality is often discussed as influencing DOM in Romance languages both synchronically and diachronically (cf. Iemmolo 2010a; Dalrymple/Nikolaeva 2011; Laca 1995; Melis 1995; Pensado 1995c). Correspondingly, besides the distinguishing and the indexing approach, the *topicality hypothesis* has been proposed as a third general explanation for DOM systems to evolve (Iemmolo in preparation: 41-47). For the purposes of this study, it suffices to keep in mind the bipartite hierarchy between topical and non-topical direct objects in (5) of which the former are predicted to take the marker more readily.

(5) **Topicality**

topical DO > non-topical DO

The following four pairs of examples from Spanish (6), Sicilian (7), Balearic Catalan (8) and colloquial Standard Italian (9) show that the *a*-marker appears with topicalized direct objects, syntactically reflected by their left- or right-dislocation, rather than with the same objects in canonical postverbal position.

(6) **Spanish**

a. *Ya* *conocía* ***ø/a*** *muchos* *estudiantes.*
already know.PST.1SG ø/DOM many students
'I already knew many students.'

b. ****ø/A*** *muchos* *estudiantes,* *ya* *los* *conocía.*
ø/DOM many students already CL.3PL know.PST.1SG
'Many students I already knew.' (Leonetti 2004: 86)

[2] For a detailed overview of both functions cf. García García (2014: 10-25) and Iemmolo (in preparation: 25-41).

(7) **Sicilian**

a. *Arrubbaru* ***ø/a*** *i* *so* *cuscini* *aieri*
kidnap.PRF.3PL ø/DOM the POSS.3PL cousins yesterday
'They kidnapped his cousins yesterday.'

b. ****ø/A*** *i* *so* *cuscini* *l'* *arrubbaru* *aieri*
ø/DOM the POSS.3PL cousins CL.3PL kidnap.PRF.3PL yesterday
'His cousins were kidnapped yesterday.' (Iemmolo in preparation: 260)

(8) **Balearic Catalan**

a. *Coneix* ***ø/*a*** *en* *Joan* / ***ø/*a*** *aquesta* *comedia.*
know.PRS.3SG ø/DOM OM PN / Ø/DOM this comedy
'S/he knows John / this comedy.'

b. *El* *coneix,* ****ø/a*** *en* *Joan.*
CL.M.3SG.ACC know.PRS.3SG ø/DOM OM PN
'S/he does know John.'

c. *La* *coneix,* ****ø/a*** *aquesta* *comedia.*
CL.F.3SG.ACC know.PRS.3SG ø/DOM this comedy
'S/he does know this comedy.' (Escandell-Vidal 2009: 855)

(9) **Colloquial Standard Italian**

a. *La* *soluzione* *ha* *soddisfatto* *solo* ***ø/*a*** *noi.*
the solution have.PRS.3SG satisfy.PTCP only ø/DOM 1PL
'The solution satisfied only us.'

b. ***?ø/A*** *noi,* *la* *soluzione* *non* *ci* *ha* *soddisfatti.*
ø/DOM 1PL the solution NEG CL.1PL have.PRS.3SG satisfy.PTCP
'We were not satisfied with the solution.' (Benincà 1986: 231f.)

While for indefinite human NPs in Spanish and plural kinship terms in Sicilian, *a*-marking is optional in postverbal position (6a, 7a), it gets obligatory once the direct object is left-dislocated (6b, 7b). For Balearic Catalan, DOM is ungrammatical in canonical object position both with human and inanimate definite NPs (8a), while it is compulsory in right-detached position (8b-c). In colloquial Standard Italian, likewise, DOM is rejected in postverbal position (9a) but is strongly preferred with a left-

dislocated direct object (9b).[3] As can be furthermore seen by all (b) sentences and (8c), the topicalized variant typically goes along with the insertion of a clause-internal resumptive clitic.

To conclude the brief overview, Table 1 displays the synchronic cut-off points for DOM for a number of Romance languages and varieties according to the literature (cf. references in Table 2 below). For the purposes of this study, besides Spanish, Sicilian and Catalan are of particular interest. As can be seen, of these three languages, Spanish has developed DOM to the highest extent showing obligatory marking of all definite human NPs. In Sicilian, by contrast, DOM with both definite and indefinite human NPs is generally optional. Regarding Catalan, the extension of *a*-marking is highly dependent upon the variety in question: in (spoken) Central Catalan for example, which I will focus on, only strong personal pronouns are compulsorily marked, whereas for proper names, definite human NPs and, to a lesser degree, indefinite human NPs, DOM has optional status. Note that the table is simplified and does not claim to be complete. As for the NP types, kinship terms as well as non-human animates have not been included. Concerning the languages, varieties of Spanish are not taken into account since this work is solely considered with European (Standard) Spanish.

[3] Here also the type of verb matters, as will be further addressed in sections 4.3, 6.1.3 and 6.2.3.

	personal pronoun 1/2	3	proper name	human definite NP	human indefinite NP
Europ. Spanish	a	a	a	a	a/ø
Romanian	pe	pe	pe	pe/ø[4]	pe/ø
Sardinian	a	a	a	a/ø	a/ø
Sicilian	a	a	a	a/ø	a/ø
Neapolitan	a	a	a	a/ø	a/ø
Calabrian	a	a	a	a/ø	a/ø
Ariellese	a	ø	ø	ø	ø
Standard Catalan	a	a	ø	ø	ø
Central	a	a	a/ø	a/ø	?a/ø
Valencian	a	a	a	a	?a/ø
Balearic non-detached	a	a	ø	ø	ø
detached[5]	a	a	a	a	a/ø
Languedoc	a	a	a	a/ø	?a/ø
Gascon	a	a	a	a/ø	?a/ø
Engadinian	a	a	a	a/ø	?a/ø
Corsican	a	a	a	ø	ø
Portuguese	a/ø	a/ø	ø	ø	ø
Galician	a	a	a	a/ø	ø
Asturian	a	a	a	a/ø	ø
Italian non-detached	ø	ø	ø	ø	ø
detached (colloquial)	a/ø	a/ø	ø	ø	ø
French non-detached	ø	ø	ø	ø	ø
detached (colloquial)	a/ø	a/ø	ø	ø	ø

Table 1. Cut-off points for DOM in Romance languages and varieties

[4] In Romanian, unmodified definite human NPs can be either accompanied by *pe* or by the enclitic definite article (cf. Chiriacescu/von Heusinger 2011).
[5] In Balearic, detached definite inanimate NPs are compulsorily marked while for detached indefinite inanimate NPs the *a*-marker is optional.

Language	References (selection)
Europ. Spanish	von Heusinger 2008; Aissen 2003
Romanian	Maiden 2016; Pană Dindelegan 2013
Sardinian	Jones 1995, 2003; Putzu 2008; Mensching/Remberger 2016
Sicilian	Iemmolo 2010b, in preparation; Guardiano 2000
Neapolitan	Ledgeway 2016b
Calabrian	Bossong 1998; Ledgeway et al. 2019
Ariellese	D'Alessandro 2017
Catalan (Standard, Central, Valencian, Balearic)	Hualde 1992; Sancho Cremades 1995; Escandell-Vidal 2009
Occitan (Languedoc/Gascon)	Bossong 1998
Raeto-Romance (Engadinian)	Bossong 1991, 1998
Corsican	Neuburger/Stark 2014; Ledgeway 2016a
Portuguese	Bossong 1998; Döhla 2014; Dubert/Galves 2016
Galician	Dubert/Galves 2016
Asturian	Prieto Entrialgo 2010
Italian	Bossong 1998; Berretta 1991
French	Bossong 1998; Fagard/Mardale 2014

Table 2. References on synchronic cut-off points for DOM in Romance languages and varieties (as displayed in Table 1)

2.2 Verbal factors

Since the focus of this study is on the impact of verbal factors on DOM, this section provides an overview of DOM-sensitive verb classes reported for the Romance languages in question, that is Spanish (2.2.1), Sicilian (2.2.2) and Catalan (2.2.3). While for Spanish, verbal factors have been taken into account more systematically in research of recent decades, similar studies are still missing for the latter two languages.

2.2.1 Spanish

Verbal factors have been considered relevant for DOM in Spanish already in early accounts (i.a. Fernández Ramírez 1986; Müller 1971; Roegiest 1979). Fernández Ramírez (1986: 159) for example states that verbs typically selecting 'real beings' as

objects, such as *favorecer* 'to favour', *impresionar* 'to impress', *inviter* 'to invite', *proteger* 'to protect', *saludar* 'to greet' or *matar* 'to kill', show a preference for the *a*-marker. He further notes that while verbs predominantly taking personal, i.e. human, direct objects prefer DOM, this does not seem to be a sufficient criterion:

> Entre los que llevan *a* predominan los que se construyen siempre o casi siempre con complemento de persona. Pero no todos los de esta clase adoptan la preposición. Encontramos sin *a* verbos tan "personales" como *convidar*, *curar*, *desnudar*, *engendrar*, *evangelizar*, *libertar*, *matar*, *raptar*, *secuestrar*, *sobornar*, *tratar* (3 pasajes) y con *a* verbos no más "personales" que los anteriores: *admirar*, *contaminar*, *convencer*, *pacificar*, *reclutar*, *satisfacer*. (Fernández Ramírez 1986: 168)

In contrast, verbs like *buscar* 'to look for', *encontrar* 'to find', *hacer* 'to make', *llevar* 'to carry', *necesitar* 'to need', *querer* 'to want' or *tener* 'to have', that take a 'mental object', occur less frequently with the marker (Fernández Ramírez 1986: 159, 168).

A number of later accounts then focused on verbal factors related to high transitivity, such as affectedness and telicity, as worked out in Hopper/Thompsons' (1980) approach to prototypical transitivity. Torrego (1999: 1787ff.) puts forward the hypothesis that DOM is influenced by telicity. She bases her claim on the observation that telic verbs like *insultar* 'to insult', *curar* 'to cure', *sobornar* 'to bribe' and *emborrachar* 'to make drunk' have grammaticalized DOM also with indefinite object NPs. In addition, Torrego claims that the insertion of DOM with atelic verbs, i.e. Vendler's (1957) states and activities, turns atelic verbs into telic ones and thus modifies the reading of the sentence. The latter point has been criticized in the literature (cf. Delbecque 2002: 95ff.; García García 2018a: 223). The first point is called into question by a recent diachronic corpus study that systematically investigates the occurrence of telic and atelic predicates with DOM in Spanish between the 14th and 20th centuries (Romero Heredero 2021, 2022). It revealed that from a diachronic perspective, telic predicates, i.e. *achievements* and *accomplishments*, on the one hand and atelic predicates, i.e. *states* and *activities* on the other hand, do not differ with regard to their correlation with *a*-marking. In addition to telicity, affectedness has also been discussed as a factor for DOM in Spanish (cf. Pottier 1968; Torrego 1998, 1999). Von Heusinger/Kaiser (2011) were the first to empirically test the impact of affectedness on DOM in a diachronic corpus study, making use of Tsunoda's (1985) transitivity scale. What they found was a partial correlation of verbs selecting an affected object and *a*-marking: as predicted by the scale, *verbs of effective action* like *matar* 'to kill' and *herir* 'to wound' were among the first to generalize the marker with definite human NPs. Yet, contrary to the predictions, *feeling verbs* like *querer* 'to like' and *temer* 'to fear' were not the least-attested class with DOM but occupied a middle position both with definite and indefinite human NPs (von Heusinger/Kaiser 2011: 612). The two verbs of this class differ strongly: *temer* 'to fear' especially has strongly favoured DOM since the 15th century. As an explanation, the authors assume that "the direct object of 'fear' has more typical properties of a subject than a prototypical object of 'like' […] This behaviour, however, has nothing to do with affectedness, but rather with the competition between the agentitivity of the participants involved in the event" (von Heusinger/Kaiser 2011: 613). For both telicity and

affectedness, one more remark has to be made: a possible influence of each of the two factors on DOM would be always *dependent* upon the animacy of the direct object. With inanimate affected objects involved in a telic event, *a*-marking in Spanish has been shown to be excluded (cf. García García 2014: 194-201).[6]

Beyond this, research of recent decades has taken the role-semantic configuration of a sentence, i.e. the relation between verbal predicate, subject and object, into consideration (e.g. Weissenrieder 1985, 1991; Delbecque 1998, 2002; García García 2007, 2014; Primus 2012a). Though not labelling it as such, Bossong (1991: 158 ff.) already mentions the role-semantic dimension among the factors that cross-linguistically determine DOM: In addition to the object-oriented factors of *inherence* and *reference*, he introduces a third domain of *constituence*, which concerns the "relative dependence/independence of the object NP with respect to the verbal predicate" (Bossong 1991: 158). Whereas semantically prototypical objects do not exist independently of the verb and remain unmarked, "self constituent objects, i.e. objects which are independent, or autonomous, with respect to the verb, tend to be positively marked" (Bossong 1991: 158). However, since Bossong also subsumes pragmatic and syntactic (in)dependence under the notion of constituence, his account cannot be seen as a proper role-semantic approach. The assumption that the whole predication has to be considered for DOM is also a core idea of Delbecque's (2002) construction grammar approach on DOM in Spanish. She advances the hypothesis that Spanish has two transitive frames, an *a* construal and a *ø* construal. These two frames are in principle available for any given transitive verb and they are argued to involve different event structures: While the *ø* construal establishes a dependency relation between subject and direct object, the *a* construal sets up a relation of interdependency between the two arguments, which means that both participants are ascribed the semantic role of initiator and goal of the event at the same time. The only verb class that, in Delbecque's view, is limited to the *a* construal is the one of "goal centered affection" (2002: 104f.) comprising *object-experiencer psych verbs* (OE-psych verbs) like *asustar* 'to frighten' or *fascinar* 'to fascinate'. The confinement of these verbs to the *a* frame is explained by the event-structural peculiarity that "the source–goal relationship is as much centered on the human entity affected as on the causal instance that triggers the affective process" (cf. Delbecque: 105).

A further role-semantic account on DOM in Spanish operates with generalized semantic roles: in his account on DOM with inanimate objects in Spanish, García García (2007, 2014) works within the framework of Dowty's (1991) proto-role model and Primus' (1999a, b, 2006) modification of it. He shows that based on the thematic distinctness of subject and direct object in a sentence, predictions can be made on the presence of the *a*-marker with inanimate objects. Accordingly, the marker is said to appear in cases where the subject does not outrank the object in terms of agentivity (García García 2007: 71, 2014: 145). In this way, the behaviour of verb classes that

[6] For a detailed corpus analysis on Spanish DOM that takes into account the interplay of the predicate's degree of transitivity in the sense of Hopper/Thompson (1980) and the conceptualization of the object, cf. Egetenmeyer (2019: 279-423).

systematically select for DOM with inanimate objects such as reversible-symmetrical predicates (e.g. *preceder* 'to precede', *seguir* 'to follow') can be explained. Note that these verbs are of an extremely low degree of semantic transitivity. Later work by Primus (2012a) attempts to apply the same kind of role-semantic explanation to animate objects, also comprising the ones that are involved in a highly transitive event. Here, it is claimed that animate objects get marked due to their *potential* to function as agents in the same event. This study will elaborate further on the latter point.

2.2.2 Sicilian

Although DOM in Sicilian has been studied both diachronically and synchronically (Rohlfs 1971, 1973; Guardiano 2000; La Fauci 1991; Romagno 2007; Iemmolo 2009, 2010a, 2010b, in preparation; Sornicola 1997, 2000, 2011; Volo 2010), the impact of verbal factors has only been scarcely investigated.

Just as for Spanish, telicity has been proposed as a factor that favours DOM in Sicilian (cf. e.g. Romagno 2007). However, as shown by a diachronic corpus study conducted by Iemmolo (2010a), telicity correlates negatively with the occurrence of *a*-marking. The same observation is made for Modern Sicilian (Iemmolo 2010a, b). Especially the claim put forward by Romagno (2007) that telic verbs (with highly affected objects) such as *ammazzari* 'to kill' show a particular preference for DOM is rejected. Guardiano (2000: 33) likewise notes that her data do not suggest an influence of affectedness on DOM in Sicilian but she does not elaborate on this point. Beyond this, at another point in her article she hypothesizes that the high frequency of the *a*-marker with deity names is due to verb type: deity names often co-occur with verbs like *pregare* 'to pray' or *ringraziare* 'to thank' which are similar to verbs like *chiedere* 'to ask' or *dire grazie* 'to say thank you'. Unlike the former, the latter verbs are constructed intransitively and thus always take *a*. In her line of argumentation, speakers might make a transfer from the intransitive to the transitive construction.

Concerning the role-semantic factor agentivity, evidence on DOM in Modern Sicilian is lacking. There are some clues, though, that in Old Sicilian DOM was favoured with verbs selecting an agent-like object, such as the ones listed in (10) (cf. La Fauci 1991: 391-393; Sornicola 1997: 72f.):

(10) Verbs taking DOM in Old Sicilian (13th-15th century)

adurari 'to adore', *aiutari* 'to help', *amari* 'to love', *ascultari* 'to listen to' *consolari* 'to console', *cunfurtari* 'to confort', *cunsiglari* 'to deliver', *cuntraddiri* 'to contradict', *firiri* 'to injure', *ingannari* 'to deceive', *offendiri* 'to offend', *honorari* 'to honour', *iastimari* 'to curse', *inculpari* 'to accuse', *inflamari* 'to inflame', *iniuriari* 'to injure', *laudari* 'to praise', *perdunari* 'to forgive', *prigari* 'to pray', *serviri* 'to serve'

In La Fauci's (1991) argumentation, DOM correlates with verbs that are thematically "dative". In any case, what the above-listed verbs have in common is that they select for a non-prototypical agent-like object. Yet these indications have neither been systematically proved in an empirical study on Old Sicilian nor have the verbs been tested for their DOM preference in Modern Sicilian.

2.2.3 Catalan

DOM in Catalan has been studied both from a synchronic and diachronic perspective (Hualde 1992; Solà 1994; Sancho Cremades 1995, 2002; Wheeler et al. 1999; Bossong 1998, 2008; Bel 2002; Aissen 2003; Escandell-Vidal 2009; Iemmolo 2010a, in preparation; Alsina 2016; Pineda 2021, in press). To begin with, it must be pointed out that when the literature refers to DOM in *Catalan*, it usually means *Standard Catalan*. In this variety, according to prescriptive grammar, DOM only occurs with personal pronouns. In other varieties, such as Valencian or (spoken) Central Catalan, the phenomenon is more widespread (cf. Table 1 in section 2.1). However, for all varieties there is only very little evidence of a potential impact of verbal factors on DOM.

Yet there are two exceptional cases in which DOM is mentioned to be compulsory in Standard Catalan, even if the object is no strong pronoun. These are (i) reciprocal constructions and (ii) comparative structures with an elided verb (cf. GIEC 2016; Pineda 2021: 248; Bossong 2008: 115). Crucially, in both cases DOM serves as a means of disambiguation, since subject and object cannot be clearly identified in the given events. Examples for both constructions are given in (11) and (12).

(11) **Reciprocal constructions**

a. *Ens* *miràvem* *l'* *un* ***a*** *l'* *altre*
REFL.1PL look at.IPRF.1PL the one DOM the other
'We looked at each other.' (Bossong 2008: 115)

b. *S'* *abraçaven* *l'* *un* ****ø/a*** *l'* *altre.*
REFL.3 hug.IPRF.3PL the one ø/DOM the other
'They were hugging each other.' (Pineda 2021: 248)

(12) **Comparative structures (elided verb)**

El *perseguia* *com* *el* *gat* ****ø/a*** *la* *rata*
CL.M.3SG.ACC chase.IPRF.3SG like the cat ø/DOM the rat
'He chased him like the cat (chases) the rat.'
(Pineda 2021: 248; cf. also Bossong 2008: 115)

In the reciprocal constructions in (11a) and (11b) the insertion of DOM guarantees the distinction between subject and direct object. Similarly, in the comparative structure with elided verb in (12), DOM clarifies the argument relation between the subject *el gat* 'the cat' and the direct object *la rata* 'the rat'. Thus, in these two cases the requirement of DOM can be explained by the role-semantic configuration. Beyond this, there are instances of *a*-marked objects of OE-psych verbs that are of interest. Royo (2017: 68) shows that for the transitive verb *sorprendre* 'to surprise', DOM

does not occur if the object is in canonical postverbal position (13a). If it is left-dislocated as in (13b), though, the *a*-marker does appear. Note that in this case both a direct (*el*) and an indirect object clitic (*li*) are possible. Yet the fact that DOM does not occur with *sorprendre* in (13a) suggests that we indeed are dealing with a transitive and not an intransitive verb.

(13) a. *La Maria sorprèn* ***ø*** *l' Arnau.*
the PN surprise.PRS.3SG ø the PN
'Maria surprises Arnau.'

b. ***A*** *l' Arnau* ***el/li*** *sorprèn*
DOM the PN CL.M.3SG.ACC/CL.3SG.DAT surprise.PRS.3SG
l' actitud de la Maria.
the attitude of the PN
'Arnau is surprised by Maria's attitude.' (Royo 2017: 68)

It is important to notice that the role-semantic configuration differs in both sentences. While (13a) can have both a causative and a non-causative reading, the sentence in (13b) only has a non-causative reading. That is, the co-arguments in the latter case are clearly less distinguishable than in the former one. This is supported by examples from Pineda/Royo (2017: 457) with the OE-psych verb *commoure* 'to move'. In (14a), the verb can have a causative reading and the *a*-marker is optional. In (14b), the verb exclusively has a non-causative reading and *a*-marking is compulsory. In this case, the use of solely the *direct* object clitic indicates the transitivity of the verb.

(14) a. ***ø/A*** *la Maria(,)* *l'* *han* *commogut* *amb*
ø/DOM the PN CL.F.3SG.ACC have.PRS.3PL move.PTCP with
mostres d' afecte
displays of affection
'They moved Maria with displays of affection.'

b. ****ø/A*** *la Maria(*,)* ***la*** *commou* *el* *record*
ø/DOM the PN CL.F.3SG.ACC move.PRS.3SG the memory
d' aquell dia
of that day
'Maria is moved by the memory of that day.' (Pineda/Royo 2017: 457)

However, the authors explicitly state that although the object *La Maria* in (14a) and (14b) "is coindexed with an accusative clitic, the *a* is a dative marker […], and not an instance of DOM" (Pineda/Royo 2017: 457, fn. 18).

All in all, despite the scarce evidence, the examples shown suggest that in certain cases, role-semantic factors can favour or require the occurrence of DOM in Catalan. Certainly, this claim needs to be proved in an empirical study where transitive and intransitive verbs have to be clearly dissociated.

2.3 Summary

This chapter provided an overview of DOM in Romance languages and outlined the main triggering factors discussed in the literature. It has been stressed that, although the phenomenon is generally widespread across Romance and seems to be influenced by the same set of nominal factors, it manifests itself to different degrees in the individual languages and varieties. As has been illustrated, the animacy and definiteness scale serve as helpful descriptive and comparative tools to capture the cut-off point of DOM in a language. Furthermore, the information-structural factor topicality can account for the fact that in some languages, the marker occurs solely or preferably in dislocated structures. Unlike nominal factors, verbal factors still lack a systematic investigation. This applies in particular to the two languages relevant here, Sicilian and Catalan, for which only sporadic observations on DOM-sensitive verbs and verb classes have been made. For Spanish DOM, studies of recent decades have revealed fruitful insights on verb-based properties that seem to correlate with *a*-marking. One main account argues that the role-semantic configuration between subject and direct object affects DOM. The present study will pursue this point further and examine DOM within a framework of generalized semantic roles.

3. (Proto-)Agentivity, agent-patient asymmetry and (morpho)syntactic argument realization

The role-semantic view on DOM adopted in this work is based on Dowty's (1991) proto-role model, Primus' modification of it (1999a, b, 2006) and Blume's linking model (1998, 2000). As argued in these models, the distribution of proto-agent and proto-patient properties for a given predicate, here termed *agent-patient asymmetry* or *thematic distinctness*, serves as a basis for predicting argument and case selection. That such an explanation is also well suited to account for DOM has been demonstrated by García García (2007, 2014), who puts forward the hypothesis that systematic occurrences of DOM with inanimate objects in Spanish are determined by the thematic distinctness of subject and direct object.

In the following, a comprehensive overview of the impact of role semantics on different dimensions of argument realization phenomena will be provided. First, I will introduce Dowty's (1991) proto-role model, which focusses on syntactic argument selection in transitive sentences (3.1). Then, modifications of the model (Primus 1999a, b, 2006; Blume 1998, 2000), which extend its applicability to morphosyntactic argument realization, are discussed (3.2). Finally, on the basis of García García (2007, 2014) I will show how the proto-role model serves to explain differential morphosyntactic argument realization (3.3).

3.1 Syntactic argument realization in Dowty's (1991) proto-role model

Dowty's (1991) approach suggests a solution to the shortcomings of traditional semantic role lists (cf. Levin/Rappaport Hovav 2005: 38-48). In these traditional views, as for example stated by Fillmore (1968), semantic roles such as agent, patient, theme or experiencer are understood as discrete and unanalysable notions. This means that there are no common properties assumed among the different roles that would allow further subcategorization. Yet from a cross-linguistic perspective on argument realization, this runs into the problem that certain semantic roles show the same encoding patterns, contradicting the assumption originally put forward in Fillmore's (1968) Case Grammar that there is a one-to-one correspondence of a particular semantic role and its syntactic expression (Levin/Rappaport Hovav 2005: 41f.). Furthermore, it is problematic to assume that each argument can be attributed to exactly one distinct semantic role and that each semantic role is only distributed once for a given predicate. A violation of the first point can be illustrated by verbs of transfer of possession: In the sentence *Phil sold the yacht to Mira* the subject *Phil* is both source and agent, while *Mira* is goal and *the yacht* theme (Levin/Rappaport Hovav 2005: 42f.). Note that also with object-experiencer psych verbs such as *to frighten* in *The ghost frightened Peter* the question arises whether the direct object *Peter* only functions as an experiencer or if it could just as well be analyzed as a patient being affected by the subject stimulus. Similarly, besides being analyzed as a stimulus, the subject *the ghost* can be ascribed the role of an agent, too. The second point is a problem with symmetrical predicates: With verbs like *to resemble*, subject and direct object can

swap places without changing the truth-conditions of the sentence. This is illustrated by the examples in (1), taken from Levin/Rappaport Hovav (2005: 43):

(1) a. Pat resembles Lee.

b. Lee resembles Pat.

The interchangeability of subject and object without any change in meaning suggests that the two arguments are assigned to the same semantic role by the predicate.[1] As will be shown later, both object-experiencer psych verbs and symmetrical predicates can be considered core cases for a role-semantic account on differential morphosyntactic argument realization (cf. sections 3.3, 4.3.1).

In order to overcome these issues, in his model Dowty (1991) elaborates the concept of two *generalized semantic roles*, a label later coined by Van Valin (1999). The two cluster roles, the proto-agent and the proto-patient role, are characterized by the following properties (Dowty 1991: 572):

(2) Contributing properties for the Agent Proto-Role:

a. volitional involvement in the event or state

b. sentience (and/or perception)

c. causing an event or change of state in another participant

d. movement (relative to the position of another participant)

(e. exists independently of the event named by the verb)

(3) Contributing properties for the Patient Proto-Role:

a. undergoes change of state

b. incremental theme

c. causally affected by another participant

d. stationary relative to movement of another participant

(e. does not exist independently of the event, or not at all)

The proto-properties are understood as entailments of the predicate in the strictly logical sense.[2] Dowty situates his approach within prototype theory (Rosch 1973),

[1] However, there might be a discourse-related difference between (1a) and (1b) in that the object expresses the standard of comparison and the subject expresses the entity that is compared to this standard (cf. Levin/Rappaport Hovav 2005: 43).

[2] These entailments are encoded in the lexical meaning of a predicate. This is also the reason why they "survive" negation or other operations at the sentence level, as pointed out by Beavers (2010: 831): "The fact that these entailments are assigned in the lexicon means that by the time a verb enters the syntax, operations that cancel or change the entailments (e.g. negation, modal operators, quantifiers) generally have no effect on argument realization, which is predetermined. Thus, in *John did not eat the sandwich*, technically no properties need hold of either argument, but the argument realization

which is supported by the following characteristics. As he stresses, the two lists are neither complete nor are the properties themselves clearly delimited, indeed, at least some of the listed entailments have "unclear boundaries" (1991: 598). That is also a reason why Dowty rejects treating them as semantic features in a decompositional approach. Such a model would suggest not only a definite set of entailments but also an equal distribution of features. For the proto-properties listed, although they give the impression of being of equal status, their internal ranking or weighting is uncertain. Dowty exemplifies this with the proto-agent properties of (2c) *causation* and (2d) *movement*: The latter is only treated as a proto-agent property when it is not caused by another participant, as for the subject in *Water filled the tank*, while caused movement, as for the object in *John threw the ball*, is a proto-patient property. From this, Dowty concludes "causation has priority over movement for distinguishing agents from patients" (1991: 574). Working with the cluster concepts of proto-agent and proto-patient, Dowty avoids defining his notions in necessary and sufficient conditions. Rather, there are more or less prototypical instances of an agent and patient while there is "no invariant entailment or set of entailments which determines access to subjecthood or objecthood" (Levin/Rappaport Hovav 2005: 54). The following examples show that each of the proto-agent properties can appear as separate entailments:

(4) Examples illustrating independence of Proto-Agent entailments (in subject NPs):

a. VOLITION ALONE: *John is being polite to Bill/ is ignoring Mary.*

b. SENTIENCE/PERCEPTION ALONE: *John knows/believes/is disappointed at the statement.*

c. CAUSATION ALONE: *His loneliness causes his unhappiness.*

d. MOVEMENT ALONE: *The rolling tumbleweed passed the rock.*

e. INDEPENDENT EXISTENCE: *John needs a new car.*

(Dowty 1991: 572f.)

By *volition*, Dowty means both deliberate action and deliberate refraining from action. However, identifying *volition* as a verbal entailment is not straightforward, as becomes evident when we consider Engelberg (2006: 270) who discusses volitionality as a licensing condition for the impersonal passive in German. He remarks that "volitionality is only very rarely lexically entailed at all", offering evidence from activity verbs like *draw*, *scribble*, *tap* or *walk*, for which the entailment can be easily cancelled in a context where the action is carried out unconsciously or unknowingly.

The property of *sentience*, which will be essential for this work, refers to cognitive, perceptive or emotional involvement of a participant. It is defined as follows: "Sentience means more than a presupposition that an argument is a sentient being; it is

facts are fixed prior to negation just the same as they are for *John ate the sandwich*". For a similar point cf. also Primus (1999a: 36f.).

rather sentience with respect to the event or state denoted by the verb: the objects of verbs like *elect*, *appoint*, *nominate* and *idolize*, *venerate* and *convict*, *acquit*, *exculpate* are necessarily human but are not entailed to know or perceive the relevant event" (Dowty 1991: 573). Dowty adds in a later passage of his paper that "volition necessarily involves sentience" (1991: 607). This implication is only unilaterally valid since sentient involvement in an event can be without volition. One further remark concerning the relationship between *sentience* and animacy has to be made. As the definition above clearly states, the animacy of a participant does not imply its sentience in a given event. Yet it is debatable whether the reverse holds true, namely that *sentience* entails the animacy of a participant as postulated by Blume (2000: 137) and García García (2014: 113, 130f.). Primus (1999a: 50f.) likewise assumes that *sentience* – similar to *control*, a notion involving *volition*, though defined more broadly – implies animacy at least in prototypical cases. Possible exceptions are "impersonal agents and experiencers such as institutions (e.g. *the church/the jury knew/reported*) and other conventionalized metaphors and metonymies such as in *this book discusses ergative languages*" (Primus 1999a: 50f.). Dowty generally does not mention possible implicational relations between his proto-agent properties and animacy. In the case of sentience, however, his following remark is of importance:

> […] the boundary of sentience is clouded by cases of computers or intelligent animals doing certain actions or being in certain states that are stereotypically reserved for human, sentient participants, and such cases are reflected linguistically in *The machine switched itself off* (Cruse 1973:21), *The dog believed you were a stranger*, *The program did that because it thinks you haven't saved the file first*, etc. (Dowty 1991: 574)

The predicates used in the examples – at least *to believe* and *to think* – entail a sentient subject participant. This role is typically filled by the prototypical sentient being, namely a human being. However, as can be inferred from Dowty's remark, this does not mean that *sentience* implies animacy or humanness. Otherwise, we would expect a mismatch in cases such as *the program thinks* or *the computer calculates* which does not arise: the sentences are perfectly grammatical. One could argue that we are dealing with figurative uses here, namely the personification of machines and that we can subsume these uses together with metaphors and metonymies under 'human being'. But that misses an important point: The above-mentioned examples illustrate that *sentience* does not overlap one-to-one with human or animate beings but that the property can be ascribed to other entities, such as intelligent machines, as well. Of course, this does not mean that they have the same cognitive abilities or show, say, the same range of emotions as human beings. That becomes clear in sentences like *#the computer falls in love* or *#the program is jealous of me*, which are odd. Rather, "artificial agents […] are only capable of a limited number of specific actions, a restriction that human agents do not have" (García García/Primus/Himmelmann 2018: 32). With respect to *sentience* this is meant to express that every artificial agent can only fill a limited number of predicates entailing sentience but not all of them at the same time, the latter function is only reserved for human beings. Perhaps this comes close to what Dowty means with "sentience itself is something that different entities can have to different degrees" (1991: 574). To sum up, I understand the unclear

boundaries of the notion of *sentience* in the following way: a slot PREDICATE(x+SENTIENT) can always be filled by a human entity, the core case of a sentient being, and this can be said to hold true regardless of context. By contrast, it is not possible to previously determine the set of non-human entities that fit into the slot above. Rather, this varies from predicate to predicate, that is from context to context.

The proto-agent property *causation* appears in nearly all of the cases together with the property *movement*. Still, as noted by Dowty (1991: 573), we cannot speak of an implicational relation here since there are cases with generic subject and object NPs such as (4c) above and of stative causatives where movement is not entailed. Movement, by contrast, appears more frequently without causation, namely with all kinds of self-induced physical activity.

With regard to the property *independent existence*, Dowty is well aware of the fact that it might be different from the other four entailments, which is why he keeps it in brackets. On the one hand, it expresses a *de re* (vs. *de dicto*) reading of the noun phrase in question, given, for example, for the subject but not for the object in *John needs a new car*. Note that this dimension concerns more the semantic domain of specificity than the verbal semantics. On the other hand, the criterion expresses that a referent "is not brought into being or destroyed by the event named by the verb but is presumed to exist before and after the event" (Dowty 1991: 573). As Dowty mentions, *independent existence* can appear as a separate proto-agent property but is necessarily implied by all other proto-agent properties (4a-d). Apparently, this challenges the assumption of the model that proto-agents do not have to share a common criterion since no proto-agent without the property of independent existence can be identified.

In summary, the discussion of proto-agent properties has revealed at least two questions: The first one concerns the relationship between the different proto-agent properties. This includes both dependencies between the five properties (*volition* entails *sentience*, all properties (4a-d) entail *independent existence*) and the question of weighting. With respect to the latter aspect, we have seen that causation seems to weigh more than movement. In section 3.2.2 I will take up this point again while discussing the impact of sentience on (morpho)syntactic argument realization. The second issue concerns the relationship between individual proto-agent properties and animacy not tackled by Dowty. As has been argued by the example of *sentience*, proto-agent properties are understood as strictly relational notions that do not imply the involvement of an animate participant though this might run against one's initial intuition. Rather, a predicate slot PREDICATE(x+SENTIENT) can be filled by intelligent agents, e.g. machines as well, always depending on the context.

In contrast to the set of proto-agent entailments, the proto-patient entailments are more difficult to isolate from each other:

(5) Examples illustrating Proto-Patient entailments independently (in object NPs):

a. CHANGE OF STATE: *John made a mistake* (coming into being […]), *John moved the rock* (indefinite change of position), *John erased the error* (ceasing to exist).

b. INCREMENTAL THEME: *John crossed the driveway/filled the glass with water* (also stationary relative to other arguments).

c. CAUSALLY AFFECTED: *Smoking causes cancer.*

d. STATIONARY RELATIVE TO ANOTHER PARTICIPANT: *The bullet entered the target/overtook the arrow.*

e. EXISTENCE DEPENDENT OF THE EVENT: *John built a house/erased an error* (Coming into and out of existence; […]), *This situation constitutes a major dilemma for us, John needs a car/seeks a unicorn/lacks enough money to buy it* (de dicto objects: no existence).

(Dowty 1991: 573f.)

By *change of state*, Dowty means "coming into existence, going out of existence, and both definite and indefinite change of state" (1991: 574). Importantly, this notion not only includes a physical but also a mental change of state, that is "coming to experience an emotion or a new mental state" (1991: 580). A mental change of state appears for example in the direct object of object-experiencer psych verbs such as *to frighten* or *to surprise* in their inchoative reading. The property *incremental theme*, discussed in detail by Dowty (1991: 567-571), goes back to Krifka (1989) and refers to a specific kind of telic predicate that – in combination with an object NP that denotes a definite quantity – expresses "a homomorphism from its (structured) theme argument denotations into a (structured) domain of events" (Dowty 1991: 567). Typical examples for incremental themes are predicates with destroyed (= affected) objects, such as *destroy a presidential finding* or *eat a sandwich* and predicates with created (= effected) objects, e.g. *build a house* or *write a letter* (Dowty 1991: 568). The three other properties (5c-e) are the converse properties of the respective proto-agent properties (4c-e). As mentioned above, the proto-patient entailments are much harder to disentangle than the proto-agent properties. So, for instance *change of state* could be given only in the presence of the proto-agent property *causation* and the proto-patient property *causally affected* (cf. also Levin/Rappaport Hovav 2005: 63). Even more notable and not explicitly mentioned by Dowty, once the entailment *incremental theme* is given, *all* other proto-patient properties (and even all proto-agent properties) are necessarily implied, too.

After having defined the two cluster concepts of proto-agent and proto-patient, Dowty (1991: 576) explains how we can make predictions about argument selection based on the accumulation of proto-agent and proto-patient properties. He formulates an argument selection principle and two corollaries:

(6) ARGUMENT SELECTION PRINCIPLE (Dowty 1991: 576):

In predicates with grammatical subject and object, the argument for which the predicate entails the greatest number of Proto-Agent properties will be lexicalized as the subject of the predicate; the argument having the greatest number of Proto-Patient entailments will be lexicalized as the direct object.

> COROLLARY 1: If two arguments of a relation have (approximately) equal numbers of entailed Proto-Agent and Proto-Patient properties, then either or both may be lexicalized as the subject (and similarly for objects).
>
> COROLLARY 2: With a three-place predicate, the nonsubject argument having the greater number of entailed Proto-Patient properties will be lexicalized as the direct object and the nonsubject argument having fewer entailed Proto-Patient properties will be lexicalized as an oblique or prepositional object (and if two nonsubject arguments have approximately equal numbers of entailed P-Patient properties, either or both may be lexicalized as direct object).

Dowty understands the argument selection principle as "a constraint on what lexical predicates may exist in a natural language, out of many imaginable ones" (1991: 576). It can be thus used to make predictions on lexicalization patterns of predicates. For a verb like *to build* or *to eat*, for example, it correctly predicts that the argument bearing the greatest number of agent properties is realized as the subject and the argument having the greatest number of proto-patient properties is realized as the direct object: *Peter eats an apple*/*Maria builds a house*. Although the argument selection principle is not intended to be a linking principle, since this would not be compatible with the monostratal concept of Montague grammar, Dowty remarks that a reformulation into a proper linking principle would be theoretically possible (1991: 576, fn. 18). Dowty proves his argument selection principle only for English transitive verbs. However, he notices that the verbs his argument selection principle predicts to have the most stable lexicalization patterns correspond to the ones that are generally considered to be "primary transitive verbs" (Andrews 1985) or verbs ranking high on a scale of transitivity (Hopper/Thompson 1980), namely the following: *build (a house)*, *write (a letter)*, *murder*, *eat*, *wash (a plate)*[3] (Dowty 1991: 577).

There are two verb classes that illustrate the innovations of Dowty's (1991: 579f.; 583-585) argument selection principle: these are symmetrical predicates (e.g. *is similar to*, *is equal to*, *is like*) and lexical doublets of psychological predicates (e.g. *fear/frighten*, *like/please*). These predicates support his argument selection principle since in both cases, subject and object argument bear approximately the same number of proto-properties. Hence, in line with corollary 1, both arguments can be realized as the subject or direct object, respectively. Note that, in (8) and (9), there is no difference in truth-conditions between the (a) and (b) sentence:

(8) Symmetrical predicates

a. The dog resembles the wolf.

b. The wolf resembles the dog.

[3] However, a cross-linguistic study on "transitivity prominence" by Haspelmath (2015) reveals that at least for case selection, achievements such as 'break', 'kill' and 'tear' slightly differ from incremental theme verbs like 'eat' and 'build': in a sample of 36 languages, the former obtain a transitivity score of 1.00 while the latter only show a score of .93. Note that this difference cannot be accounted for in terms of proto-role properties since incremental theme verbs show more proto-role entailments than the mentioned achievement verbs.

(9) Psychological predicates

a. Peter feared the ghost. [subject-experiencer (SE-) psych verb]

b. The ghost frightened Peter. [object-experiencer (OE-) psych verb]

In (8), both subject and direct object bear the same proto-agent entailment, namely *independent existence*. Thus, both arguments can be either realized as subject and direct object without a change in meaning. In order to follow Dowty's argumentation for the psych verbs in (9), it seems helpful to have a closer look at the proto-properties entailed for the two arguments in (9a) and (9b), listed in detail in Table 1 below:

Verbs	**Arg1 (Subject)**		**Arg2 (Direct Object)**	
	Proto-agent	**Proto-patient**	**Proto-agent**	**Proto-patient**
(9a) *to fear* (SE-psych)	+ sentience		+ causation	
Number of proto-properties	1	0	1	0
(9b) *to frighten* (OE-psych)	+ causation		+ sentience	± change of state
Number of proto-properties	1	0	1	0-1

Table 1. Dowty's (1991: 579f.) analysis of SE- and OE-psych verbs

According to Dowty (1991: 579f.), both verbs in (9) entail *causation* to the stimulus and *sentience* to the experiencer, so that both arguments are assigned exactly one proto-agent property. This is argued to lead to lexical doublets of the *fear/frighten* type since both arguments are equally good subjects and direct objects. In the case of OE-psych verbs, the choice of construction type can be further motivated by its aspectual behaviour: With reference to Croft (1986), Dowty further points out that whereas SE-psych verbs are always stative, OE-psych verbs can have both a stative and an inchoative reading (stative: *Dogs have frightened Peter for years* vs. inchoative: *The dog frightened Peter in the street*). In the inchoative reading, the object bears the proto-patient property *change of state*. So, while according to Dowty, subject and direct object are equal in proto-agent properties, "they are unequal in that one is a 'better' Patient, so it must be the direct object according to the selection principle in [31]" (1991: 580).

There are several inconsistencies with the presented analysis. First of all, in his analysis Dowty misses his own previous point, namely that the proto-agent property *causation* necessarily implies the proto-patient property *causally affected* for the other argument (1991: 574; cf. also Kailuweit 2005: 104). This causes some problems for the class of SE-psych verbs like *to fear* which, consequently, must be assigned *causally affected* for the subject if the object entails *causation* (which is highly debatable as discussed below). As a result, the direct object would have two proto-agent

properties and no proto-patient entailment while the subject would have two proto-agent properties as well, plus one proto-patient property (cf. Table 2). If argument selection was dependent on counting of proto-properties in a strict sense, we would expect the actual subject to be realized as the direct object. Similarly, also for OE-psych verbs the proto-patient entailment *causally affected* has to be assigned to the object experiencer. Furthermore, Dowty does not take into account the proto-agent property of *independent existence*. This might be either due to the fact that he is unsure about its status in general or that he does not consider it to be relevant for the argument selection of psych verbs. The following table provides a revised version of Table 1, which, as I have argued, captures the assumptions of the proto-role model more adequately:

Verbs	**Arg1 (Subject)**		**Arg2 (Direct Object)**	
	Proto-agent	**Proto-patient**	**Proto-agent**	**Proto-patient**
(9a) *to fear* (SE-psych)	+ sentience + independent existence	+ causally affected	+ causation + independent existence	
Number of proto-properties	2	1	2	0
(9b) *to frighten* (OE-psych)	+ causation + independent existence		+ sentience + independent existence	+ causally affected ± change of state
Number of proto-properties	2	0	2	1-2

Table 2. Revised version of Dowty's (1991: 579f.) analysis of SE- and OE-psych verbs

Since psych verbs constitute a challenge for every linking model and are of importance throughout this work, I will sketch two other problems inherent in Dowty's account. These consider the assumed one-to-one correspondence of aspectual structure and construction type (SE- vs. OE-psych verb) on the one hand and the validity of the *causation* entailment for all types of stimuli on the other hand. As for the first point, Dowty (1991: 580) put forward the assumption that the construction type of OE-psych verbs is due to the possible *change of state* entailment for the experiencer. Kutscher (2009: 28) criticizes that this aspectually motivated explanation fails to account for the cross-linguistically frequent construction type */nom/dat* with the experiencer in object position (e.g. with the German verb *gefallen* 'to like') since these verbs do not have an inchoative but only a stative reading. Moreover, it should be added that there are also SE-psych verbs in English that have an inchoative interpretation such as *get angry with*, *get mad at* or *get bored with* (Croft 1993: 63). Contrary to the argument selection principle, in a sentence like *Mary got angry with Peter*, *Mary* as the subject argument would be assigned two proto-patient properties (*causally affected*, *change of state*) while *Peter* as the (prepositional) object does not bear any proto-patient property.

As for the second point, the view that the *causation* entailment is valid for the stimuli of *all* types of psych verbs is now largely rejected (Croft 1993: 61f.; Pesetsky 1995: 55-58; Blume 2000: 151; Levin/Rappaport Hovav 2005: 15; Kailuweit 2005: 106). In his causal structure analysis, Croft (1993: 61f.) contrasts causative psych verbs on the one hand (10) and stative psych verbs on the other hand (11). While the former can be broken down into three different subparts, the latter "consist of only a single stative segment" (1993: 62). This is proved by the insertion of a means clause (*by…*), which is only possible for verbs consisting of an initial causing subpart (10).

(10) + CAUSATIVE

a. Peter (intentionally) annoyed Mary by/ *because of his rude behaviour.
[OE-psych]

b. Peter's answer (*intentionally) annoyed Mary by/ because of its aggressive tone.
[OE-psych]

(11) - CAUSATIVE

a. Peter (*intentionally) interested Mary *by/ because of his kind behaviour.
[OE-psych]

b. Peter's book (*intentionally) interested Mary *by/ because of its sarcastic tone.
[OE-psych]

c. Peter likes Mary *by/ because of her kind behaviour.
[SE-psych]

The means clauses in (10) thereby specify the actual stimulus of the psych relation. It seems to make no difference if a human (potentially volitionally acting) (10a) or an inanimate (10b) subject is considered. For stative psych verbs, such as the SE-psych verb *to like*, but also some OE-psych verbs like *to interest* (Blume 2000: 151), a means clause cannot be constructed (11a-c). Here, only a causal clause is possible, which "is used to describe a source of causation prior to the verbal segment in the causal chain" (Croft 1993: 62). This analysis suggests that we may have to differentiate between causation as a lexical entailment, which is given in (10), and the source of causation (11), which is not lexicalized in the verb. If such differentiation is correct, the choice of an OE-encoding for the verbs of type (11a-b) could not be explained in terms of Dowty's selection principle: English verbs like *to interest* would entail *independent existence* to both of their arguments and *sentience* to their object. As a result, the object would have one proto-agent property more and, according to the argument selection principle, should be lexicalized as the subject.

For argument realization patterns of lexically causative and non-causative psych verbs, Croft (1993: 63) makes the following prediction: whereas the former type can be characterized by stable argument realization patterns, namely as OE-psych verb, only the latter type allows for a greater range of variation. Croft's prediction is supported by Levin/Rappaport Hovavs' (2005: 22f.) analysis of psych verbs. In their line of argumentation, causative OE-psych verbs show an agent-patient asymmetry that

makes them "close to prototypical transitive verbs" with the only difference being that they select for an animate, instead of an inanimate, direct object. This is borne out by Haspelmath's (2015: 143) finding that the causative psych verb 'frighten' cross-linguistically nearly always receives a transitive encoding (in .98 of the cases), whereas non-causative psych-verbs such as 'like' (.78) and 'fear' (.53) allow for more variation. This evidence seems at first contradictory to Kutscher's (2009) account on the construction variance of psych verbs: in her view, the variance covers all types of psych verbs and is due to the causal bidirectionality between stimulus and experiencer. She supports her hypothesis with evidence from psycholinguistic and neurolinguistic research (2009: 60). Still, Kutscher's account might be compatible with the presented construction stability for causative OE-psych verbs if we distinguish two different types of causation, namely the verbal entailment *causation* and the implicit causality of the stimulus which is inherent to every emotional expression. In case both are present for a given verb, it is the lexical entailment that might determine the argument realization pattern.[4] Section 3.2.2 will use Blume's (2000) linking model to show how the construction variance of non-causative psych verbs (and role-semantically similar verbs) can be captured within a proto-role approach.

There are two other remarks to be made. First, Dowty's model is not applicable to sentences in the passive voice (cf. Primus 1999a: 47; Levin/Rappaport Hovav 2005: 59). While this holds true, it must be underlined as by Ackerman/Moore (2001: 12) that this was not intended by Dowty, "since the entailment sets of those predicates related as active/passive pairs are not generally thought to be different". The possible criticism thus only relates to the broader applicability of the proto-role model within a theory of argument realization but does not cause contradiction within the model itself, which is "rather modest in scope" (Levin/Rappaport Hovav 2005: 59). It is hence left to further research to determine how Dowty's generalization on lexical predicates can be extended to syntactic alternations, such as active/passive.

Secondly, due to its restrictedness to transitive verbs, Dowty's approach does not allow for more general questions on argument realization, which concern for example the choice of a transitive vs. an intransitive case frame. This would not only be of interest for languages that have morphological case (e.g. accusative vs. dative selection in German) but also for a language like English. As shown in Haspelmath's (2015: 139) study on transitivity prominence, English has a transitivity score of .58 (comparable to German with .56). So, also with respect to English, the more general question arises to what extent role-semantic properties influence the choice of encoding, i.e. if a verb is transitive or intransitive, for example selecting a prepositional phrase such as *listen to*, *shout at*, *look for*, *play with*, etc. (cf. Haspelmath 2015: 141).[5] Here, it would be especially interesting to compare semantically related predicates

[4] Cf. Kutscher (2009: 60): "Ein Emotionsausdruck weist stets eine implizite Kausalität auf. Der Stimulus ist immer als der die Emotion auslösende Ursachenfaktor konzeptualisiert. Diese Konzeptualisierung bestimmt nicht den Konstruktionstyp des Psychverbs."

[5] This is relevant also with respect to the encoding of psych verbs, which in English can also get a prepositional object, e.g. *worry about* or *appeal to* (cf. Levin 1993: 192f.).

like English *to see* and *to look at*. While the first only implies *sentience* for its subject, the second entails *sentience* and *volition* for its subject. As for the object, in both cases no proto-properties are entailed at all. However, the two verbs differ with regard to their encoding: *to see* is transitive, whereas *to look at* gets a prepositional encoding. Interestingly, this difference reflects the cross-linguistic tendency for these two verbs, in that 'see' overwhelmingly choses a transitive encoding (.93), whereas 'look at' only has a transitivity score of .73. Moreover, apart from the applicability to transitive vs. prepositional encodings in English, the further question arises of whether Dowty's model can also be adopted in order to predict more specific kinds of morphosyntactic argument realization, such as DOM in Spanish (cf. García García 2014: 118ff.).

We can conclude that, in general, Dowty's description of semantic roles as two cluster concepts, one proto-agent and one proto-patient role, offers many advantages to traditional models of semantic roles. In this way, it overcomes the need for respecting necessary and sufficient conditions for certain semantic roles. Moreover, the possible combination of proto-agent and proto-patient properties for one argument (as e.g. in the case of OE-psych verbs) allows for a more exact analysis of certain arguments. The same applies for the innovation that one predicate could assign an equal number of proto-properties to its arguments (as e.g. in the case of symmetrical predicates). However, I have argued that Dowty's argument selection principle seems to be violated in more cases than he suggests. If we want to apply the model cross-linguistically, we would therefore need some modifications and specifications: in order to account not only for syntactic but also morphosyntactic argument realization and to be applicable to ergative languages, too, we would need a consideration of morphological case instead of the grammatical roles subject and object. This is precisely what has been put forward in later versions of Dowty's model by Primus (1999a, b, 2006) and Blume (1998, 2000). In this way, we can approach a possible application to differential morphosyntactic argument realization.

3.2 Application to morphosyntactic argument realization

In the following, I will introduce two modified versions of Dowty's (1991) approach that reformulate proto-roles as predictors of morphosyntactic argument realization (Primus 1999a, b, 2006; Blume 1998, 2000). In this way, one weakness of Dowty's argument selection principle, namely its restrictedness to syntactically transitive two-place predicates, would be overcome. As the morphosyntactic linking principles of Primus (1999a, b) and Blume (2000) demonstrate, proto-roles qualify for making cross-linguistically valid claims on case selection. Both models integrate their semantic case selection principles into an Optimality Theory (OT) model, in which these compete with formal constraints on case selection. This allows both constraints to be violated and ranked relatively to each other depending on language-specific preferences. We will see that both models provide important clues for a role-semantic treatment of DOM: Primus' (1999a, b, 2006) co-argument dependency model serves as the basis for a role-semantic explanation of DOM with inanimate objects in

Spanish (3.3). Blume's (1998, 2000) model, which focusses on the selection of marked case frames such as */nom/dat*, brings two main innovations that I will adopt in chapter 4: these consider the formulation of participant roles in relation to subevents as well as an inner-ranking of proto-agent properties.

3.2.1 Primus (1999a, b, 2006)

In this section, I will sketch the most important innovations of Primus' proto-role model (1999a, b, 2006) for the purposes of this work. These consider, on the one hand, the definition of roles and, on the other hand, the formulation of a morphosyntactic linking principle.

With regard to the proto-role definition, Primus deviates from Dowty's approach in that her proto-agent and proto-patient properties are not understood as predicate entailments but as basic predicates, i.e. as primitive predicates, that result from the decomposition of lexical meanings of predicates. With respect to their definition, the proto-agent and proto-patient basic predicates, listed in (12) and (13), show great overlaps with Dowty's (1991) lists of predicate entailments.

(12) Proto-agent basic predicates (Primus 1999b: 141)

a.	CONTROL(x, …)	x controls/controller
b.	CAUSE(x, …)	x causes/causer
c.	MOVE(x, …)	x moves/mover
d.	EXPER(x, …)	x experiences/experiencer
e.	POSSESS(x, …)	x possesses/possessor

(13) Proto-patient basic predicates

a.	CONTROL(x, y, …)	y is controlled
b.	CAUSE(x, y, ...)	y is caused
c.	MOVE(x, y, ...)	y is moved
d.	EXPER(x, y, ...)	y is experienced
e.	POSSESS(x, y, ...)	y is possessed

Dowty's entailment of *volition* is replaced by CONTROL(x, …), which is understood as a broader notion also involving the capacity to start and stop an event as well as taking responsibility for it (Primus 1999a: 36f.). Furthermore, a basic predicate of POSSESS(x, …) is added, which is not only of relevance for possession verbs such as 'posses' or 'belong to' but – as will be shown – is especially helpful for the role-semantic analysis of recipients of ditransitive change-of-possession verbs like 'give' (Primus 2006: 55ff.). However, a main deviation from Dowty's model is a conceptual one: Primus assumes two different dimensions of role-semantic information, namely

(i) *involvement* and (ii) *co-argument dependency*. The involvement dimension deals with the distribution of the basic predicates listed in (12) and (13). More precisely, if a transitive predicate entails a proto-agent basic predicate for an argument x, e.g. CONTROL(x, …), the converse proto-patient role can be derived for the relevant argument y (CONTROL(x, y,…)). So, the proto-patient in a given event is defined by its dependency on the involvement of the proto-agent. The concept of co-argument dependency was inspired by Dowty's proto-agent property of *independent existence*: As indicated in section 3.1, this property is entailed for all other proto-agent properties and thus seems to be of particular status among Dowty's entailments. In contrast to the various kinds of involvement relations that are lexically encoded, the notion co-argument dependency concerns a more general, binary understanding of causal dependency that is situated at the structural level. While the proto-agent is characterized by being causally independent from its co-argument and can therefore function as the only argument of a predicate, the proto-patient is defined by its co-argument dependency from the proto-agent. This is expressed by the ranking Proto-agent $>_{dep}$ Proto-patient and the impossible converse ranking *Proto-patient $>_{dep}$ Proto-agent (Primus 2006: 56-59, 74-79).

One further point of note is that Primus (1999a), unlike Dowty (1991), discusses the relation between the proto-agent basic predicates and animacy. As already argued in 3.1, the proto-agent property of *sentience* does not logically imply the animacy of a participant. The negation of a reverse entailment, namely that the animacy of a participant entails its sentience in a given event, is less controversial. This can easily be shown using the following example: *Mary saw John but he had no perception of it*. If the animacy of *John* would entail his sentience in the event, the cancelation of the implication in the second half of the sentence would make the utterance infelicitous. Nonetheless, Primus notices that animacy "may increase the agentivity of the participant on the basis of the conversational implicature that animate causers *(John killed the bird)* are interpreted as controllers *(John killed the bird intentionally),* if there is no information from the context to the contrary" (1999a: 58). She formalizes this implicature, which has been previously stated by Holisky (1987: 119) and more generally formulated but not formalized by Foley/Van Valin (1984: 32), in the following way:

(14) $\forall x[\text{CAUSE}(x, \ldots) \,\&\, \text{ANIMATE}(x) +> \text{CONTROL}(x, \ldots)]$[6]

That CONTROL(x, ...) has the status of a pragmatic implicature in the sense of Grice (1975: 56ff.) and is not anchored in the lexical meaning of the verb can be proved by the possibility to cancel it in sentences such as *John accidentally killed the bird*. Moreover, the possibility for verbs like *to kill* to select a subject incapable of controlling a situation, e.g. in *The cold killed the bird* reveals that CONTROL(x, ...) cannot

[6] Primus (1999a: 49f.) presents the following two premises for the givenness of an implicature (P-CONTROL and P-EXPER refer to the *prototypical* controlled or experiental situation, respectively):

1. $\forall x \forall y[\text{CONTROL}(x \ldots) \rightarrow \text{CAUSE}(x \ldots)]$
2. (a) $\forall x[\text{P-CONTROL}(x, \ldots) \rightarrow \text{ANIMATE}(x)]$
   (b) $\forall x\ [\text{P-EXPER}(x, \ldots) \rightarrow \text{ANIMATE}(x)]$

have the status of a semantic entailment. García García (2014: 132f.) further shows that these systematic relationships are only found between animacy and the proto-agent role but can be disproved for animacy and the proto-patient role. Due to the interpretation of animate participants as agentive participants in a given event, he refers to the former as *potential proto-agents*. However, most of the correlations drawn between animacy and agentivity consider the subject function and the constellation of an argument bearing exclusively proto-agent basic predicates. There are some clues, though, that the overlap of animacy and proto-patient properties in object function may also lead to a conversational implicature of the following type:

(15) ∀x[CAUSE(x, y...) & ANIMATE(y) +> EXPER(y, x...)]

The implicature I propose can be formulated in the following way: "You may interpret proto-patients that are human as experiencers (in absence of any information to the contrary)". This can be illustrated using the verb of contact by impact *to hit*:[7]

(16) a. Peter hit the door with a stick.

b. Peter hit Mary with a stick. +> Mary felt pain

c. Peter hit Mary with a stick but she did not feel any pain or even notice.

With the help of (16a), we can prove that the verb *to hit* does not imply EXPER(y, x...) for its object: a door is not capable of being sentient in any way and the sentence is fully grammatical. In the case of a human object, like *Mary* in (16b), we typically interpret that the act of hitting not only applys to the physical but also to the psychological dimension, in that Mary felt pain. We thus interpret Mary as being an experiencer with respect to the event denoted by the predicate. This implicature is cancellable as (16c) illustrates. How does the implicature arise? First of all, the basic predicate EXPER(y, x...), or in Dowty's (1991) terms, the proto-agent property of *sentience*, regularly occurs in object position, namely with OE-psych verbs such as *to frighten* or *to worry*. In the case of causative OE-psych verbs, *sentience* overlaps with

[7] In some Germanic languages, e.g. in Dutch and German, verbs of contact by impact show a particularity in case selection: Whereas e.g. German *schlagen* 'to hit' selects a transitive case frame with an animate object (e.g. *Peter schlägt die Frau.* 'Peter hits the woman'), inanimate objects are encoded as prepositional phrase (e.g. *Peter schlägt gegen den Tisch.* 'Peter hits (against) the table'). The encoding alternation has been intensively studied in recent years and has been given different explanations (cf. de Swart 2014; de Hoop 2015; de Swart & de Hoop 2018; Fleischhauer 2018). De Swart (2014) in his account considers the encoding difference in Dutch as an instance of *paradigmatic DOM*, based on the paradigmatic argument selection principle of Ackerman/Moore (2001) in combination with Dowty's (1991) and Primus' (1999a) proto-role model. His basic assumption is that animate objects are better patients since, in comparison to inanimate objects, they are assigned one proto-patient property more, namely *sentience*. This analysis is problematic for two reasons: First, Dowty (1991) and Primus (1999a) clearly do not consider *sentience* a proto-patient but a proto-agent property. Consequently, sentient objects should be analyzed performing less well as patients since patient-properties overlap with a proto-agent property. Second, the status of *sentience* as an entailment for Dutch verbs such as *slaan* 'to hit' is questionable since it can be cancelled in context: *Peter sloeg Mary met een stok, maar ze voelde geen pijn en merkte het niet eens op omdat ze verdoofd was* 'Peter hit Mary with a stick, but she did not feel any pain or even notice because she was sedated' (comparable to English *to hit* in 16c). Therefore I propose that it should rather be treated as a conversational implicature which of course needs further review (cf. section 4.2.1 f. below).

the proto-patient property *causally affected* and, in an inchoative reading, also with *change of state*. Secondly, human beings are the core cases of sentient beings. They are the only kind of entities that can perform in any given predicate, which implies *sentience* for an argument. Put together, it is comprehensible that human patients are interpreted as being sentient with respect to the event, if not the contrary is indicated. Dowty (1991: 572, fn. 16) already made a similar point when dissociating the notion of *predicate entailment* from other sorts of 'entailments':

> It is important here to distinguish entailments of the PREDICATE from what follows from any one sentence as a whole (e.g. entailments that may arise in part from NP meanings. etc.). For example, if *Mary slapped John* is true, and John is a normal human, then, slapping being the kind of action it is, we would conclude that John necessarily perceives something (and we would do likewise from the majority of sentences using *slap*). But it does not follow that the direct object of slap is entailed to have the P-Agent property of sentience, since we can also felicitously say *Mary slapped the table* or *Mary slapped the corpse*. However, the object of *awaken* does have the P-Agent entailment of sentience, as is revealed by the anomaly of *#Mary awakened the table/ the corpse*. (Dowty 1991: 572, fn. 16)

For the purposes of this study, it is important to keep in mind that the relationships between animacy and the proto-agent role not only apply to subjects but seem to also play a role for objects. I will come back to this discussion in chapter 4.

Besides the roles of proto-agent and proto-patient, Primus assumes a third cluster concept, namely that of *proto-recipient* (1999a: 54ff., b: 145), which serves as a "convenient label for thematic roles that combine Proto-Agent and Proto-Patient properties" (1999a: 54). Due to this characteristic, it is of particular interest for this study. The proto-recipient covers the traditional semantic roles of addressee, benefactive and recipient. Consider the two following examples:

(17) *Peter gave Mary an apple.*

∀x∀y∀z[GIVE(x, y, z) → P-CONTROL(x, BECOME(POSS(y, z)))]

(18) *Peter showed Mary his new car.*

∀x∀y∀z[SHOW(x, y, z) → P-CONTROL(x. BECOME(EXPER(y,)))]

(Primus 1999a: 54)

In (17), the verb *to give* induces a change of possession. The subject *Peter* (x) is a proto-agent relative to the new possessor *Mary* (y) since it has control over the giving-event. Consequently, *Mary* (y) is a proto-patient relative to *Peter* (x). At the same time, *Mary* (y) is a proto-agent relative to the direct object *an apple* (z) having the proto-agent basic predicate POSSESS(y, z). In (18), we can identify a similar distribution of proto-agent basic predicates with the difference being that *Mary* (y) is a proto-agent relative to the direct object *his new car* (z) not because it takes the first position of POSSESS(y, z) but of EXPER(y, z) since the indirect object can be attributed *visual experience* (Primus 1999a: 55). The dependency relations between proto-agents, proto-patients and proto-recipients are captured by Primus in the following hierarchy:

(19) Proto-Agent <θ Proto-Recipient <θ Proto-Patient

In the thematic hierarchy, proto-recipients thus occupy an intermediate position between proto-agents and proto-patients: on the one hand, their involvement is dependent on the involvement of a proto-agent in a CONTROL(x,…) or CAUSE(x,…) predicate, while on the other hand, they are the first participant of a second embedded predicate (POSSESS(y,…) or EXPER(y,…)). The proto-patient, by contrast, is dependent on the involvement of the proto-recipient and thus takes the second position of the POSSESS(y,…) or EXPER(y,…) basic predicate (Primus 1999a: 55).

In a further step, Primus (1999a, b) formulates an argument selection principle that is based on the accumulation of proto-agent and proto-patient basic predicates. Unlike Dowty (1991: 576), her principle is not limited to syntactic argument realization in a transitive sentence, i.e. the realization of subject and direct object, but also considers morphosyntactic argument selection, i.e. the selection of case frame:

(20) Argument selection principle (Primus 1999a: 61)

> The Principle of Morphosyntactic Expression of Thematic Information:
>
> For any language L, for any participants that are syntactic arguments, and for any distinct morphosyntactic coding categories A, B and C such that A and B are the highest ranking coding categories in L:
>
> (a) The greater the number of Proto-Agent basic relations a participant accumulates in the unmarked reading of the predicate (or sentence), the more likely it is coded by A.
>
> (b) The greater the number of Proto-Patient basic relations a participant accumulates in the unmarked reading of the predicate (or sentence), the more likely it is coded by B.

Similar to Dowty's (1991: 576) argument selection principle, Primus takes the accumulation of proto-agent and proto-patient basic predicates of an argument as indicators for their linking behaviour: while an argument with a great number of proto-agent basic predicates in accusative languages is linked to the nominative (A= nominative), an argument bearing a high number of proto-patient basic predicates is encoded in the accusative (B= accusative). The major departure from Dowty's (1991) proposal is that Primus (1999a: 61f.) replaces grammatical relations with case categories. This modification offers several advantages: on the one hand, her principle is not only applicable to accusative but also to ergative languages.[8] This is achieved by the general assumption that proto-agent and proto-patient involvement relations are not ranked relatively to each other. On the other hand, the use of grammatical relations such as subject and direct object would blur the dissociation to the second type of role-semantic information, namely co-argument dependency.

[8] In the ergative parameter, category A would be the ergative and category B the absolutive (or nominative) (Primus 1999a: 61).

The linking principle has some corollaries of which the first four are listed below, whereby category C can be a dative or another oblique case:

(21) Corollary I: The fewer Proto-Agent basic relations a syntactic argument accumulates, the more likely it is coded by a category distinct from A (i.e. by B or C).

Corollary II: Given two syntactic co-arguments with Proto-Agent properties, the argument with the greater number of Proto-Agent properties is more likely to be coded by A.

Corollary III: The fewer Proto-Patient basic relations a syntactic argument accumulates, the more likely it is coded by a category distinct from B (i.e. by A or C).

Corollary IV: Given two syntactic co-arguments with Proto-Patient properties, the argument with the greater number of Proto-Patient properties is more likely to be coded by B.

Thus, the case category C can be described both by a weaker number of proto-agent basic predicates than A and by a weaker number of proto-patient basic predicates than B. However, the dative is characterized by a weak number of proto-agent properties, as a further semantic default suggests. The so-called *dative default*, presented in (22), is of importance for this work since it points to an influence of agentivity on case selection:

(22) Dative-Default (Primus 1999a: 66):

If a syntactic verb argument is in the dative, the corresponding semantic argument (i.e. participant) has a small number of Proto-Agent properties as a strong preference.

First of all, it is notable that unlike common linking principles, the dative default is formulated in the direction FORM (CASE) → FUNCTION (WEAK PROTO-AGENT) and not the other way around (FUNCTION (WEAK PROTO-AGENT) → FORM (CASE)). Primus herself admits, "the reverse Dative-Default that maps the relevant thematic roles onto the dative case is weaker" (1999a: 66). This weakness of the reverse default that shrinks its usability as a linking principle has already been criticized by Kailuweit (2004: 91) and García García (2014: 127). This point can be illustrated with the following German examples:

(23) a. *Das Buch gefällt Peter.* /nom/dat

the book.NOM please.PRS.3SG PN.DAT

'The book pleases Peter.'

b. *Peter mag das Buch.* /nom/acc

PN.NOM like.PRS.3SG the book.ACC

'Peter likes the book.'

c. *Das Buch interessiert Peter.* /nom/acc

the book.NOM interest.PRS.3SG PN.ACC

'The book interests Peter.'

All three sentences (23a-c) show non-causative psych verbs that assign the proto-agent basic predicate EXPER(x,y) to one of their arguments. As such, it is an argument bearing a small number of proto-agent properties. However, in all three cases the realization of the weak proto-agent differs: whereas in (23a), it is realized as a dative, it is coded with nominative case in (23b) and with accusative case in (23c). Hence, the dative default does not qualify as a reliable linking principle for the case selection of experiencers of non-causative psych verbs.

In addition, a cross-linguistic account reveals that the dative default in its original direction must be weaker than originally formulated by Primus: The assumption of the default, that in */nom/dat* case frames, it is the dative that expresses a weak proto-agent has been proved for languages such as German, Italian, Spanish, Korean and Russian but disproven for other languages, e.g. Samoan and Tongan (Blume 2000: 157). In the latter two languages, in a */nom/dat* case frame, the experiencer is expressed in the nominative and the stimulus is encoded in the dative.

An important part of Primus' (1999a, b) proto-role model is the assumption of a formal case principle that competes with the above-mentioned semantic case principle in (20). The formal *subcategorization principle* is defined as follows:

(24) The Subcategorization Principle:

> Ideally, the assignment of a lower ranking formal category (e.g. case) by a predicate P implies asymmetrically the assignment of a higher ranking formal category by P; the higher the rank of a formal category is, the less restricted the class of predicates that assign it. (Primus 1999: 24)

It states that whenever a lower-ranking formal category of the case hierarchy has been assigned (e.g. accusative), the assignment of a higher-ranking category (here nominative) is implied. Both the semantic case principle and defaults and the formal subcategorization principle are understood as violable. They are modelled within OT suggesting that both function as constraints that compete with each other. It is thereby highly language-specific how formal and semantic constraints are ranked to each other. For German, the following ranking of constraints is assumed (Primus 1999a: 67):

(25) Distinctness and Nominative Requirement >> NOM/ACC-preference for polyvalent verbs and semantic defaults

Violations of the nominative requirement are rare and can be attested for instance by one-place predicates with non-nominative experiencers. Here, */acc/* case frames like (26a) are clearly less frequent than */dat/* ones (26b) which are still productive (Primus 1999a: 67):

(26) a. *Mich friert.* */acc/*

1SG.ACC freeze.PRS.3SG

‘I am freezing.’

b. *Mir* *ist* *kalt.* */dat/*

1SG.DAT be.PRS.3SG cold

‘I am cold.’

In later versions of her model, Primus (2002, 2004, 2006) formalizes her linking principle in a more precise way. Moreover, the set of basic predicates and their representation differ slightly:

(27) a. ctrl(x,s) x controls the situation s denoted by the predicate

b. exp(x,y) x is sentient of y

c. phys(x,y), phys(x) x physically contacts or moves y, x moves or is active

d. poss(x,y) x possesses y

(Primus 2006: 55)

As shown by the proto-agent basic predicates listed in (27), the proto-agent basic role CAUSE(x,…) has been deleted. This is due to its redundancy since in Primus’ model the givenness of a proto-agent role already establishes a co-argument dependency relation between two arguments with the proto-agent being causally independent from the proto-patient. One important innovation is a distinction between minimally and maximally involved roles: “θ^{max} is an abbreviation for a role with a large number of consistent Proto-Role properties; θ^{min} means that an argument accumulates a smaller number of consistent Proto-Role properties or none at all” (Primus 2006: 61). Thus, depending on the number of proto-properties, a proto-agent can be either A^{max} or A^{min} and a proto-patient either P^{max} or P^{min}. Yet Primus admits that “the cut off point between A^{max} and A^{min} on the one hand and P^{max} and P^{min} on the other hand can only be decided on a language specific basis” (2006: 81, fn. 6).

Based on the distinction of maximally and minimally involved roles, Primus (2004: 381, 2006: 62) presents the following schema on case linking in accusative languages (n: not highest ranked case; m: not 2nd ranked case):

(28) Accusative Ranking (n≠1; m≠2)

a. A^{max} / 1C >> A^{max} / nC

∨ ∧

∨ ∧

A^{min} / 1C A^{min} / nC

b. P^{max} / 2C >> P^{max} / mC

∨ ∧

∨ ∧

P^{min} / 2C P^{min} / mC

In the horizontal order, the ranking for case selection is thus as follows: If we deal with a maximally involved proto-agent (A^{max}), the constraint to code this argument in the nominative is higher ranked to code it in a non-nominative case. At the same time, the constraint to code a maximally involved proto-patient (P^{max}) in the accusative dominates the constraint to code this argument in another non-accusative case (Primus 2004: 381f.). For minimally involved roles, i.e. A^{min} and P^{min}, a similar ranking of constraints cannot be established. This is reflected in variable case patterns for these roles: "Case variation is explained in our approach by the fact that minimally involved roles do not participate in the invariant ranking proposed for maximally involved roles" (Primus 2006: 67). In the vertical order, the constraints that A^{max} is coded in the nominative and P^{max} in the accusative dominate the constraints that A^{min} or P^{min} are encoded in these cases, respectively. Correspondingly, the constraints that A^{min} takes a non-nominative and P^{min} takes a non-accusative case are higher ranked than the constraints that A^{max} and P^{max} are encoded this way.

However, as I will further argue, the linking principle only holds true under the premise that θ^{max} and θ^{min} are not only distinguished in terms of quantity but also in terms of quality. That is, we have to assume a weighting of different proto-agent basic predicates. This can be illustrated with the following four examples. In (29), we have a highly transitive predicate (*to eat (an apple)*). Here, the subject *Peter* is a maximally involved proto-agent bearing the three roles ctrl(x,s), exp(x,y) and phys(x,y), while the direct object *an apple* consequently is a maximally involved proto-patient:

(29) Highly transitive verb: *to eat (an apple)*

Peter eats an apple.

ctrl(x, exp(x, y), phys(x,y)) A^{max} (3 roles) + P^{max} (3 roles)

Yet considering another highly transitive verb, namely *to break*, we are confronted with the following theoretical problem: The predicate only entails one proto-agent role to its subject argument, that is phys(x,y). Consequently, the direct object is assigned the proto-patient role *physically affected*:

(30) Highly transitive verb: *to break*

Peter accidentally broke the glass.

phys(x,y) A^{max} (1 role) + P^{max} (1 role)

Hence, we only have one proto-role for subject and direct object. However, in order to fulfil the linking constraints for case selection shown above in (28) and not

contradict typological studies showing that 'break' verbs are always transitively encoded (Haspelmath 2015: 143), we have to assume the proto-roles in (30) to be maximally involved (A^{max} and P^{max}). For a similar reason, we have to treat the experiencer proto-role of a non-causative psych verb such as *to like* in (31) as minimally involved (A^{min} and P^{min}). The same holds true for the proto-role of possession for a verb like *to own* in (32). These two types of predicates typically show cross-linguistic variation in case marking, which in Primus' linking principle is only allowed for minimally involved roles.

(31) Non-causative psych verb: *to like*

Peter likes Maria. A^{min} (1 role) + P^{min} (1 role)

(32) Possession verb: *to own*

Maria owns a house. A^{min} (1 role) + P^{min} (1 role)[9]

Thus, in order to make Primus' model compatible with empirically attested linking tendencies, we have to assume that not only the quantity but also the quality of the proto-properties is decisive. More precisely, the proto-agent basic roles of exp(x,y) and poss(x,y) seem to be weaker than the other two roles (ctrl(x,s) and phys(x,y); while ctrl(x,s) entails exp(x,y)). The suggested role-internal ranking can be formulated in the following way: ctrl(x,s), phys(x,y) > exp(x,y), poss(x,y). While Primus also assumes a ranking Agent $>_{invol}$ Experiencer, it is only justified in terms of quantity: "A volitional, active Agent accumulates more Proto-Agent properties than an Experiencer, which is only characterized by the entailment of sentience" (2006: 66). The fact that such ranking is rather qualitative is reflected in theories of argument realization assuming a priority of causal notions including control over experiencers (Levin/Rappaport Hovav 2005: 117-125, cf. references therein). With respect to the examples of this study, it is confirmed by the cross-linguistically invariant linking of causative psych verbs in comparison to the construction variance found with non-causative psych verbs[10] as discussed in section 3.1 with reference to Croft (1993). In the following section, I will present a proto-role model and linking principle (Blume 2000) that has systematically implemented a weighting of proto-agent properties.

[9] Primus (2006: 61) does not explicitly mention that possessors fall under A^{min}. The reason to subsume them under this notion lies in the observation that possession verbs show a similar construction variance to psych verbs (Primus 1999a: 70f.; Blume 2000: 211-213).

[10] Despite its variance in case linking, Primus (2004: 392-397, 2006: 76-78) also assumes non-causative psych verbs to bear the co-argument dependency relation Experiencer $>_{dep}$ Stimulus, which in her model is understood to be a special case of the thematic dependency relation Proto-agent $>_{dep}$ Proto-patient. Following Primus' argumentation, this underlying dependency relation shows up in word order in that the experiencer is preferred to be structurally linked to the first position (e.g. for verbs like *mögen* 'to like', *lieben* 'to love', *gefallen* 'to like', *interessieren* 'to interest' in German, 2004: 395-397). However, as has been shown by a more recent experimental study of Verhoeven (2014), experiencer-first effects are determined by language-specific properties and are due to the lexical entry of an individual predicate in a language. More precisely, they are restricted to languages that, like German or Greek, have a specific type of OE-psych verbs, namely non-causative ones. In languages like Turkish or Mandarin Chinese, that do not make this distinction and have only ± causative predicates, experiencer-first effects are considerably less attested.

A further complication of the model that leads to false predictions in linking concerns interaction verbs such as *to help*. Cross-linguistically, these verbs often select marked case frames such as *helfen* in German, which is */nom/dat* (Blume 1998, 2000). However, the analysis of proto-roles in terms of Primus' (1999a, b, 2006) model reveals that the subject is a maximally involved proto-agent having phys(x,y) and exp(x,y) while the object is a maximally involved proto-patient:[11]

(33) Interaction verb: *to help*

Maria helps Peter with the dishes.

exp(x, y), phys(x,y) A^{max} (2 roles) + P^{max} (2 roles)

With respect to case selection, this leads to false predictions since maximally involved roles are said to select an invariant linking pattern. As I will show in the next section, linking patterns of interaction verbs can be more adequately captured in a model that (i) defines proto-roles as relations of participants in (implied and presupposed) subevents and that (ii) does not assign a corresponding proto-patient role for *any* proto-agent role (Blume 2000: 159f.).

In summary, Primus' (1999a, b, 2006) linking principle can be seen as an important innovation in comparison to Dowty's (1991) argument selection principle. Due to its formulation in terms of morphosyntactic cases it allows more fine-grained predictions on argument coding for both accusative and ergative languages. Unlike Dowty, who considers argument selection to be only triggered by lexical properties of predicates, Primus makes a difference between a semantic and a formal case principle. Whereas the former concerns lexical properties and hence the involvement of a participant, the latter refers to the so-called subcategorization principle stating that once a lower-ranking formal category of the case hierarchy has been assigned (e.g. accusative), the assignment of a higher-ranking category (nominative) is implied. Both constraints are violable and are in competition with each other. As a result, the relative ranking of both constraints to each other is determined by language-specific preferences. However, I have put forward some points of criticism with regard to Primus' model: these concerned the definition of θ^{max} and θ^{min}, which, if only based on the quantity of basic predicates, leads to false predictions for linking. Moreover, Primus' model cannot account for the cross-linguistically attested preference of interaction verbs for selecting marked case frames. As for the selection of */nom/dat* case frames, Primus' (1999a: 66) dative default, formulated as a semantic default ascribing a weak number of proto-agent properties to an argument coded in the dative, cannot be

[11] Primus (2002: 32f.) presents a different analysis in which the object of German *helfen* 'to help' is analyzed as an A^{min} since the object participant is autonomously active in a presupposed event. In this way, the selection of a */nom/dat* case frame should be accounted for. In my view, this analysis is hardly compatible with her model for two reasons: Firstly, the model generally does not define basic predicates in relation to (implied and presupposed) subevents of two-place verbs. Secondly, in a consequent analysis, the subject of *helfen* would be assigned phys(x,y) and would hence be an A^{max}. Thus, besides being an A^{min}, the object would also be a P^{max}. For A^{max} and P^{max}, Primus' linking principle predicts invariant case selection. A possible solution would be to assume variation in linking if A^{min} and P^{max} overlap for an argument.

qualified as a linking principle in the strict sense due to its direction from form to function. Additionally, it only seems the case for a subset of languages but does not qualify as a licensing condition for dative encoding cross-linguistically. In the next section, I will present an alternative proto-role account and linking model of Blume (1998, 2000), which proposes a solution to the points mentioned based on a wide range of cross-linguistic data.

3.2.2 Blume (1998, 2000)

The linking model elaborated by Blume (1998, 2000) aims at proposing a comprehensive account on cross-linguistic preferences in the selection of marked case frames. It is based on the proto-role model by Dowty (1991) and Primus' (1999a)[12] modification of it. On the one hand, Blume herself makes linking-relevant modifications to the definition of proto-roles: these are (i) an integration of the temporal structure of the denoted events as well as (ii) a differentiation between potent and non-potent proto-agent properties. On the other hand, she more precisely states the previously introduced argument selection principle of Primus (1999a). As will be shown, Blume's semantic licensing condition for marked case frames accounts for a greater range of linking patterns than Primus' dative default. In particular, it can explain both the construction variance of non-causative psych verbs and the preference of interaction verbs to select for a marked case frame. For the present treatment on DOM, we will further see that overgeneralizations of the licensing condition are of particular interest.

In a first step, I will illustrate Blume's (2000) version of the proto-role model. The following list introduces the relevant proto-agent (34) and proto-patient (35) properties of her model (Blume 2000: 169f.; translation: S.M.). The role-semantic relation in (36) applies if neither proto-agent nor proto-patient properties are assigned to an argument.

(34) **Proto-agent properties:**

The following properties of a participant A at an atomic subevent s^i in the event structure S of a verb are proto-agent properties:

a. A controls s^i; including all involved participants (CONTROL);

b. A is aware of s^i including all involved participants/ A perceives something in s^i/ feels something in s^i (SENTIENCE);

c. A strives for a certain aim in s^i/ has a certain interest with respect to s^i (INTEREST);

d. A is autonomously active in s^i or fulfils a certain function/ a purpose (ACTIVITY/ FUNCTION).

[12] Blume (1998, 2000) in her work refers to an unpublished manuscript by Primus (1994). Since it expresses the same ideas as her later published monograph (Primus 1999a), I will refer only to the latter.

A participant who has at least one proto-agent property in an atomic subevent of S is the agent of this subevent; if he has all proto-agent properties in an atomic subevent, he is the proto-agent of this subevent.

(35) **Proto-patient properties:**

The following properties of a participant P at an atomic subevent s^i in the event structure S of a verb are proto-patient properties:

a. P is controlled by another participant A in s^i (CONTROLLED);

b. P is affected by the physical or mental activity of another participant A in s^i (AFFECTED).

A participant who bears the proto-patient property 'affected' in an atomic subevent of S is a patient of this subevent; if he bears all proto-patient properties in an atomic subevent he is a proto-patient of this subevent.

(36) **'Thematic':**

A participant P in an atomic subevent s^i of the event structure S of a verb to which none of the statements in [(34)] and [(35)] applies has the thematic role 'thematic'.

I will discuss the individual properties in comparison to the two previously presented models. In general, Blume's understanding of proto-roles overlaps both with Dowty's (1991) and with Primus' (1999a, b) models. One basic similarity with Dowty is understanding proto-agent and proto-patient properties as lists of entailments and not as basic predicates as in Primus' model. This decision is justified as follows: In Primus' (1999a, b) model, the basic predicates (CONTROL(x,…), CAUSE(x,…), MOVE(x,…), EXPER(x,…), POSSESS(x,…)) are characterized by the co-argument independency of the proto-agent argument. Unlike the proto-patient, which is always dependent on a proto-agent and can thus never occur without the latter, a proto-agent is causally independent, as reflected in the possible occurrence of a proto-agent as the single argument of a predicate. However, Blume (2000: 130f.) doubts the applicability of this dependency notion on the role-semantic relation between proto-agent and proto-patient: Assuming the proto-agent argument x to be always obligatory and the proto-patient y, in contrast, to appear only optionally, Primus presupposes that one and the same predicate (e.g. CONTROL(x,…)) can either be a one-place or a two-place predicate. However, this contradicts the basic distinction of predicate types: "Ein Prädikat, das die Sättigung *eines* Arguments fordert, kann aber per definitionem nicht mit einem Prädikat identisch sein, das die Sättigung *zweier* Argumente verlangt" (Blume 2000: 131; italics: S.M.). Consequently, in her view, the dependency relation for a two-place predicate cannot be determined unequivocally since none of the two arguments x and y can occur without the other. Yet Blume (2000: 131) stresses that she does support Primus' (1999a, b, 2006) view in that proto-patient properties are defined by their dependency on a corresponding proto-agent property (e.g. (35b) depends on (34d)). The chief deviation from Primus' model is that for a two-place predicate, proto-agent properties (such as (34a) *control*) can also be assigned to an argument x in a subevent *without* assigning the converse proto-patient property to the other argument y (Blume 2000: 170). This can be illustrated

with an example comparing the role-semantic analysis of one and the same sentence for the three models discussed:

(37) *Peter looks at the apple tree.*

(1) Dowty (1991): x: volition, sentience; y: -

(2) Primus (1999a, b): CONTROL(x, EXPER(x, y))

(3) Blume (2000):

$\forall x \forall e \forall e^1$ [look at(x, e) → e^1 [x- CONTROL, ACTIVITY, SENTIENCE, y- THEMATIC]]

With regard to the subject argument, the assignment of proto-agent properties is quite similar in the three models: they all assign *volition* (or *control*) and *sentience* to x, while in Blume's (2000) model, *activity* is additionally implied for x. For the direct object, though, the three analyses differ: whereas Primus (1999a, b) assumes y to be a maximally involved proto-patient, Dowty (1991) does not assign any proto-properties to y. Blume's (2000) analysis is comparable to Dowty's with the difference that she makes use of the explicit label *thematic* for arguments that are not assigned any proto-agent or proto-patient properties in any subevent.

The definition of proto-roles as relations of participants to subevents brings another deviation concerning the establishment of causal relations: in Dowty's (1991) as well as in Primus' (1999a, b) early models, *causation* is defined as a proto-agent role being on a par with other proto-agent roles like *volition*/*control*, *sentience* or *movement*. This is shown in the following examples:

(38) *Peter killed the pig.*

(1) Dowty (1991): x: causation, (movement), independent existence;
y: causally affected, (stationary movement), change of state, dependent existence

(2) Primus (1999a, b): CAUSE(x, BECOME(DEAD (y))

(3) Blume (1998, 2000):

$\forall x \forall y \forall e \forall e^1 \forall e^2$ [kill(x,y,e) → e^1 [x-ACTIVITY, y-AFFECTED] > e^2 [y-ø]]

Thus, while *causation* is an ordinary proto-agent property in Dowty's and Primus' model, in Blume's analysis it "follows from the subsequent ordering of the subevents and from the fact that they share an argument; i.e. causation is not a relation between arguments but between subsequent or parallel subevents" (1998: 278, fn. 8).[13]

Blume's model offers two main innovations that, as is further argued, are the prerequisites for formulating more precise predictions for linking phenomena than Primus' (1999a, b, 2006) approach. This is true in particular for the selection of marked case

[13] This can be formalized in the following way (Engelberg 2000a: 333):
[(47)] a. E-STR: ($\rightarrow_i$ $\mathbf{e^1}$) < CAUSE ($\rightarrow_i$ $\mathbf{e^2}$)…
b. …→ CAUSE(e^1,e^2) & (e^1 < e^2)

frames. The first is the mentioned definition of proto-roles as relations of participants to subevents (Blume 1998: 266, 2000: 164ff.). This idea is based on Engelberg's (1995, 2000a, b) *Lexical Event Structure Theory* which makes the basic assumption "that lexical representations of verbs should be based on the notion of events and their properties, i.e., their mereological structure, the length of subevents and the temporal relations between them as well as the relations between subevents and their participants" (2000b: 332). Verb meanings are represented as lexical event structures (LES) and include the following information:

(39) a. **Complexity of events:**
Verbs refer to events that are internally structured in the sense that they can consist of different subevents (e^1, e^2,…) and a possible result state (s).

b. **Sorts of subevents:**
The subevents are durative (e^{DUR}) or punctual (e^{PCT}).

c. **Relations between subevents:**
A subevent can precede another subevent (<) or subevents can be temporally parallel or to a large degree overlapping (<>).

d. **Participation in subevents:**
The event participants which correspond to the verb arguments are not necessarily involved in all subevents, but rather only in some of them; semantic functions like 'control', 'move', 'volition', etc., relate participants and subevents.

e. **Implication vs. presupposition:**[14]
The occurrence of a subevent is either entailed ($\rightarrow_I$) or presupposed ($\rightarrow_P$) by the open proposition that constitutes the verb's meaning, i.e., by an expression like 'verb(**x**,**y**,**e**)'.
(Engelberg 2000b: 332)

I will illustrate how these characteristics are formalized with the example of the lexicon entry (short notation) of German *fangen* 'to catch' (Engelberg 2000a: 335; translation: S.M.):

(40) *fangen*: x^{nom}, y^{acc}

E-STR: $(\rightarrow_P \mathbf{e^1}: y^{move}) < (\rightarrow_I \mathbf{e^{2[PCT]}}: x^{impulse,\ sentience\ldots}, y^{affected})$[15]

$< (\rightarrow_I \mathbf{s}: x,y)$

The meaning of *fangen* 'to catch' involves two subevents: In the sentence *Sabine hat den Baseball gefangen* 'Sabine caught the baseball', there is a first presupposed subevent (e^1) which denotes the flying of the baseball and a second implied subevent (e^2) that implies Sabine's grabbing of the ball. Furthermore, the verb implies a result state (s) in which the agent controls the ball. Importantly, the lexicalization of *fangen*

[14] Engelberg (2000a: 68-78) shows that the distinction between implied and presupposed subevents provides a clue for understanding the puzzling and seemingly contradictory behaviour of some punctual verbs. So e.g. the fact that some resultative punctual verbs can occur in the progressive (*Rebecca was winning the race/ was arriving*) while others cannot *(??Rebecca was noticing that/??That was astonishing Rebecca*) can be explained as follows: only the former but not the latter presuppose a durative preceding subevent to which the progressive refers (2000a: 74f.; 2000b: 263).

[15] Engelberg (2000a: 206-210) argues that for action verbs, such as unergatives like *to walk*, *to jump* or *to clap*, volition (or intention) is not a necessary condition but more a possible concomitant. Instead, from a neurological perspective, these verbs imply the *impulse* for an action to take place.

and its event time only refer to e^2 (Engelberg 2000a: 34, 335). The assumption of a presupposition in e^1 is can be tested by its consistency in an interrogative clause (41b) and under negation (41c) (cf. Engelberg 2000a: 34):

(41) a. *Sabine hat den Baseball gefangen.* → *Der Baseball ist geflogen.*
'Sabine caught the baseball.' 'The baseball flew.'

b. *Hat Sabine den Baseball gefangen?* → *Der Baseball ist geflogen.*
'Did Sabine catch the baseball?' 'The baseball flew.'

c. *Sabine hat den Baseball nicht gefangen.* → *Der Baseball ist geflogen.*
'Sabine did not catch the baseball.' 'The baseball flew.'

In Blume's (1998, 2000) model, the specification of proto-roles for subevents and the distinction between implied and presupposed subevents is particularly important to account for the linking of interaction verbs. Cross-linguistically, these verbs show a preference for marked case frames, such as */nom/dat* or */abs/dat*. The class of interaction verbs is rather loosely defined comprising mainly two-place verbs of communication and social gesture (e.g. 'to greet', 'to call for', 'to listen to') but also motion (e.g. 'to follow', 'to meet') and so-called 'obey' verbs ('to obey', 'to serve') (Blume 1998: 273f.). With respect to the distribution of proto-properties, following Blume (2000: 158), the defining characteristic of interaction verbs is that both participants are independently acting proto-agents. This can be illustrated by an example:

(42) *Peter hörte seinem Bruder zu.*

'Peter listened to his brother.'

$\forall x \forall e \forall e^1 \forall e^2$ [listen to(x,y,e) $\rightarrow_P e^1$ [y- ACTIVITY] > $\rightarrow_I e^2$ [x- CONTROL, SENTIENCE]]

In Blume's model, it is assumed that the object participant *seinem Bruder* 'his brother' is autonomously active in a first presupposed subevent (e^1) and that the subject participant *Peter* is entailed to undergo a controlled perception event in the second subevent (e^2). Interaction verbs can be thus said to denote complex[16] events.

The givenness of a true presupposition in case of the first subevent can be verified by the test introduced in (41) above:

[16] The present analysis of interaction verbs as complex events brings counter evidence to the strict constraint made by Kaufmann (1995: 201), stating that verbs can only denote subsequent subevents when these are causally connected: in the case of interaction verbs, the two subevents are in a rather loosely causal, consecutive or concessive relation to each other (Blume 2000: 166f.). For instance, the verb 'contradict' implies that x utters p in e^2 although y is presupposed to have uttered something in e^1 that is incompatible with p. However, as noted by Engelberg (2000: 332f.), these relations are only found when the first subevent is *presupposed*. Hence, the strict constraint may still hold for the case that both subevents are *implied*: E-STR: ($\rightarrow_i e^1$) < CAUSE ($\rightarrow_i e^2$). As for the meaning of the above-mentioned consecutive or concessive relations, Engelberg further hypothesizes that instead of relations between *events* we rather deal with relations between *propositions* here (ibd.: 333).

(43) a. *Peter hörte seinem Bruder zu.* → *Sein Bruder hat etwas gesagt.*
'Peter listened to his brother.' 'His brother said something.'

b. *Hat Peter seinem Bruder zugehört?* → *Sein Bruder hat etwas gesagt.*
'Did Peter listen to his brother?' 'His brother said something.'

c. *Peter hat seinem Bruder nicht zugehört.* → *Sein Bruder hat etwas gesagt.*
'Peter did not listen to his brother.' 'His brother said something.'

Since both arguments in Blume's (2000) analysis are independently acting proto-agents, she has to argue which argument is linked to the nominative (subject) and which to the accusative/dative (object). Recall that Dowty (1991) and Primus (1999a, b), who do not consider presupposed but only implied proto-properties/ basic predicates, unequivocally predict the linking of *Peter* to subject position (nominative) because it is the only argument implying proto-agent properties. Blume (2000: 159) states that although interaction verbs have two independently acting participants, these verbs do not show converse linking patterns. Rather, the opposite seems to be the case: the argument of the implied subevent is always linked to the nominative, while the argument of the presupposed subevent is linked to the accusative/dative. Hence, in an event structure model specified for implied and presupposed subevents, the linking pattern can be captured in the following lexical redundancy rule:

(44) The agent in the first implied subevent of a temporally ordered sequence of subevents in the event structure of a verb is coindexed with the nominative or absolutive. (Blume 1998: 269; cf. also Blume 2000: 224)

The additional question why, from a cross-linguistic perspective, interaction verbs systematically select marked case frames is dealt with further below.

The second innovation of Blume's model is the assumed ranking of the proto-agent properties. She distinguishes between *potent* and *non-potent* proto-agent properties. Under potent proto-agent properties, she subsumes *control* and *activity/function* while *sentience* and *interest*[17] are defined as non-potent proto-agent properties (2000: 137). The motivation behind this ranking is its explanatory potential for argument linking patterns: verbs that solely assign non-potent proto-agent properties to one argument and no properties to the second, such as the German uncontrolled perception verb *schmecken* 'to taste' in (45a), cross-linguistically do not encode the weakly agentive argument in the nominative per default. The same applies to possession verbs like *gehören* 'to belong to' in (45b) which do not entail any proto-properties at

[17] With respect to the newly introduced proto-agent property *interest,* Blume (2000: 138) notes: "'Interesse' hat weder in Primus' noch in Dowtys Modell eine Entsprechung, denn diese Eigenschaft setzt mehr Anteilnahme am Geschehen voraus als bloße willentliche Beteiligung, vgl. Dowtys 'Volition'". It is, for instance, entailed for the initiator of action verbs with benefactives (e.g. German *jemandem etwas waschen* 'to wash something for someone'): Here, the subject participant gets active because s/he wants to do a favour to the object participant. It is further included that the latter should profit from the situation. Note that the proto-agent property *interest* is understood to imply *sentience* (2000: 137).

all to their arguments. Rather, these verb classes allow for multiple encodings, such as the linking of the weakly agentive argument to the dative for the two German verbs mentioned:

(45) a. *Die Steaks schmecken ihnen nicht.* /nom/dat

the steak.NOM taste.PRS.3PL 3PL.DAT NEG

'They do not like the steaks.' (Blume 2000: 142)

$\forall x \forall y \forall e \forall e^{l}$ [schmecken(x,y,e) $\rightarrow_{I}$ e^{l} [x- THEMATIC, y- SENTIENCE]]

b. *Das Haus gehört dem Verein.* /nom/dat

the house.NOM belong to.PRS.3SG the association.DAT

'The house belongs to the association.'

$\forall x \forall y \forall e \forall e^{l}$ [gehören(x,y,e) $\rightarrow_{I}$ e^{l} [x- THEMATIC, y- THEMATIC]][18]

In the case non-potent proto-agent properties co-occur with potent agentive properties for one argument, though, nominative selection can be unequivocally assigned:

(46) *Maria betrachtet das Gemälde.* /nom/acc

PN.NOM look at.PRS.3SG the painting.ACC

'Maria looks at the painting.'

$\forall x \forall y \forall e \forall e^{l}$ [betrachten(x,y,e) $\rightarrow_{I}$ e^{l} [x- CONTROL, SENTIENCE, y- THEMATIC]]

Note that an inner ranking of proto-agent properties had already been considered in previous accounts. Dowty (1991: 574) explicitly leaves the question of weighting the individual proto-properties unanswered, stating only that his proto-agent property *causation* is higher ranked than *movement*. With respect to *sentience*, Dowty (1991: 573) is unsure whether it "should or should not be classed separately from perception". Regarding this point, Engelberg (2000a: 194-198) argues that Dowty's proto-agent property *sentience* should be considered a disjunction of two properties, namely *knowledge* and *perception* (whereas *perception* further stands in complementary distribution with *introspective "perception"*, i.e. emotional states). The subsumption of both properties under the notion of *sentience* is considered an arbitrary grouping being motivated mainly by Dowty's (1991: 576) argument selection principle that is based on the accumulation of proto-properties (Engelberg 2000a: 196). If *sentience* was composed of two disjunct properties, both counted as 1 for argument selection, then Dowty's argument selection principle would not fully apply anymore (e.g. in the case of *to frighten*, the object-experiencer would have one agentive property more than the subject-stimulus). As a straightforward solution to this problem, Engelberg (2000a: 197) proposes – comparable to Blume's (2000) idea – a weighting

[18] The non-potent proto-agent properties *interest* and *sentience* for the dative argument are conveyed by conversational implicature (cf. Blume 2000: 142).

of the different proto-agent properties with the sentience-based properties (*knowledge* and *perception*/*introspective perception*) being of less weight for linking than the others.[19] It is furthermore worth noting that Blume's (2000) assumed role-internal ranking suits well Primus' (2002, 2004, 2006) distinction of minimally (A^{min}, P^{min}) and maximally (A^{max}, P^{max}) involved roles. As has been argued in the previous section (3.2.1), though being formulated in terms of quantity, the cut-off point between minimally and maximally involved roles seems to be more adequately captured as one of quality: for two-place predicates, A^{min} is assigned to exp(x,y) and to poss(x,y) while the other two basic predicates, ctrl(x,s) and phys(x,y), are A^{max} regardless of how many other proto-agent roles they accumulate.

Based on her modified version of the proto-role model, Blume (2000) develops a universal linking principle that is aimed at accounting for the selection of marked valencies. Like Primus (1999a, b), Blume (2000: 200, 221-224) works within Optimality Theory assuming case selection is the result of the competition of formal and semantic violable constraints. In her model, she makes the following three assumptions (47), (49) and (52) (Blume 2000: 200; translation: S.M.): First, there are two constraints for unmarked valencies shown in (47):

(47) Universally unmarked syntactic verb valencies

A verb valency VV of any language L in which ≥ 1 argument positions are formally specified for case is unmarked iff in VV

1. Each argument position is linked to a distinct case

2. An argument position specified for a lower ranked case in L is only included if other semantic argument positions are specified for all higher ranked cases

So, in case of the 2nd unmarkedness constraint in (47), it is supposed that e.g. if in a verb valency dative case is assigned to an argument, there are two other arguments that obtain accusative and nominative case. This constraint has been previously introduced in Primus' (1999a: 24) subcategorization principle. The ranking of cases referred to can be also stated in terms of Primus' (1999a: 2) case hierarchy:

(48) nominative/absolutive <m accusative/ergative <m dative <m other oblique cases

Second, Blume assumes that for marked valencies, the two constraints in (47) are violated to different degrees:

(49) Universally marked syntactic verb valencies:

[19] Note that in more recent approaches, operating with the concept of *prominence* (cf. Himmelmann/Primus 2015 for a definition), it is suggested that neither accumulation nor a fixed ranking of proto-properties is decisive but that the prioritization of certain properties is rather construction-specific. Hence, in this line of argumentation, for each construction, the *prominent* proto-properties have to be identified (as illustrated for German active and passive constructions by Kretzschmar et al. 2019). This view is compatible with Blume's (2000) account: The assumed ranking of *potent* and *non-potent* proto-agent properties should be understood as being modelled in order to account for morpho-syntactic linking and hence serves construction-specific purposes.

> A VV of any language L, in which ≥ 1 valency positions are formally specified for case, is marked exactly when it violates one or more of the unmarkedness conditions for syntactic valencies. The more violations of these conditions VV exhibits, the more marked it is. Thereby violations of the 1st condition have double weight and multiple violations of a condition are counted multiple times.

Accordingly, a violation of the 1st unmarkedness constraint of (47) is given in the case of */acc/acc* valencies, such as for the Icelandic verb *bresta* 'to lack' (50a). This highly marked valency nowadays alternates with a less marked, though still marked, */nom/dat* valency (50b)[20] (Blume 2000: 222f.):

(50) a. *Mig* *brestur* *kjark.* */acc/acc*
1SG.ACC lack.3SG.PRS courage.ACC
'I lack courage.'

b. *Mér* *brestur* *kjarkur.* */nom/dat*
1SG.DAT lack.3SG.PRS courage.NOM
'I lack courage.' (Blume 2000: 222f.)

A violation of the 2nd unmarkedness constraint of (47) can be identified for */nom/dat* valencies, such as for the German verb *gefallen* 'to like' (51a), and to a greater degree, for the antiquated */acc/gen* case frame of German *erbarmen* 'to feel pity' (51b). In Modern German, such highly marked case frames tend to be constantly replaced by less marked ones. In the case of *erbarmen* for instance, this happens by the insertion of an additional valency position, which is linked to the nominative:[21]

(51) a. *Mir* *gefällt* *das* *Buch.* */nom/dat*
1SG.DAT like.PRS.3SG the book.NOM
'I like the book.'

b. *Mich* *erbarmt* *seiner.* */acc/gen*
1SG.ACC feel pity.PRS.3SG 3SG.GEN
'I feel pity for him.' (Blume 2000: 222)

c. *Ich* *erbarme* *mich* *seiner.* */nom/acc/gen*
1SG.NOM feel pity.PRS.1SG 1SG.ACC 3SG.GEN
'I feel pity for him.' (Blume 2000: 223)

[20] The replacement of accusative by dative "subjects" is a well-attested development of Icelandic also referred to as "dative sickness" (cf. Barðdal 2008: 158f. and references therein).
[21] The tendency for nominative experiencers is often described as a Standard Average European (SAE) development (cf. Allen 1995; Dahl 1990; Lazard 1990; Haspelmath 1998; Seefranz-Montag 1983).

Thirdly, and most importantly for this study, the semantic licensing condition for marked valencies is introduced:

(52) Licensing condition for marked two-place syntactic valencies

> A two-place disambiguated verb, which shows all characteristics of semantic transitivity except for the asymmetric role assignment to its realizable arguments and is therefore of restricted semantic transitivity, has marked form properties, which manifest themselves e.g. in a marked two-place syntactic valency.

The notion of semantic transitivity is defined by Blume (2000: 186; translation: S.M.) based on four criteria, which in order to be a verb of high semantic transitivity have to be simultaneously fulfilled:

(53) A disambiguated verb V is semantically transitive if V bears at least one of the following properties:

a. V opens argument positions for at least two distinct arguments arg1 and arg2;

b. V shows the valency relation of involvement to arg1 or arg2;

c. V assigns (potent) proto-role properties to arg1 or arg2;

d. The thematic roles of arg1 and arg2 show a clear asymmetry that manifests itself in the qualitative (or quantitative) distribution of proto-role properties

The more of the criteria (53a-d) apply to a verb, the more semantically transitive it is. Following (52), the decisive property for marked valencies to be licensed is the lack of an asymmetric distribution of proto-role properties (53d). Blume (2000: 190; translation: S.M.) mentions three different constellations of this *non-asymmetric* role assignment:

(54) Non-asymmetric assignment of proto-roles in an argument relation:

1. all realizable arguments show proto-agent properties;

2. all realizable arguments bear the role *thematic*;

3. the arguments partly bear the role *thematic* and partly non-potent proto-agent properties.

I will now illustrate the notion of semantic transitivity by applying the four criteria listed in (53) to seven German verbs, most of them taken from Blume (2000: 186-188). Table 3 indicates for each verb which criteria apply and which case frame the verb selects. For criterion (iv), which must be absent according to the licensing condition for marked case frames in (52), it is further noted which constellation of non-asymmetric role assignment (54) is given.

Verb	(i) two argument positions	(ii) at least one argument is a participant in e	(iii) at least one argument has (potent) proto-properties	(iv) clear asymmetry with respect to the distribution of proto-properties	case frame
töten 'to kill'	+	+	+	+	*/nom/acc*
sich verspäten 'to be late'	-	+	-	-	*/nom/acc*
kosten 'to cost'	+	-	-	- (2.)	*/nom/acc*
helfen 'to help'	+	+	+	- (1.)	*/nom/dat*
gefallen 'to like'	+	+	+	- (3.)	*/nom/dat*
gehören 'to belong to'	+	-	-	- (2.)	*/nom/dat* (but: often copula constructions)
ähneln 'to resemble'	+	-	-	- (2.)	*/nom/dat* (but: often /nom/PP)

Table 3. Semantic transitivity and case frames for a selection of German verbs

The Verb *töten* 'to kill' fulfils all four criteria and can thus be classified as a highly semantically transitive verb that does not meet the licensing condition in (52). As noted by Blume (2000: 191), a high degree of semantic transitivity unilaterally implies the syntactic transitivity of a predicate and requires the selection of an unmarked case frame, such as */nom/acc* for the given verb in German. Likewise, the verbs *sich verspäten* 'to be late' and *kosten* 'to cost' do not license marked case frames, though for different reasons: While for the reflexive verb *sich verspäten* only one semantic participant can be identified, the verb *kosten* does select two semantic arguments but none of it is a participant in the event and bears proto-properties. As predicted, these verbs do not select a marked but an unmarked */nom/acc* case frame. The verbs *helfen* 'to help' and *gefallen* 'to like', by contrast, are the only two verbs that according to (52) license marked valencies. These two verbs, which take a */nom/dat* case frame in German, represent two prototypical verb classes that cross-linguistically show a preference for marked case frames: uncontrolled perception verbs (including non-causative psych verbs) and interaction verbs (Blume 2000: 202-211). As for the first class, the lack of an asymmetric proto-role assignment is due to the fact that one argument (the experiencer) bears the non-potent proto-agent property *sentience* while the other argument only bears the role *thematic*. Since no potent proto-properties are assigned, no predictions for linking can be made, which is cross-linguistically reflected in the selection of converse linking patterns for these verbs. Considering the class of interaction verbs, the role assignment is non-asymmetric because both arguments are assigned (implied or presupposed) proto-agent properties (either potent or non-potent) while none of the arguments bears proto-patient properties. For argument linking, as already mentioned in section 3.2.1, these verbs do not show converse linking patterns

since the argument of the first implied subevent is linked to the nominative per default (Blume 1998: 269, 2000: 224).

As we can see from Table 3, there are two other verbs, *gehören* 'to belong to' and *ähneln* 'to resemble', that also select for a marked */nom/dat* case frame but do not meet the three criteria required in order to be classified as a verb of restricted semantic transitivity. Thus, they are seemingly counterexamples to the licensing condition in (52). Blume (2000: 211-217), however, presents a comprehensible analysis of these verbs arguing that they can be treated as overgeneralizations to the licensing condition and thus do not contradict her model: both the possession verb *gehören* 'to belong to' and the verb denoting an identity relation *ähneln* 'to resemble' do not entail any proto-properties to their arguments. Consequently, in Blume's model, both arguments bear the role *thematic*. Moreover, none of the two arguments is a participant in the event. This is why these verbs are referred to as *minimally transitive verbs*. In the case of possession verbs, Blume draws an analogy to the verbs of uncontrolled spontaneous perception because these verbs implicate *sentience* for the possessor. As for verbs that denote an identity relation, we can identify a completely symmetric argument relation. In Blume's argumentation, the lacking asymmetric distribution of proto-role properties is the most striking characteristic of verbs of limited semantic transitivity, thus the absence of the other characteristics can be compensated (2000: 214).

However, there is another class of verbs, namely *semantically highly transitive verbs* that seems to systematically overgeneralize the licensing condition for marked case frames. This class includes verbs such as 'to touch', 'to pester' and 'to kiss', which are found with marked valencies for example in Finnish or in Czech (Blume 2000: 211). While these verbs do bear an asymmetric assignment of proto-properties and hence fulfil all four criteria of semantic transitivity, they can be shown to have at least one reading in which the object argument is also assigned proto-agent properties. In this reading, they come close to interaction verbs. In Blume's analysis, it is the presence of proto-agent properties that licenses marked case frames for these verbs and that sets up an argument relation that can be described as *pseudo-symmetric* (since one argument also bears proto-patient properties) (2000: 217).

Blume develops her theory based on a data set of agentive two-place verbs with different marked valencies for eleven languages.[22] A special focus thereby lies on the class of interaction verbs. As Blume shows, the degree of "extension" of marked case frames, i.e. the priority of the semantic licensing condition in (52) over the unmarkedness constraints in (47), or vice-versa, is highly language-specific. While in languages like Czech or Finnish, marked valencies are strongly preferred for verbs of restricted semantic transitivity, a language like Romanian rarely violates the unmarkedness constraints and thus prefers unmarked case frames. Consequently, it can

[22] These are: German, Finnish, Icelandic, Latin, Polish, Romanian, Czech, Hungarian, Maori, Samoan and Tongan (Blume 2000: 233-255).

be expected that overgeneralizations of the licensing condition show up in Czech or Finnish but not in Romanian (Blume 2000: 200f.).

Blume's licensing condition for marked case frames shows two main advantages in comparison to Primus' (1999a: 66) dative default: Firstly, it is more suitable as a linking principle since it takes the distribution of proto-role properties as a starting point in order to make predictions on case selection. Recall that Primus' dative default was formulated in the reverse direction, namely from dative encoding to role-semantic properties. Certainly, the restricted semantic transitivity of a verb does not impose it to select a marked valency. Rather, it is due to language-specific preferences if the semantic licensing condition has priority over the unmarkedness constraints mentioned in (47). As shown by Blume (1998: 274f.), in a data sample of 236 interaction verbs from seven languages with all of them satisfying the licensing condition for marked case frames, only 64% (151) actually selected a */nom/dat* or */abs/dat*, i.e. a marked valency, whereas 19% (45) were attested with unmarked case frames (the rest appeared with PPs or adverbials or were realized as a three-place predicate). Secondly, Blume's licensing condition turns out to be valid for a broader range of languages than Primus' dative default: "Entgegen der verbreiteten Annahme, daß die Kodierung des Experiencers die zentrale Funktion des Dativs ist […], gibt es nämlich auch Sprachen, in denen der Stimulus statischer Wahrnehmungsverben mit /Nom/Dat- oder /Abs/Dat-Valenz im Dativ erscheint" (Blume 2000: 203). This linking pattern can be observed in Austronesian languages such as Samoan and Tongan. However, in one and the same language it is restricted if either the experiencer or the stimulus constantly selects the dative (Blume 2000: 228).[23] Hence, in this regard, Primus' (1999a: 66) dative default still holds true for a characterization of arguments with dative encodings in languages like German, Spanish, Russian or Korean.

Kailuweit (2004: 95) criticizes Blume's linking principle for the following reason: Concerning for example non-causative psych verbs in German, it predicts that a linking of the experiencer to the nominative (*lieben* 'to love') is equally good as to the dative (*gefallen* 'to like') and to the accusative (*interessieren* 'to interest'). In his view, this does not adequately capture linking preferences in German and Romance languages:

> Aus der grundsätzlichen Irrelevanz der PAE [Proto-Agens-Eigenschaft] "Bewusstheit" für das Linking folgt, dass nicht-kausative Gefühlsverben mit Nominativ- (*lieben*), Dativ- (*gefallen*) und Akkusativ-Experiencer (*interessieren*) in gleichem Maße den Linkingregeln entsprechen (2000: 151). Dies ist unbefriedigend, da für nicht-kausative Gefühlsverben eine Grundabfolge Experiencer-Stimulus anzunehmen ist (cf. Primus 2004), was sich auch darin zeigt, dass Nominativ-Experiencer-Verben im Deutschen und Romanischen eine größere Gruppe bilden als Dativ- oder nicht-kausative Akkusativ-Experiencer Verben. (Kailuweit 2004: 95)

[23] It must be noted that there are a few exceptions to this generalization: the German psych verbs *zugetan sein* 'to be fond of', *entgegenfiebern* 'to look forward to' or *entgegenbangen* 'to fear for' link the stimulus to the dative and the experiencer to the nominative (cf. Primus 2004: 391; Kailuweit 2005: 127, fn. 44).

I do agree with Kailuweit in that Blume's linking principle does not allow for precise predictions about the likelihood of the different case frames for non-causative psych verbs. Yet Blume (2000: 232) herself mentions that her linking principle is deliberately less restrictive than other approaches in order to account for the range of converse linking patterns attested cross-linguistically for uncontrolled perception verbs, which include non-causative psych verbs. Further restrictions, such as the above-mentioned preference of the linear order "experiencer before stimulus", have been shown to be not universally valid but language-specific depending on lexical properties of the psych verbs in question (Verhoeven 2014). Beyond this, as illustrated by Klein/Kutscher (2002) for German psych verbs, case frame selection may also be determined by lexical economy: As they show, many psych verbs in Contemporary German are either polysemous with a highly transitive physical verb (e.g. *ausrasten* 'to freak out' vs. *ausrasten* 'to come out') or from a diachronic perspective can be traced back to a physical reading of the same predicate (e.g. *hassen* 'to hate', which goes back to 'to chase sb. due to hostile feelings'). In their argumentation, it is the non-psychological reading or origin of a psych verb that determines its case frame satisfying the "principle of Lexial Economy"[24] (Klein/Kutscher 2002: 20). Hence, for a comprehensive understanding of linking patterns in a given language, one might have to take into account the diachronic development of a verbal lexeme as well. Blume's universal linking principle, by contrast, "just restricts the possible valency frames for verbs with a given meaning, correctly predicting that certain verbs can assume a */nom/dat* pattern in one language and a */nom/acc* pattern in another, yet still accounting for the non-arbitrariness of the */nom/dat* (or any other marked) valency option" (Engelberg 2006: 260).

In conclusion, Blume's (1998, 2000) model offers the following innovations in comparison to Primus' (1999a, b) model: First, the introduction of an event structure model that distinguishes implied from presupposed subevents leads to a more precise description of interaction verbs than previously formulated proto-role models lacking the reference to the temporal structure of a predicate. Second, Blume's dissociation of potent and non-potent proto-agent properties is highly relevant for the explanation of converse linking patterns of uncontrolled perception verbs (and possession verbs). Here, due to the lack of potent proto-agent properties, no argument is linked to the nominative per default. Third, Blume's definition of proto-patient properties is more adequate than Primus' co-argument dependency relation in that it is not part of *every* proto-agent role. Simultaneously, the idea that proto-patient properties can only

[24] [(67)] Principle of Lexical Economy:
Lexical entries are as simple as possible.
[(68)] Corollary 1 to the Principle of Lexical Economy:
Each verbal lexeme has only one case frame. This case frame holds for each reading of the lexeme and must therefore be compatible to all of its readings.
[(70)] Corollary 2 to the Principle of Lexical Economy:
For each verbal lexeme exists one reading which is decisive for the selection of its case frame. It is the reading putting the strongest restrictions on the selection of the case frame. This reading is called the "strongest reading" of a verb. Any case frame that is compatible with the strongest reading of a verb is compatible with any other reading, too. (Klein/Kutscher 2002: 20f.)

appear as the converse of corresponding proto-agent properties is upheld as an important novelty of Primus' model.

With these premises made, Blume establishes a linking principle for marked case frames that consists of a competitive model, comparable to the OT-model in Primus (1999a, b). In her model, marked case frames are licensed in case of limited semantic transitivity. This condition applies both to uncontrolled perception verbs and to interaction verbs, which, from a cross-linguistic perspective, both systematically select marked case frames (*/nom/dat* or */abs/dat*). Blume's OT-model can account for a wide range of languages. Thereby, it is language-specific if the unmarked constraint or the marked constraint (limited semantic transitivity) is higher ranked. Yet Blume does not consider differential morphosyntactic argument realization. This latter point was first addressed by García García (2007, 2014) for DOM in Spanish and will be explored in the following section.

3.3 Application to differential morphosyntactic argument realization: DOM in Spanish (García García 2007, 2014)

The linking principles I have discussed so far in this chapter all considered syntactic argument selection or morphosyntactic case selection. In the following, I will demonstrate that they are also suitable to account for differential morphosyntactic argument realization. Among the role-semantic approaches on DOM that have been put forward in the past decades, García García (2007, 2014) is the first one to make use of Dowty's (1991) proto-role model and Primus' (1999a, b, 2006) modification of it in order to predict DOM in Spanish. The point of departure for his theory is the finding that with certain verb classes, DOM in Spanish systematically occurs with inanimate objects. These occurrences, as shown with the reversible predicate *suceder* 'to succeed' in (55a) and with the attribution verb *caracterizar* 'to characterize' in (55b), cannot be captured under an animacy- and definiteness-based account: the relevant object NPs are not only inanimate but, as can be seen in (55a), may also refer to a non-specific indefinite object.[25]

(55) a. *Un sustantivo sucede* ****ø/a*** *un artículo.*
a noun succeed.PRS.3SG ø/DOM a article
'A noun comes after an article.'

b. *La euforia caracteriza* ***??ø/a*** *la situación.*
the euphoria characterize.PRS.3SG ø/DOM the situation
'Euphoria characterizes the situation.' (García García 2018a: 228, 230)

[25] It has to be noted, though, that the indefinite object in (55a) has a generic reading. In contrast to other non-specific indefinite NPs, generics are said to rank high in the referentiality scale, being sometimes compared to proper names, i.a. due to their rigid reference to an entire class (Carlson 1977).

The occurrence of DOM with inanimate objects, which tends to be obligatory rather than only optional, can mainly be detected for the following verb classes (García García 2018a: 227):

(56) DOM-sensitive verb classes:

a. Verbs of sequencing (e.g. *preceder* 'to precede', *suceder* 'to succeed')

b. Verbs of replacement (e.g. *sustituir* 'to substitute', *reemplazar* 'to replace')

c. Verbs of competition (e.g. *vencer* 'to win', *derrotar* 'to defeat')

d. Verbs of attribution (e.g. *caracterizar* 'to characterize', *definir* 'to define')

e. Verbs of naming (e.g. *considerar* 'to consider', *llamar* 'to call')

The innovation of García García's account lies in the fact that it, as Dowty's (1991: 576) argument selection principle, considers the accumulation of proto-agent and proto-patient properties of both subject and object to predict the morphosyntactic form of the direct object argument. This deviates from Primus' (2012a: 81) extended dative default, a revised version of the dative default discussed in section 3.2.1:

(57) The dative is used for an argument with a low number of proto-agent properties that are i) subcategorized by the verb or ii) assigned according to the intrinsic meaning of the direct object referent.

One problem of the dative default, as stated by García García (2014: 140) referring to Kailuweit (2004: 91), is that it does not qualify as a linking principle: although it nicely connects the common function of indirect objects and marked direct objects in Spanish in that both express the – subcategorized or intrinsic – proto-agent role of the given argument, it does not tell which agentive arguments are realized as indirect and which as marked direct objects. With respect to the latter, it further does not specify whether DOM is necessary or only optional. Moreover, it does not account for the ungrammaticality of DOM in ditransitive constructions. This can be illustrated by the following five examples in (58), which can be all subsumed under the generalization in (57):

(58) a. Subcategorized proto-agent properties (human indirect object, ditransitive verb):

*La madre le da **el/*al** niño ***el/al** padre.*

the mother CL.3SG.DAT give.PRS.3SG the/DOM.the child the/IO.the father

'The mother gives the child to the father.' (Primus 2012a: 67)

b. Subcategorized proto-agent properties (human indirect object, two-place verb):

***ø/A** Pepe le gusta la musica.*

ø/IO PN CL.3SG.DAT like.PRS.3SG the music

'Pepe likes the music.'

c. Subcategorized proto-agent properties (human direct object, transitive verb):

Pepe asusta ****ø/a*** *la mujer.*

PN frighten.PRS.3SG ø/DOM the woman

'Pepe frightens the woman.'

d. Subcategorized proto-agent properties (inanimate direct object, transitive verb):

Los días siguen ****ø/a*** *las noches.*

the days follow.PRS.3PL ø/DOM the nights

'The days follow the nights.'

e. Intrinsic proto-agent properties (human direct object, transitive verb):

Pepe ve ****ø/a*** *la mujer.*

PN see.PRS.3SG ø/DOM the woman

'Pepe sees the woman.'

Crucially, in case of a co-occurrence of (57i) and (57ii), as in the case of ditransitive constructions, the subcategorized proto-agent properties are said to be prior to the intrinsic ones (cf. Primus 2012a: 81, fn. 8). This is illustrated by (58a) where the intrinsic proto-agent argument cannot be marked "in favour of the subcategorized recipient" (Primus 2012a: 81, fn. 8). Note that a higher ranking of (57i) with respect to (57ii) would also suggest a difference between (58c) and (58e): while the verb *asustar* 'to frighten' subcategorizes *sentience* for its object argument, the object of *ver* 'to see' can only be a 'proto-agent' due to the intrinsic meaning of *la mujer* 'the woman'. However, such difference cannot be captured in terms of grammaticality, at least not for Spanish, where DOM with human definite objects has reached a high degree of grammaticalization.

Given the variety of morphosyntactic phenomena to which the extend dative default applies in Spanish, García García (2007, 2014) aims at formulating a more precise linking principle in order to account for the obligatory *a*-marking of inanimate objects in cases such as (58d) above. For this purpose, he introduces the generalization of thematic distinctness (García García 2007: 71, 2014: 145):

(59) Unbelebte Objekte mit *a*-Markierung (Hypothese des Agentivitätsgefälles):

> Weist das DO die gleiche oder eine höhere Anzahl an Proto-Agenseigenschaften als das S auf, so erhält das DO eine *a*-Markierung.

It is predicted that DOM must occur with inanimate objects in Spanish when the subject does not outrank the object in terms of agentivity. García García (2014: 146) crosschecks his hypothesis with a further generalization formulated in (60):

(60) Unbelebte Objekte ohne *a*-Markierung:

> Weisen das S eine maximale Anzahl an Proto-Agenseigenschaften und das DO eine maximale Anzahl an Proto-Patienseigenschaften auf, so erhält das DO keine *a*-Markierung.

The second hypothesis is argued to be a sort of logical proof of the first one: if *a*-marked inanimate objects in Spanish occur in case of a deviation from the prototypical distinctness of subject and direct object, then those cases in which thematic distinctness between the two arguments is fulfilled should not allow an *a*-marking of the direct object.

What is crucial for the present account is its understanding of agentivity: based on the co-argument dependency model of Primus (1999a, b, 2006: 55-60), agentivity is defined in a broad sense. So not only the involvement dimension, which addresses the distribution of proto-agent and proto-patient basic roles to the arguments, is included, but also the co-argument dependency of the arguments. Following Primus (2006: 55-60), co-argument dependency is the main criterion that distinguishes a proto-agent from a proto-patient: a proto-agent is always causally independent from a proto-patient and the proto-patient is always defined by its dependency from the proto-agent (Proto-agent $>_{dep}$ Proto-patient). If no proto-agent basic roles are assigned to any argument, which means that there is no involvement relation given, a co-argument dependency relation between the arguments cannot be established. This lack of co-argument dependency means that both arguments qualify as weak proto-agents (A^{min}). This can be illustrated using two examples with reversible predicates, a verb of sequencing in (61a) and a verb of replacement in (61b):

(61) a. *El carnaval precede* ***ø/a*** *la Cuaresma.*
the carnival precede.PRS.3SG ø/DOM the Lent
'Carnival precedes Lent.' (García García 2014: 2)

b. *En este cóctel el vodka puede sustituir* ***ø/a*** *la ginebra.*
in this cocktail the vodka can.PRS.3SG substitute.INF ø/DOM the gin
'In this cocktail, vodka can be substituted by gin.' (García García 2018a: 226)

In (61a) the verb *preceder* 'to precede' expresses a temporal ordering relation between the two arguments *el carnaval* 'carnival' and *la Cuaresma* 'Lent'. In this relation, none of the arguments is defined by its dependency from the other. As shown by García García (2018a: 228), this can be proved by the expression of the same truth-conditional meaning with the converse predicate *suceder* 'to come after' where subject and direct object change position: *La Cuaresma sucede *ø/al carnaval* 'Lent comes after carnival'. In (61b), the situation is comparable in that the verb of

replacement *sustituir* ‘to substitute’ sets up a relation between the two arguments *el vodka* ‘vodka’ and *la ginebra* ‘gin’ that lacks co-argument dependency. In contrast to *preceder* ‘to precede’, this relation is none of temporal ordering but rather located in the spatial domain. Both gin and vodka can fulfil the same function as ingredients for obtaining the cocktail in question (García García 2018a: 230f.). Again, the lack of co-argument dependency can be checked by the fact that both arguments are interchangeable without imposing a change in truth-conditions (*En este cóctel la ginebra puede sustituir *ø/al vodka.)*. Thus, the preference of reversible predicates to take DOM can be given a functional explanation by the need to disambiguate subject and direct object, which show the same degree of agentivity. The class of attribution verbs, such as *caracterizar* ‘to characterize’, presented in the initial example *La euforia caracteriza ??ø/a la situación* in (55b), deviates even more from the prototypical agent-patient asymmetry of a transitive sentence. Here, the direct object is more agentive than the subject. The latter cannot even be ascribed argument status: it rather serves as a predicate in the given situation. The object, by contrast, qualifies as a weak proto-agent since it bears co-argument independency (García García 2014: 171f.).

There are several pieces of evidence that support the presented generalization of thematic distinctness as a linking principle for *a*-marking in Spanish. Firstly, I suggest that the scope of the generalization can even be extended to non-differential morpho-syntactic argument realization in Spanish, namely to indirect objects of psych verbs such as *gustar* ‘to like’ and of possession verbs such as *pertenecer* ‘to belong to’. Consider the following two examples:

(62) a. *A María le gusta la casa.*
IO PN CL.3SG.DAT like.PRS.3SG the house
‘Maria likes the house.’

b. *A María le pertenece la casa.*
IO PN CL.3SG.DAT belong to.PRS.3SG the house
‘The house belongs to Maria.’

In both sentences, the indirect object outranks the subject in terms of agentivity: in terms of Dowty’s (1991) proto-role model, it bears the proto-agent properties *sentience*[26] and *independent existence* whereas the subject only bears *independent existence*. Proto-patient properties are not assigned. In accordance with the generalization of thematic distinctness, the object *María* is *a*-marked. This reveals the connection

[26] In the case of possession verbs, it can be discussed if they bear *sentience*. Probably it is more a conversational implicature as it can be cancelled in certain contexts, such as*: A María le pertenece la casa, pero no lo sabrá hasta que se abra el testament* ‘Maria owns the house, but she won’t know that until the will is opened’ (cf. also Blume 2000: 212). In that case, however, subject and direct object would still bear the same degree of agentivity (both have *independent existence*) and thus correspond to the prediction the generalization of thematic distinctness makes.

between agentivity and *a*-marking in Spanish grammar: thematic distinctness seems to have an impact on the morphosyntactic expression of an object beyond syntactic transitivity. This is also in line with the extended dative default of Primus (2012a: 81) assuming a common semantic role of *a*-marked intransitive and transitive objects in Spanish. I will take up this issue in section 4.1. Yet one could raise the question why the DOM-sensitive verb classes in Spanish are encoded as transitive verbs at all, since they show a sharp deviation from the prototypical agent-patient asymmetry. García García (2014) proposes an answer to this question which follows from a comparison of different readings that the predicates in question show: it becomes evident that verbs like *preceder* 'to precede', *suceder* 'to succeed, *seguir* 'to follow', *sustituir* 'to substitute' etc., besides their reversible reading in which co-argument dependency is absent, also have at least one more – main – reading in which a clear-cut agent-patient asymmetry can be attested (for an overview cf. García García 2014: 170). In accordance with the generalization of thematic distinctness, the insertion of *a*-marking is not possible in the latter role-semantic configuration.

Secondly, further evidence for García García's (2014) approach comes from Spanish verbs that always show a maximal degree of thematic distinctness in that the subject bears a high number of proto-agent and the direct object, conversely, all possible proto-patient properties. Relevant verb classes are consumption, destruction and creation verbs. Based on a corpus study in the ADESSE comprising 2,012 sentences with verbs from these three classes, it is demonstrated that these do not show *a*-marking in 99.7 % of the cases (2014: 197). Moreover, a further clue for the blocking of DOM with maximal proto-patient arguments is the fact that even left dislocated, topicalized objects cannot take the marker without making the sentence ungrammatical. Usually, the left periphery is a position that strongly favours DOM, in particular when both subject and object are inanimate entities as in (63a) and (63b) below:

(63) a. Creation verb:

El/*Al *problema* *más* *grave* *lo* *ha*

the/DOM.the problem most serious CL.M.3SG.ACC have.PRS.3SG

creado *la* *ignorancia.*

create.PTCP the ignorance

'The most serious problem has been created by ignorance.'

b. Destruction verb:

La/*A ***la*** *casa* *la* *ha* *derrumbado* *la*

the/DOM the house CL.F.3SG.ACC have.PRS.3SG tear down.PTCP the

tormenta.

storm

‘The house was torn down by the storm.’

c. Consumption verb:

La/*A	***la***	*manzana*	*(se)*	*la*	*comió*	*Pepe.*
the/DOM	the	apple	REFL.3	CL.F.3SG.ACC	eat.PST.3SG	PN

‘Pepe ate the apple.’ (García García 2014: 197f.)

This finding calls into question the indexing approach as a functional explanation for DOM based on Hopper & Thompsons’ (1980) understanding of transitivity (cf. García García 2014: 198f.). Apparently, telic and affected inanimate objects do not allow for DOM. This is an important piece of counterevidence against the view that telicity and affectedness are factors directly influencing DOM in Spanish (cf. e.g. Torrego 1999; Næss 2004, 2007; von Heusinger/Kaiser 2011). Rather, it suggests that an impact of these factors, if at all, is limited to cases where they interact with the object’s animacy. Note that this would be still in line with Hopper & Thompsons’ (1980) concept of transitivity since they assume an increase in transitivity of a sentence if an object is highly individuated, including i.a. its animacy and referentiality. This is reflected by the morphosyntactic encoding of Spanish direct objects, which are typically *a*-marked in the case of a human referential participant and unmarked in the case of a human non-referential or an inanimate participant (Hopper/Thompson 1980: 256). In contrast, Primus (1999a: 27f., 57ff.) and Blume (2000: 181ff.) argue that the morphosyntactic marking of an argument, as compared to an argument that stays unmarked, rather points to a lower than to a higher degree of semantic transitivity. This assumption results from a different understanding of *transitivity*: Unlike Hopper/Thompson (1980), who aim at replacing a syntactic notion of transitivity with a (multifactorial) semantic one, Primus and Blume clearly differentiate between these two dimensions. (Morpho)syntactic and semantic transitivity are understood as two independent notions overlapping only in one configuration, as stated by Blume (2000: 191):

> Die semantische Transitivität eines Verbs ist unabhängig von seiner syntaktischen Transitivität, wobei allerdings starke semantische Transitivität syntaktische Transitivität einseitig impliziert. D.h., daß Verben, die alle Merkmale semantisch transitiver Verben aufweisen […], sich immer auch durch alle Merkmale syntaktischer Transitivität auszeichnen, daß sie also mehrstellige unmarkierte Valenzen haben […].

In Primus’ approach, (morpho)syntactic transitivity is captured by her case hierarchy, shown again for illustrative purposes (1999a: 2):

(64) nominative/absolutive <m accusative/ergative <m dative <m other oblique cases

It is assumed that for a two-place verb with one argument encoded in the nominative (/absolutive), the morphosyntactic transitivity of the predicate increases the higher the second argument is ranked in the case hierarchy (i.e */nom/dat* verbs are of lower morphosyntactic transitivity than */nom/acc* verbs) (1999a: 28). As noted by Primus, although valid for many languages, the hierarchy is not universal. For Spanish,

lacking morphological case but showing differential morphological marking of direct objects and non-differential marking of indirect objects, we might formulate the following versions of the case hierarchy:

(65) a. ø-marking <m *a*-marked direct object <m *a*-marked indirect object <m prepositional marking (*a*, *con* etc.)

b. ø-marking <m *a*-marking <m prepositional marking (*a*, *con* etc.)

I will not discuss here which version (65a) or (65b) is more appropriate. However, in both cases *a*-marked objects would be of lower morphosyntactic transitivity than ø-marked ones. Accordingly, Primus argues that in the case of DOM we would deal with a decrease rather than with an increase of morphosyntactic transitivity (Primus 1999a: 58).[27] In the remainder of this work, I adopt the view that (morpho)syntactic transitivity should be differentiated from the semantic concept.

There is another point left to be made in favour of García García's (2014) approach. So far, the applicability of the generalization of thematic distinctness as a linking principle on DOM has only been proved for inanimate objects. Now, the question of how *a*-marked animate objects enter into the picture has to be addressed. Consider the two following examples:

(66) a. *Pepe conoce* ****ø/a*** *esta actriz.*

PN know.PRS.3SG ø/DOM this actress

'Pepe knows this actress.'

b. *Pepe conoce* ***ø/*a*** *este libro.*

PN know.PRS.3SG ø/DOM this book

'Pepe knows this book.'

In (66a) we see that the human definite direct object *esta actriz* 'this actress' obligatorily takes DOM whereas the marker would be ungrammatical with the inanimate direct object *este libro* 'this book' in (66b). The predicate, in both cases *conocer* 'to know', entails *sentience* for its subject while the object is assigned no proto-properties at all. Evidently, the generalization of thematic distinctness is not applicable in a strict sense here: the subject outranks the object in terms of agentivity. As argued by García García (2014: 129ff.) and Primus (2012a), although animate objects like in (66a) cannot be classified as proto-agents, they are *potential* proto-agents. These are defined as "Partizipanten, die in der Lage sind, das vom Verb denotierte Ereignis herbeizuführen" (García García 2014: 133f.). This is proved for the object *esta actriz* 'this actress' of (66a) above, which is potentially sentient in the given event of

[27] Note, though, that Primus (1999a: 58) seems to subsume Spanish and Romanian DOM under *adpositional* marking. This is misleading since in both cases, whether marked or unmarked, we deal with a proper direct object argument.

knowing, but not for the object *este libro* ‘this book’ of (66b) above, which cannot fulfil this function:

(67) a. *Esta actriz conoce ***ø/a** Pepe.*
this actress know.PRS.3SG ø/DOM PN
‘This actress knows Pepe.’

b. *#Este libro conoce ***ø/a** Pepe.*
this book know.PRS.3SG ø/DOM PN
‘#This book knows Pepe.’

Importantly, the notion of potential agentivity is understood as being independent of the animacy of a participant. Nonetheless, García García (2014: 134) and Primus (2012a: 84) assume that the following inference from animate to potential agentive participants can be drawn: “Argumente, die sich auf belebte bzw. menschliche Referenten beziehen, werden als Partizipanten verstanden, die ein Ereignis potenziell kontrollieren, wahrnehmen oder über einen anderen Partizipanten verfügen können” (García García 2014: 134). Primus (2012a: 85) motivates such an inference by abductive reasoning, referring to Levinson (2000). Since, in her argumentation (2012a: 85, 1999a: 50), the proto-agent properties *sentience* and *volition* (*control*) in the prototypical cases unilaterally imply the animacy of the participant in question, it is plausible to assume the reverse, too, namely that an animate participant is interpreted to potentially bear the mentioned agentive properties in a given event. However, this inference has to be treated with caution. As we have seen in section 3.1, a unilateral implication from the proto-agent properties of *sentience* and *volition* (or *control*) to animacy (or humanness) does not apply in a strictly relational understanding of proto-properties. Instead, the link between agentivity and animacy seems to be weaker and, as I proposed, can be formulated in the following way: ‘An argument slot of a given predicate assigning proto-agent properties to its participant can be context-independently filled by human beings’.

Yet there is another aspect that needs further clarification: there seem to be different levels of potential agentivity for an object argument. On the one hand, potential agentivity is understood with regard to the (hypothetical) ability of the object participant to bear the proto-agent properties assigned to the subject in the given event. This case can be illustrated by (66a) and (67a) above, where the former object *la actriz* ‘the actress’ can fulfil the subject role of *conocer* ‘to know’ since it can have *sentience*, too. On the other hand, it can be identified as an understanding of potential agentivity that rather concerns the relation between the object and the event in question. I will illustrate this with the examples in (68):

(68) a. *Pepe golpeó ***ø/a** la actriz.*
PN hit.PST.3SG ø/DOM the actress

‘Pepe hit the actress.’

b. *Pepe golpeó* ***ø/*a*** *la mesa.*

PN hit.PST.3SG ø/DOM the table

‘Pepe hit the table.’

As already mentioned in section 3.2.1 with reference to Dowty (1991: 572), there are verbs entailing proto-patient properties to its direct object that in combination with an animate or human object typically evoke the interpretation that the participant is sentient with respect to the event (Dowty uses an example with the verb *to slap*). I classified this as a conversational implicature given that it can easily be cancelled in context (by adding e.g. *pero ella no lo notó* ‘but she did not notice it’ in (68a)). With regard to the (68b), *la mesa* ‘the table’ cannot be interpreted to be sentient in the denoted event. Hence, there is a difference in agentivity that exclusively concerns the two direct objects in (68a) and (68b). Of course, *la actriz* ‘the actress’ could also be interpreted as a potential proto-agent in that it would be able to take the subject role, which *la mesa* ‘the table’ cannot. Does this ‘double’ potential agentivity make the object in (68a) even more potentially agentive? To answer this question, we need a more fine-grained definition of potential agentivity in a first step. Especially, clear testable criteria must be elaborated that specify how good or less good potential proto-agents can be identified.

For the purpose of the present section, it suffices to stay with the first notion of potential agentivity. Based on this assumption, Primus (2012a: 75), following García García (2007, 2014), establishes the following working hypothesis in order to account for DOM with animate objects in Spanish:

(69) A differential object marker is licensed by an object whose intrinsic meaning properties qualifies it as a proto-agent in the situation denoted by the predicate.

So, while in the case of DOM with inanimate objects we deal with an *actual* need for disambiguation since subject and direct object are not distinct enough in terms of their semantic roles, in the case of DOM with animate objects the need for disambiguation is merely *potential* and more abstract: due to the intrinsic agent-properties of an animate or human participant it could be interpreted as agentive in the given event. In case of subcategorized proto-patient properties for an animate object participant this could lead to a “role overlap” of proto-patient and potential proto-agent properties (Primus 2012a: 73). This blurs the thematic distinctness between subject and direct object: “the animacy of the proto-patient leads to a situation in which asymmetric co-argument dependency cannot be established unequivocally” (2012a: 74).

In conclusion, Dowty’s (1991) model of generalized semantic roles and Primus’ modification of it (1999a, b, 2006) can be used to account for differential morpho-syntactic argument realization in Spanish. More precisely, García García (2007, 2014) explains the systematic occurrences of certain verb classes with *a*-marked inanimate objects by a lack of thematic distinctness between subject and direct object

argument. Such an agentivity-based explanation also looks very promising with regard to DOM with animate objects. However, more theoretical work is needed in this respect, especially to clarify the notion of potential agentivity and to formulate predictions on different verb classes for their preference to get DOM. Though García García's approach is only based on Dowty's (1991) and Primus' (1999a, b, 2006) work, it is perfectly compatible with the ideas developed in Blume (1998, 2000) as will be shown in section 4.1.

3.4 Summary

In this chapter we have seen that a model of generalized semantic roles, such as Dowty's (1991) proto-role model and its modifications, serve as a basis for a unified account on different argument realization phenomena. These cover the following three dimensions:

(70) a. Case frame (ACC, DAT…)

b. Subject vs. DO

c. ø DO vs. marked DO

The three dimensions are ordered from general to more specific argument realization: (70a) concerns the encoding of a predicate as transitive or intransitive, and under the premise a language has morphosyntactic case, it determines the case frame for a given verb (cf. Primus 1999a, b; Blume 2000). Argument realization in (70b) captures the more specific case of syntactic argument realization for a transitive verb and thus corresponds to Dowty's (1991) argument selection principle, i.e. the selection of subject and direct object. In (70c) it is further zoomed in on the morphosyntactic encoding of direct objects; here DOM, as exemplified for Spanish, is concerned.

As for morphosyntactic linking in (70a), it has been argued that the choice of marked case frames such as */nom/dat* can be adequately explained by the limited semantic transitivity of the verbs in question (Blume 2000). With regard to the inventory of proto-agent properties, I follow Blume in distinguishing potent proto-agent properties (*volition*, *causation*, *movement*) from non-potent proto-agent properties (*sentience*). Making this distinction, one can also better account for the construction variance of (non-causative) psychological predicates in dimension (70b) than Dowty (1991) did: When a predicate only entails non-potent proto-agent properties to its participants, argument linking cannot be established unequivocally. Thus, for a transitive verb, both the subject and the direct object could be linked to the stimulus or the experiencer, respectively. Concerning the dimension of differential morphosyntactic argument realization (70c), we have seen that thematic distinctness between subject and direct object argument serves as an explanation for DOM with inanimate objects in Spanish (García García 2007, 2014): these are *a*-marked when the subject does not outrank the direct object in terms of agentivity. We have further recognized that in order to apply this generalization to DOM with animate objects in Spanish we need a notion of *potential* agentivity (García García 2014; Primus 2012a). Yet this

notion needs further clarification: in particular, whether potential agentivity of the object refers both to *inherent* and *relational* agentivity. Furthermore, as Primus' (2012a: 81) extended dative default suggests, there are at least two different forms in which agentivity information can be assigned to a direct object: on the one hand, proto-agent properties are subcategorized by the verb and on the other hand, they are acquired due to the intrinsic properties of the object referent. In the following chapter, I will further elaborate on the notion of agentivity and propose a scale that captures different types of agentivity.

4. Communicated agentivity for the object and DOM: Definition of verb classes

We have seen that models operating with generalized semantic roles offer explanatory potential for various dimensions of argument realization phenomena. Starting from Dowty's (1991) proto-role model, introduced to account for argument selection of transitive predicates in English, Primus (1999a, b, 2006) and Blume (1998, 2000) suggest a cross-linguistic applicability to morphosyntactic linking. As for differential morphosyntactic argument realization, García García (2007, 2014) argues that DOM with inanimate objects in Spanish is determined by the thematic distinctness of subject and direct object. The *a*-marker is thereby predicted to appear when the subject does not outrank the object in terms of agentivity. We have further seen that this explanation also looks promising for DOM with animate objects in Spanish (García García 2014; Primus 2012a). However, in order to make more fine-grained predictions for animate objects of different verb classes to get *a*-marked, we need to take a closer look at different levels of agentivity and clearly dissociate it from animacy. The aim of this chapter is twofold. As a first step, the concept of agentivity will be stated in more precise terms. As has been already pointed out in the previous chapter, a predicate can encode agentivity information for an object in different ways: It can *entail* its agentivity, like for example *sentience* for the experiencer object of psych verbs. Furthermore, it can *presuppose* the object's agentivity in a prior subevent, as has been shown for interaction verbs (Blume 1998, 2000). A third option is the object's *potential* agentivity stating that the intrinsic meaning properties of an object qualify it to act as a proto-agent in the event denoted by the predicate (Primus 2012a). In the following I will strengthen the hypothesis that, taking role-semantic configurations into account, non-differential and differential morphosyntactic argument realization are closely connected by the example of *a*-marked indirect and direct objects in Spanish (4.1). The subchapter brings together Blume's (1998, 2000) and García García's (2007, 2014) approach showing that they make similar predictions. Then, I will present a number of diagnostics that should help to delimit different types a predicate can communicate agentivity for its object (4.2). These can be basically subsumed under four types of communicated agentivity for the object, which are further proposed to be arranged on a four-point agentivity scale. As a second step, a connection between these fine-grained levels of agentivity and DOM is established. Accordingly, for each stage of the hierarchy, relevant verb classes can be defined and their preference for DOM will be documented across Romance languages (4.3). On this basis, a working hypothesis for the relevant verb classes to get DOM is presented (4.4).

4.1 Agentivity and *a*-marking in Spanish

As we have seen in section 3.3, the generalization of thematic distinctness (García García 2007: 71, 2014: 145) allows for reliable predictions on DOM with inanimate objects in Spanish. The *a*-marker is thereby expected to occur obligatorily when the

thematic relation between subject and direct objects deviates from the prototypical agent-patient asymmetry, that is when the direct object is (i) equally agentive as or (ii) even more agentive than the subject. This linking principle for DOM with inanimate objects in Spanish can be compared to Blume's (2000: 200) licensing condition for marked valencies according to which marked case frames are licensed by the restricted semantic transitivity of a predicate. The decisive criterion for restricted semantic transitivity is the non-asymmetric assignment of proto-roles between the two arguments. It is given when (i) both arguments bear proto-agent but no proto-patient properties, (ii) the role *thematic* and non-potent proto-agent properties are assigned or when (iii) both arguments bear the role *thematic*. Thus, both DOM in Spanish and the cross-linguistic selection of marked valencies can be argued to depend on the same role-semantic configuration, namely a deviation from the prototypical asymmetric assignment of proto-roles in a transitive event.

As already mentioned in section 3.3, this close connection between DOM and marked valencies is supported by the fact that in Spanish, the generalization of thematic distinctness seems to be extendable to intransitive ("dative") encodings of two-place verbs. Let us have a look at the following examples:[1]

(1) *María le gusta a Juan.* [María]$_{\text{THEMATIC}}$ [a Juan]$_{\text{SENTIENCE}}$

PN CL.3SG.DAT like.PRS.3SG IO PN

'Juan likes Maria.'

(2) *A Juan le pertenece una casa.* [a Juan]$_{\text{THEMATIC}}$ [una casa]$_{\text{THEMATIC}}$

IO PN CL.3SG.DAT own.PRS.3SG a house

'Juan owns a house.'

(3) *A Juan le falta dinero.* [a Juan]$_{\text{THEMATIC}}$ [dinero]$_{\text{THEMATIC}}$

IO PN CL.3SG.DAT lack.PRS.3SG money

'Juan is short of money.'

All three examples show a similar role-semantic configuration. In (2) and (3), taking Blume's (2000) terms, both co-arguments bear the role *thematic*.[2] In (1), the subject,

[1] Note that in all examples, the indirect object clitic *le* occurs obligatorily. For this double indirect object marking in Spanish cf. e.g. Montrul (1996) and Dufter/Stark (2008).

[2] Typically the indirect object participant is interpreted as *sentient* relative to the event (cf. section 3.3, fn. 26). Note furthermore that, interestingly, the examples with *faltar* 'to lack' and *pertenecer* 'to own' could be paraphrased with the lexical doublets *necesitar* 'to need' and *tener* 'to have' where the former indirect object would occupy the subject position with the former subject being semantically incorporated:

(2) *A Juan le pertenece una casa.* = *Juan tiene casa.*
'Juan owns a house.'

(3) *A Juan le falta dinero.* = *Juan necesita dinero.*
'Juan needs money.'

too, exhibits the role *thematic* while the indirect object is entailed *sentience*. Thus, in (2) and (3), subject and indirect object are equally agentive, whereas in (1) the indirect object is even more agentive than the subject. In all three cases we are hence dealing with a sharp deviation from the canonical agent-patient asymmetry of a transitive sentence. With respect to morphosyntactic encoding, all three verbs select for an indirect object, that is they exhibit *non-differential a*-marking. This suggests that García García's (2007: 71, 2014: 145) generalization of thematic distinctness might be suitable to predict not only *a*-marking of inanimate direct objects but also of indirect objects. If this proves true, his generalization might be extended in the following way:

(4) *a*-marked objects in Spanish (to be proved):

> If a (direct or indirect) object bears an equal number as or a higher number of proto-agent properties than the subject it is *a*-marked.

In comparing Blume's (2000) and García García's (2014) account it is also important to note that within the former model, the Spanish reversible predicates being crucial for a role-semantic approach on DOM (e.g. *preceder* 'to precede' or *seguir* 'to follow') would not meet the licensing condition for marked valencies due to their *minimal* semantic transitivity. Instead, their cross-linguistic occurrence with marked case frames is accounted for as an overgeneralization of the licensing condition: the choice of a marked case frame is claimed to be due to their *fully symmetric* argument relation since the non-asymmetric argument relation is seen as the most striking characteristic of verbs of restricted semantic transitivity (cf. Blume 2000: 214).

There is another case of overgeneralization of the licensing condition that is especially interesting for this study: In languages that show a preference for marked valencies such as Czech or Finnish, the fact that both arguments bear proto-agent properties in a subevent can also lead to a marked case frame even if proto-patient properties are also assigned (Blume 2000: 217). Dealing with DOM in Spanish, this case is crucial since in the most prototypical cases, animate objects in a typical transitive scenario and not inanimate ones of a minimal transitive predicate are *a*-marked. So, for instance, verbs with a clear-cut agent-patient asymmetry such as *matar* 'to kill' and *herir* 'to wound' show a DOM preference in the diachrony of Spanish (von Heusinger 2008). These verbs have the particularity to select for an animate object. Following Primus' (2012a: 73) argumentation, we can account for the occurrence with DOM in these cases by the permanent "role overlap" between proto-patient and potential proto-agent properties for the object. Hence DOM "occurs whenever a noun phrase subcategorized for patient properties acquires potential agent properties due to its intrinsic meaning" (Primus 2012a: 73). So, we must distinguish between an *actual* lack of thematic distinctness and one that is only *potential*. In the first case, the generalization of thematic distinctness applies in a strict sense, while in the

The existence of lexical doublets of possession verbs is predicted by Blume (2000: 212) and explained by the fact that these verbs typically do not allow for passivization. The doublets thus guarantee that both arguments can be topicalized.

second case, we only deal with an overgeneralization where thematic distinctness is blurred because the object bears a role overlap between *implied* proto-patient and *potential* proto-agent properties.

As we shall further see, DOM with animate objects in Spanish not only can be accounted for by reference to (pragmatically conveyed) *potential* agentivity. Rather, proto-agent properties can also be anchored in the verbal semantics and thus be *entailed* for the direct object. In this case, we can find an *actual* overlap of entailed proto-agent and proto-agent properties, being attested for example for the objects of OE-psych verbs such as *asustar* 'to frighten', *enojar* 'to annoy' or *ofender* 'to offend' in their causative reading. Bear in mind that Primus' extended dative default (2012a: 81), presented here again for convenience, exactly captures these two different sources of proto-agentivity:

(5) The dative is used for an argument with a low number of proto-agent properties that are i) subcategorized by the verb or ii) assigned according to the intrinsic meaning of the direct object referent.

With this default, irrespective of its applicability as a linking principle, Primus shows that indirect and direct *a*-marked objects in Spanish share a common role-semantic property, namely the agentivity of the argument in question. While in her account subcategorized proto-agent properties are seemingly associated with indirect objects of two-place predicates and ditransitive verbs, intrinsic agentive properties are given for animate *a*-marked direct objects. Note, however, that the above-mentioned OE-psych verbs implying *sentience* for their objects can be likewise subsumed under i). I will present a deeper analysis of verbs and verb classes that entail agentivity for their direct object with examples from different Romance languages in section 4.3.1.

In addition, there is another lexically anchored agentivity encoding in a predicate, namely *presupposed* proto-agent properties. As has been shown in section 3.2.2, these properties are of relevance for the selection of marked case frames in the case of interaction verbs like 'help' or 'obey', which presuppose an initial subevent in which the object participant acts as a proto-agent. Here, I will only delineate why the consideration of presupposed proto-agent properties serves to account for instances of DOM with inanimate objects in Spanish, too. As shown by García García (2014: 178ff.), verbs of competition including verbs like *vencer* 'to defeat', *atacar* 'to attack' or *derrotar* 'to win' systematically, though only optionally, also select DOM with inanimate objects. Two relevant examples for *vencer* and *derrotar* are given in (6):

(6) a. *porque ya sabe vencer **al** destino*
because already know.PRS.3SG defeat.INF DOM.the destiny
'Because he already knows how to defeat destiny.'

b. *El viejo ha derrotado **a** la soledad [...]*
the old man have.PRS.3SG win.PTCP DOM the loneliness

'The old man has defeated loneliness […].' (García García 2014: 186, 183)

In both examples, we find DOM with an abstract entity as direct object, *el destino* 'destiny' and *la soledad* 'loneliness', while the subject is human in both cases. Checking the examples with regard to the generalization of thematic distinctness in a strict sense, the *a*-marker would not appear here: Since the verbs *vencer* and *derrotar* – also in the given readings – imply a potent proto-agent property (*autonomous activity*) for the subject and the proto-patient property *affected* for the object, both arguments are role-semantically distinct. Consequently, the generalization of thematic distinctness would predict the object marker to be absent in both (6a) and (6b). There is yet another role-semantic particularity with respect to the direct objects of verbs of competition: they function as *potential agents* (García García 2014: 178) given that an event of defeating or winning against a competitor requires a previous competition or a fighting event in which both participants are active. As will be substantiated by empirical diagnostics in section 4.2.1, in these cases the agentivity of the object argument is *presupposed* by the lexical meaning of the predicate. The role-semantic constellation for verbs of competition can thus be summed up as follows: Concerning the proto-properties entailed by the predicate, subject and direct object show a clear thematic asymmetry. Yet, given the implied proto-patient properties for the object overlap with presupposed proto-agent properties in a previous subevent, thematic distinctness may still be blurred to a certain degree.[3]

In summary, *a*-marked (inanimate) direct and indirect objects in Spanish, although concerning different dimensions of linking, can be linked to a similar role-semantic configuration: the non-asymmetric distribution of proto-roles between co-arguments. Based on Blume (2000) it was proposed that besides the two scenarios where (i) the object outranks the subject in terms of agentivity and (ii) subject and object are equally agentive, also (iii) the simple presence of proto-agent properties for both arguments – despite a potential overlap with proto-patient properties for the object – may lead to morphological marking. It was further stated that the agentivity of the object can show up on different levels of meaning being either *entailed* or *presupposed* by the predicate or only *potential*, which means rather pragmatically conveyed. What is needed is clear diagnostics for the different levels of agentivity that in a next step allow for a differentiation of testable verb classes.

4.2 Types of communicated agentivity for the object

As indicated in the previous section, there are different ways of communicated agentivity for the object which, at least in Spanish, seem to matter for DOM. Role-semantic configurations with an agentive object are still understudied and require further specification in a first step, independently of the phenomenon of DOM. For this purpose, I will present empirical diagnostics to isolate different proto-agent properties

[3] Other verbs that also show DOM with inanimate objects in Spanish and that presuppose a similar previous activity of the object participant are *ayudar* 'to help', *temer* 'to fear' and *obedecer* 'to obey'. Sticking to the account of Blume (1998, 2000), unlike verbs of competition, these verbs do not entail proto-patient properties to their objects.

for the object and to identify different types of encoding in the verbal semantics (4.2.1). I will both review several diagnostics presented in the literature, where these were mainly applied to subjects, and introduce new ones. For the sake of simplicity, agentivity diagnostics will be mainly illustrated by English examples. The tests support a differentiation of four agentivity levels (*entailed agentivity*, *presupposed agentivity*, *potential agentivity*, and if none of the tests applies, *unspecified for agentivity*). I suggest that these levels of communicated agentivity for the object can be arranged on a four-point agentivity scale (4.2.2). Note that, although the diagnostics are mostly restricted to the agentivity of objects, at the end of this section a connection to the agentivity of subjects is also made, accounting for the relevance of the relation between both co-arguments of a transitive scenario elaborated previously.

4.2.1 Empirical diagnostics

In the following, empirical diagnostics for identifying the object's agentivity are presented. The tests in (a) and (b) consider the verbal entailments *volition* and *sentience*, respectively. Though it does not have the status of an involvement property, I argue that the criterion of *no co-argument dependency* in (c) can be also subsumed under *entailed agentivity*. In (d) and (e) diagnostics to identify two different types of *presupposed agentivity* for the object are presented. The diagnostics in (f) and (g) aim at isolating pragmatically conveyed proto-agent properties for the object, that is a *conversational implicature of sentience* on the one hand and the *subject potential* of the object argument on the other hand. In (h) a summary and an overview of the diagnostics is given.

a. Volition entailed for O

Working within the proto-role framework, I will stick to Dowty's (1991) notion of *volition*, defined as "volitional involvement in an event or state" (1991: 572), and Primus' (1999a) and Blume's (2000) similar though more general notion of *control*. This latter notion also captures the ability of an agent "to start and stop the event at will" (Primus 1999a: 37) as well as the agent's responsibility for it.

To begin with, *volition* is only rarely implied for an object argument. In accusative languages such as English, a volition entailment is never *only* attested for the object but not for the subject argument (Dowty 1991: 585). This corresponds to Blume's (2000: 143) linking principle that predicts potent proto-agent properties involving *control* to be linked per default to nominative case. The rare cases in which *volition* is entailed for both subject and object are symmetric predicates such as *to marry*, *to divorce* or *to date*. For these predicates, the *understood reciprocal object alternation* (Levin 1993: 36f.) applies with no change in truth-conditions between the transitive (7a) and the intransitive variant (7b):

(7) a. Kim married Sandy.

b. Kim and Sandy married. (Dowty 1991: 584)

It is shown further below how the volitional involvement of *Sandy* can be tested in both variants. Before, I will take a look at the Spanish translation of (7): As we can see, in contrast to English, *casarse* 'to marry' does not get a transitive encoding but is expressed by a reflexive and the preposition *con* 'with' (8a). This is characteristic for Spanish where reciprocity is reflected by the use of the reflexive pronoun *se*.

(8) a. *Kim se casó con Sandy.*
PN REFL.3 marry.PST.3SG with PN
'Kim married Sandy.'

b. *Kim y Sandy se casaron.*
PN and PN REFL.3 marry.PST.3PL
'Kim and Sandy married.'

However, not all verbs allowing the *understood reciprocal object alternation* do in fact express reciprocity in both variants. Rather, for verbs like *to kiss*, *to hug* or *to embrace*, a truth-conditional asymmetry between the transitive and the intransitive sentence can be ascertained (Dowty 1991: 583f.). In English, the transitive verb *to kiss* as in (9a) entails the proto-agent properties *volition*, *sentience* and *movement* for the subject while the direct object is assigned the proto-patient property *causally affected*. The intransitive variant (9b), in contrast, is only felicitous if both participants are volitionally involved at the same time.[4]

(9) a. Kim kissed Sandy.

b. Kim and Sandy kissed. (Dowty 1991: 584)

With regard to Spanish, we can even identify a different encoding between reciprocal *casarse* 'to marry' and non-reciprocal *besar* 'to kiss': while the former only shows up as a reflexive verb, the latter is transitive in case of a role-semantic asymmetry between subject and object (10a) but requires the use of the reflexive *se* when it has a reciprocal reading such as in (10b):

(10) a. *Kim besó a Sandy.*
PN kiss.PST.3SG DOM PN
'Kim kissed Sandy.'

b. *Kim y Sandy se besaron.*
PN and PN REFL.3 kiss.PST.3PL

[4] Blume (2000: 215) even distinguishes three different readings of 'kiss' each of which is associated with a different argument structure: *'kiss'1* as 'physical contact' (as in (9a)), *'kiss'2* as 'social gesture', with an object participant that is sentient via conversational implicature, and *'kiss'3* as 'exchange of affection' (as in (9b)).

'Kim and Sandy kissed.'

Since *volition* typically occurs as an entailment for the subject, empirical diagnostics to extract this property are commonly formulated only for the subject but not for the object argument. For German, Primus (2012b: 18) uses the pseudo-cleft sentence *Was x unterließ, war…* in order to distinguish action verbs with a (potentially) controlling subject participant from stative predicates with a non-controlling subject. This test should be more restrictive than the test *what x did was…* mentioned in Cruse (1973: 18f.) for extracting the agent-feature *volitive*. I translated Primus' test into English (*what x refrained from was…*) and applied it to three predicates:

(11) *what x refrained from was…*

a. What Peter refrained from was looking at the actor.

b. What Peter refrained from was opening the window.

c. #What Peter refrained from was knowing Mary.

Yet, what this test singles out is controllability rather than an actual *volition* or *control* entailment: This is why it applies both to *to look* (11a) that implies *volition* for its subject and to *to open* (11b) where the volitional involvement of the subject follows only by conversational implicature (this can be proved by its cancellability: *Peter accidentally opened the window with his shoulder*). The test does not work for cognitive states such as *to know* since these are not controllable by the subject participant. For my purposes, I adapt the test as follows: if a verb entails *volition* for the object, the situation denoted by the predicate must be controllable for the object participant. Hence the above-mentioned test should apply for the object as well:

(12) *what y refrained from was…*

a. (*Kim kissed Sandy*) #What Sandy refrained from was kissing Kim.

b. (*Kim and Sandy kissed*) What Sandy refrained from was kissing Kim.

c. (*Kim married Sandy*) What Sandy refrained from was marrying Kim.

d. (*Kim and Sandy married*) What Sandy refrained from was marrying Kim.

While for *to marry*, the test works as well for the transitive (12c) as for the intransitive (12d) variant, for *to kiss* it is acceptable for the intransitive (12b) but less good for the transitive variant (12a). This corresponds to Dowty's (1991: 583f.) analysis stating that in contrast to *to marry*, *to kiss* shows a truth-conditional asymmetry between the two variants: while the former is reciprocal in both variants, the latter shows reciprocity only in the intransitive reading but displays a role-semantic asymmetry in the transitive use (12a).

Another test used by Primus (2012b: 18f.) is the insertion of (non-)volitional adverbials such as *voluntarily*, *involuntarily, accidentally*, *intentionally*, *unintentionally*. This test applies to controllable agents, that is to verbs where *volition* or *control* is

conveyed for the subject participant by conversational implicature and hence can be stated explicitly (13a) or cancelled in context (13b):

(13) *(non-)volitional adverbials*

a. Mary intentionally broke the vase.

b. Mary accidentally broke the vase.

c. #Mary intentionally knew the answer.

d. #Mary accidentally knew the answer.

(14) *(non-)volitional adverbials*

a. ??Mary intentionally kissed Peter.

b. ??Mary accidentally kissed Peter.

c. #Mary intentionally married Peter.

d. #Mary accidentally married Peter.

However, besides for states (13c-d), the test broadly fails to apply for verbs that entail *volition* (14a-d) since in the latter case volitional adverbials are redundant and non-volitional adverbials that "cancel" the entailment seem odd. Thus, this test, apart from the question of its extendibility to objects, is not selective enough for the purposes of this study.

Instead, I propose the simple test *although y did not want/intend to…*, exemplified in (15), which cancels a volitional involvement of the object. The cancellation should be impossible if *volition* is implied.

(15) *although y did not want/intend to…*

a. Mary kissed Peter although *he* did not want/intend to kiss *her*.

b. Mary embraced Peter although *he* did not want/intend to embrace *her*.

c. Mary married Peter #although *he* did not want/intend to marry *her*.

d. Mary dated Peter #although *he* did not want/intend to date *her*.

As expected, the concessive clause is acceptable for the predicates *to kiss* (15a) and *to embrace* (15b) but (without additional contextual information) unacceptable for *to marry* (15c) and *to date* (15d). This suggests that the latter but not the former encode a *volition* entailment for the object in their verbal semantics.

b. Sentience entailed for O

The proto-agent property *sentience*, unlike *volition*, occurs frequently in object position. It is found in particular with OE-psych verbs such as English *to disturb* or *to annoy* as well as with dative experiencer verbs like German *gefallen* 'to like'. I repeat

here Dowty's definition of *sentience* for convenience: "Sentience means more than a presupposition that an argument is a sentient being; it is rather sentience with respect to the event or state denoted by the verb: the objects of verbs like *elect, appoint, nominate* and *idolize*, *venerate* and *convict, acquit, exculpate* are necessarily human but are not entailed to know or perceive the relevant event." (1991: 573). The presence of a *sentience* entailment is illustrated for classic propositional attitude verbs (*to know*, *to believe*), stative perception verbs (*to see*) and stative psych predicates (*to fear*, *to be disappointed at*) (Dowty 1991: 573). Although Dowty subsumes *sentience* and *perception* under one single proto-agent property, he notes that he was not sure if one should draw a distinction between both properties.

Engelberg (2000a: 194-198), who discusses Dowty's proposal, argues that the proclamation of a proto-agent property *sentience* as a single lexical entailment is problematic. Instead, *sentience* should be analyzed as a cover term of two disjunct semantic relations, namely *perception* and *knowledge*. He points out the following problems of a notion of *sentience* as defined by Dowty (1991): Firstly, Engelberg (2000a: 195f.) demonstrates that *perception* and *knowledge* represent two independent semantic relations. Whereas a *knowledge* relation can relative straightforwardly be argued to hold between an individual and a proposition, for a *perception* relation there are two options: one could either 'perceive something' or 'perceive something as something'. In the former case, the object is either an individual (16a, b) or an event (16c), while in the latter case, it is a proposition about an individual (17a, b) or about an event (17c):

(16) a. I see an elm.

b. I listen to an operetta.

c. I hear her singing an aria.

(17) a. I see that it is an elm.

b. I hear that it is an operetta.

c. I hear that she is singing an aria. (Engelberg 2000a: 195; translation: S.M.)

Only the examples in (17) but not the ones in (16) imply *knowledge* for their subject argument. This can be proved by a negation test that explicitly excludes *knowledge* for the subject participant in (16) (cf. Engelberg 2000a: 195, fn. 68). While the test applies to 'perceive something' (18), it is odd for 'perceive something as something' (19):

(18) a. I saw an elm but at that time, I did not know that it was an elm.

b. I listened to an operetta but at that time, I did not know that it was an operetta.

c. I heard her singing an aria but at that time, I did not know that it was an aria.

(19) a. I saw it was an elm #but at that time, I did not know that it was an elm.

b. I heard it was an operetta #but at that time, I did not know that it was an operetta.

c. I heard she was singing an aria #but at that time, I did not know that it was an aria.

Hence, following Engelberg (2000a: 195), we can subsume three different semantic relations under Dowty's *sentience* notion, which can be formalized as follows (translation: S.M.):

(20) a. KNOW(x,p,e)

b. $PERCEPT_1$(x,e',e) or $PERCEPT_1$(x,y,e)

c. $PERCEPT_2$(x,P(e'),e) or $PERCEPT_2$(x,P(y),e)[5]

The relation in (20b), which is independent of (20a), is taken to be the core perception relation, while the relation in (20c) corresponds to the examples in (17) above where *knowledge* is also implied for the subject. Further support for the independence of (20a) and (20b) is given by Engelberg's (2000a: 196) examples of verbs that can either imply only *perception*, but not necessarily *knowledge* such as *beleidigen* 'to insult', only *knowledge* like *verzichten* 'to forgo' or both *perception* and *knowledge* such as *schreiben* 'to write' for their subject participant. Secondly, Engelberg poses the question of why both properties have been subsumed together under the disjunct notion of *sentience*. From a lexical semantic perspective, this grouping must be regarded as arbitrary. Consequently, if the counting of *knowledge* and *perception* under the label of *sentience* as "1" cannot be given a semantic explanation, then Dowty's argument selection principle is not fully due to lexical semantic properties but also dependent on a seemingly arbitrary grouping of properties (Engelberg 2000a: 196). Thirdly, he identifies a contradiction between Dowty's (1991: 573) criterion for a sentient participant "to know or perceive the relevant event" and the application of this criterion to the example verbs such as in *John knows/believes/is disappointed at the statement*. From this perspective, *sentience* would only be present if the participant, here *John*, would indeed *know* that he knows/believes/is disappointed at the statement. Leaving this complex epistemological question aside, another remark by Engelberg (2000a: 197) with regard to psych verbs should not remain unmentioned here. He subsumes these verbs under *introspective perception* which is argued to stand in complementary distribution to *perceptive perception* verbs. Thus, in these cases Dowty's counting of *sentience* as "1" for the subject of verbs like *to fear* or *to see* would be indeed justified. Yet if, besides simple (*introspective*) *perception*, *knowledge* is also present (as formalized in (20c) above), the two properties should count as "2", which would not be consistent with the attested converse linking patterns of non-causative psych verbs. Therefore, as already discussed in section 3.2.2,

[5] Cf. Engelberg (2000a: 195):
[(59)] a. WISS(x,p,e)
b. $WAHRNEHM_1$(x,e',e) bzw. $WAHRNEHM_1$(x,y,e)
c. $WAHRNEHM_2$(x,P(e'),e) bzw. $WAHRNEHM_2$(x,P(y),e)

instead of a counting algorithm a different weighting of proto-agent properties should be taken into account with *sentience* being of less weight for linking than other proto-agent properties.

In this study, I will continue employing the cover term *sentience*. In the following diagnostics, I will, however, keep apart the different sub-properties. Concerning the terminology of the sub-properties, I follow a recent proposal of Kretzschmar et al. (2018a) who argue in favour of the differentiation of (1) *cognitive state*, (2) *perception* and (3) *emotion* as independent sentience features. That is to say, the presence of one of these properties suffices to ascribe the proto-agent property *sentience* to a given verb. Though they use a different terminology, their proposal is broadly in line with the different semantic relations Engelberg (2000a) argues to be subsumed under the notion of *sentience*. Table 1 compares the two accounts: the main difference is that Engelberg subsumes (2) *perceptive* and (3) *introspective perception* under one category (PERCEPTION$_1$), but declares them to be mutually exclusive, whereas Kretzschmar et al. (2018a) directly differentiate between three independent sub-features. Similarly, possible conjunctions of sub-properties fall under the relation PERCEPTION$_2$ in Engelberg's account, while there is no such cover relation for accumulations of sub-features in Kretzschmar et al (2018a).

(Independent) sub-properties of *sentience*	
Engelberg (2000a: 195)	**Kretzschmar et al. (2018a)**
(1) knowledge	(1) cognitive state
PERCEPTION$_1$: (2) perceptive ∨ (3) introspective	(2) perception
	(3) emotion
PERCEPTION$_2$: ((2) perceptive perception ∧ (1) knowledge)	(1) cognitive state ∧ (2) perception
∨ ((3) introspective perception ∧ knowledge)	(1) cognitive state ∧ (3) emotion

Table 1. Sub-properties of *sentience*

In a manual on sentience verbs in German, Kretzschmar et al. (2018b) propose semantic tests that allow us to isolate the three sub-properties. These tests work with a negation of the entailment in question and should thus be non-applicable in case the sub-property is lexically encoded, as displayed in (21):

(21) a. Cognitive state: *x knows something, #but has no idea of it*

(e.g. John knows the answer #but has no idea of it)

b. Perception: *x perceives something, #but has no idea of it*

(e.g. John sees the book #but has no perception of it)

c. Emotion: *x feels something, #but has no feeling of it*

(e.g. John loves Mary #but has no feelings for her/no feeling of it)

Since these tests are constructed for extracting entailments for the subject argument, I slightly modified them in order to be applied to objects, too. They are presented in

(22)-(24) below for the verbs *to fascinate*, *to frighten*, *to resemble*, *to greet*, *to call*, *to help, to defeat*, *to cure*, *to injure*, *to insult*, *to see*, *to admire*, *to know*.

(22) **Cognitive state:** *x did something, #but y had no idea of it*

a. The woman fascinated/frightened Peter ?but he had no idea of it.

b. The woman resembled Peter but he had no idea of it.

c. The woman called/greeted/helped Peter but he had no idea of it.

d. The woman defeated Peter but he had no idea of it.

e. The woman cured/injured Peter but he had no idea of it.

f. The woman insulted Peter but he had no idea of it.

g. The woman saw Peter but he had no idea of it.

h. The woman admired/knew Peter but he had no idea of it.

(23) **Perception:** *x did something, #but y had no perception of it*

a. The woman fascinated/frightened Peter ?but he had no perception of it.

b. The woman resembled Peter but he had no perception of it.

c. The woman called/greeted/helped Peter but he had no perception of it.

d. The woman defeated Peter but he had no perception of it.

e. The woman cured/injured Peter but he had no perception of it.

f. The woman insulted Peter but he had no perception of it.

g. The woman saw Peter but he had no perception of it.

h. The woman admired/knew Peter but he had no perception of it.

(24) **Emotion:** *x did something, #but y has no feeling of it*

a. The woman fascinated/frightened Peter #but he had no feeling of it.

b. The woman resembled Peter but he had no feeling of it.

c. The woman called/greeted/helped Peter but he had no feeling of it.

d. The woman defeated Peter but he had no feeling of it.

e. The woman cured/injured Peter but he had no feeling of it.

f. The woman insulted Peter but he had no feeling of it.

g. The woman saw Peter but he had no feeling of it.

h. The woman admired/knew Peter but he had no feeling of it.

As can be seen, the tests turn out to be negative for nearly all verbs with an exception of the OE-psych verbs *to fascinate* and *to frighten*. For these two verbs, an entailment of *cognitive state* (22a) and *perception* (23a) is hard to determine but an entailment of *emotion* (24a) can be clearly demonstrated by the negation test.

c. No co-argument dependency

Unlike *volition* and *sentience*, the criterion *no co-argument dependency* can only be identified when considering the relation between subject and object. It goes back to Primus' (1999a, b, 2006) co-argument dependency model, which has been presented in section 3.2.1. As a short repetition, according to the model not only involvement properties of the participants (e.g. the proto-agent properties *sentience* and *volition*) can qualify for being a proto-agent in an event but also the lack of the (structural) co-argument dependency relation between subject and object. Thus, if both have argument status but are not assigned any involvement properties, they are both conceived as (weak) proto-agents.[6] The fulfilment of this criterion can thus only be assessed ex negativo, that is by excluding the presence of any proto-properties for both arguments.

Heuristics that may serve to identify verbs (or readings of verbs) to which the criterion *no co-argument dependency* applies are (i) the change of subject and object without change of truth-conditions or the existence of (ii) lexical diatheses or (iii) lexical converses. Since for a majority of the predicates in question a passive option is rather disfavoured or not available because they are syntactically intransitive, these can all be regarded alternative strategies to topicalize the object. Note that, although these predicates do not properly *entail* the object's agentivity, depending on the verb, new roles or predicates can be defined that further specify the relation between the two co-arguments. This is shown by García García (2014: 147-170) who introduces i.a. the role-semantic representations BE-LIKE(x, y) → BE-LIKE(y, x) for reversible-symmetrical predicates and AFTER(EXIST(x), EXIST(y)) → BEFORE(EXIST(y), EXIST(x)) for reversible-converse predicates. In (25)-(27), the three heuristics mentioned are illustrated by examples:

(25) **Change of subject and object without change in truth-conditions (reversible-symmetrical predicates, cf. García García 2014: 128, 159):**

a. German *ähneln* 'to resemble' (intrans.):

Der Apfel ähnelt der Birne. → *Die Birne ähnelt dem Apfel.*

'The apple resembles the pear.' → 'The pear resembles the apple.'

[6] In Blume's (2000: 212) version of the proto-role model, which does not make use of a separate notion of causal dependency, this configuration is characterized only with reference to verb lexeme-specific criteria: besides the lack of proto-properties assigned by the verb to arg1 or arg2, the verb does not stand in the valency relation of involvement to arg1 or arg2. As a consequence, these predicates are of *minimal semantic transitivity*. Examples are the German verbs *ähneln* 'to resemble' and *entsprechen* 'to correspond to' (Blume 2000: 214) where both subject and object bear the role *thematic*.

BE-LIKE(x, y) → BE-LIKE(y, x)

b. Spanish *igualar* ‘to equal’ (trans.):

Los productos chinos igualan a los productos europeos.

→ *Los productos europeos igualan a los productos chinos.*

‘Chinese products are equal to European products.’

→ ‘European products are equal to Chinese products.’

BE-LIKE(x, y) → BE-LIKE(y, x)

(26) **Lexical diatheses (cf. García García 2010: 151f.):**

a. Spanish *caracterizar* ‘to characterize’ (trans.)/*ser* ‘to be’ (copula):

(1) *La euforia caracteriza a la situación centroamericana.*

‘Euphoria characterizes the Central American situation.’

(2) *La situación centroamericana es eufórica.*

‘The Central American situation is euphoric.’

b. Spanish *pertenecer* ‘to belong to’ (intrans.)/*tener* ‘to have’ (trans.):

(1) *La casa pertenece a Pedro.*

‘The house belongs to Pedro.’

(2) *Pedro tiene una casa.*

‘Pedro has a house.’

(27) **Lexical converses (reversible-converse predicates, cf. García García 2014: 163):**

Spanish *preceder* ‘to precede’ (trans.)/*seguir* ‘to follow’ (trans.):

(1) *El verano precede al otoño.*

‘Summer precedes autumn.’

BEFORE(EXIST(x), EXIST(y)) → AFTER(EXIST(y), EXIST(x)

(2) *El otoño sigue al verano.*

‘Autumn follows summer.’

AFTER(EXIST(x), EXIST(y)) → BEFORE(EXIST(y), EXIST(x))

As exemplified by German intransitive *ähneln* ‘to resemble’ in (25a) and Spanish transitive *igualar* ‘to equal’ in (25b), subject and object can change places without change in truth-conditions. As for lexical diatheses, two examples from Spanish are given: these consider transitive *caracterizar* ‘to characterize’ and the copula *ser* ‘to be’ on the one hand (26a) and intransitive *pertenecer* ‘to belong to’ and transitive

tener ‘to have’ on the other hand (26b). The heuristic of lexical converses is illustrated by Spanish transitive *preceder* ‘to precede’ (27a) and *seguir* ‘to follow’ (27b).

d. Autonomous activity presupposed for O

Dowty (1991: 552), before introducing his set of proto-agent and proto-patient properties, already remarks that some of his *lexical entailments* might rather be treated as *presuppositions*. Blume’s (1998) modification of his model then states a differentiation between entailed and presupposed proto-agent properties: just as a verbal entailment, she analyzes presupposed agentivity of an argument as being encoded in the lexical semantics of a predicate. For instance, the interaction verb *to contradict implies* that the subject participant says something in a subevent, while it *presupposes* at the same time that the object participant has claimed the opposite in a previous subevent.[7]

Presuppositions are here understood in their semantic sense, namely as implicit but necessary requirements for an expression to have a meaning (cf. Strawson 1950).[8] This can be logically reformulated in the following way (cf. Levinson 1994: 177):

(28) One sentence A semantically presupposes another sentence B, iff:

a. in all situations, where A is true, B is true

b. in all situations, where A is false, B is true

A characteristic of presuppositions, already observed by Frege (1892), is that they survive the negation of a statement. This is what sets them apart from entailments and serves as a common diagnostic used in linguistics to identify presuppositions. Beyond this, a presupposition is constant under a change of mood as for example in questions (cf. Engelberg 2000a: 34). Both tests are illustrated in (29): As can be seen for the predicate *to defeat*, regardless of whether we deal with an affirmed (29a) or negated (29b) sentence or with a question (29c), the implicit sentence *Mary and Peter were active in a tennis match* referring to a previous subevent in which both participants were active is still true.

(29) a. *Mary defeated Peter in the tennis match.*

[7] Engelberg (2000a: 332f.) points out that the (consecutive or concessive) relations that apply between the two subevents of Blume’s (1998) interaction verbs might rather be treated as relations between *propositions about (sub)events* than between proper *(sub)events*.

[8] Using the example *The king of France is wise* Strawson (1950: 330) develops the argument that what are now called *presupposition failures* lead to problems in assigning truth values to a statement:
“To say, ‘The king of France is wise’ is, in some sense of ‘imply’, to *imply* that there is a king of France. But this is a very special and odd sense of ‘imply’. ‘Implies’ in this sense is certainly not equivalent to ‘entails’ (or ‘logically implies’). And this comes out from the fact that when, in response to his statement, we say (as we should) ‘There is no king of France’, we should certainly *not* say we were *contradicting* the statement that the king of France is wise. We are certainly not saying that it’s false. We are, rather, giving a reason for saying that the question of whether it’s true or false simply doesn’t arise.”

→P Mary and Peter were active in a tennis match.

b. *Mary did not defeat Peter in the tennis match.*

→P Mary and Peter were active in a tennis match.

c. *Did Mary defeat Peter in the tennis match?*

→P Mary and Peter were active in a tennis match.

It is not straightforward to determine the presupposed proto-agent properties of the object participant. There is certainly some kind of autonomous movement of *Peter* but he should also be volitionally involved and thus sentient relative to the presupposed subevent. For the time being, I therefore make use of Blume's (1998, 2000) vague (potent) proto-agent property *autonomous activity*. Since for my purposes, negation is sufficient as a criterion to differentiate presuppositions from verbal entailments, I will only use this one. It can be seen that the negation test *x did not* PREDICATE$_a$ *y* →$_P$ *y was active before…* works in order to extract presupposed autonomous activity for the verbs *to help* (30e), *to defeat* and *to stop* (30f). It furthermore applies to verbs of auditory perception such as *to listen to* and some communication and social gesture verbs like *to thank* (30d). In contrast, the test is not applicable to the verbs *to frighten* and *to fascinate* (30a), *to resemble* (30b), *to call* and *to greet* (30c), *to cure*, *to injure* and *to insult* (30g) as well as *to see* (30h), *to admire* and *to know* (30i).

(30) **Negation Test: Presupposed autonomous activity for O**

x did not PREDICATE$_a$ *y* →$_P$ *y was active before…*

a. *The woman did not fascinate/frighten Peter.*

→P #Peter was active before.

b. *The woman did not resemble Peter.*

→P #Peter was active before.

c. *The woman did not call/greet Peter.*

→P #Peter was active before.

d. *The woman did not listen to/thank Peter.*

→P Peter was active before (= produced some chain of sounds/ did something to the favour of the woman).

e. *The woman did not help Peter.*

→P Peter was active before (= striving for some aim).

f. *The woman did not defeat/stop Peter.*

→P Peter was active before (= in competition with the woman/ performing some activity).

g. *The woman did not cure/injure/insult Peter.*

→P #Peter was active before.

h. *The woman did not see Peter.*

→P #Peter was active before.

i. *The woman did not admire/know Peter.*

→P #Peter was active before.

e. Sentience presupposed for O

Besides *autonomous activity*, there are also interaction verbs that presuppose *sentience* for the object participant in a previous sub-"event". Making use of the subclassification of *sentience* introduced in b., I will present separate diagnostics for *cognitive state*, *perception* and *emotion*.

Let me start with the sub-property *cognitive state*. Such a property is typically presupposed for the object participant of "cheating verbs" like 'cheat on', 'deceive' and 'manipulate' as well as with "corruption verbs" such as 'bribe', 'buy off' or 'corrupt'. As illustrated by the verb *to manipulate*, a prior cognitive state of the object participant can not only be ascertained in an affirmed sentence (31a) but also under negation (31b) and in a question (31c):

(31) a. *Mary manipulated Peter with her behaviour.*

→P Peter had some cognitive state before (= which did not correspond to what Mary wanted from him).

b. *Mary did not manipulate Peter with her behaviour.*

→P Peter had some cognitive state before (= which did not correspond to what Mary wanted from him).

c. *Did Mary manipulate Peter with her behaviour?*

→P Peter had some cognitive state before (= which did not correspond to what Mary wanted from him).

I used the negation test *x did not* $PREDICATE_a$ *y* $\rightarrow_P$ *y had some cognitive state before…* to isolate a presupposed cognitive state for the object. The only verbs that pass the test are *to bribe* (32e), *to cheat on* and *to manipulate* (32f), and possibly also *to blackmail* and *to threaten* (32g). In contrast, it does not apply to the verbs in (32a-d) and (32h-k).

(32) **Negation Test: Presupposed cognitive state for O**

x did not $PREDICATE_a$ *y* $\rightarrow_P$ *y had some cognitive state before…*

a. *The woman did not fascinate/frighten Peter.*

→P #Peter had some cognitive state before.

b. *The woman did not resemble Peter.*

→P #Peter had some cognitive state before.

c. *The woman did not call/greet/help Peter.*

→P #Peter had some cognitive state before.

d. *The woman did not listen to Peter.*

→P #Peter had some cognitive state before.

e. *The woman did not bribe Peter.*

→P Peter had some cognitive state before (= he knew that he would profit from some state of affairs).

f. *The woman did not cheat on/manipulate Peter.*

→P Peter had some cognitive state before (= he believed in some state of affairs)

g. *The woman did not blackmail/threaten Peter.*

→P ?Peter had some cognitive state before (= he might have been aware of his unfavourable situation).

h. *The woman did not defeat Peter.*

→P #Peter had some cognitive state before.

i. *The woman did not cure/injure/insult Peter.*

→P #Peter had some cognitive state before.

j. *The woman did not see Peter.*

→P #Peter had some cognitive state before.

k. *The woman did not admire/know Peter.*

→P #Peter had some cognitive state before.

Presupposed perception for the object is mostly found with communication and social gesture verbs such as the ones mentioned by Blume (1998: 273): 'answer someone', 'write to someone', 'greet someone', 'call for someone', 'wave to someone', 'congratulate', 'thank', 'read to', 'threaten'. Taking the example of *to greet*, one can see that the possible perception of the object participant and his/her ability to react are constant in an affirmed (33a) and negated (33b) sentence as well as in a question (33c):

(33) a. *Mary greeted Peter in the street.*

→P Peter would have been able to perceive the action and to react.

b. *Mary did not greet Peter in street.*

→P Peter would have been able to perceive the action and to react.

c. *Did Mary greet Peter in the street?*

→P Peter would have been able to perceive the action and to react.

Accordingly, I employ the negation test *x did not* $PREDICATE_a$ *y* →$_P$ *y would have been able to perceive the action and to react…* to isolate the object's presupposed perception. While this diagnostic is clearly fulfilled by the verbs *to call* and *to greet* (34d), *to bribe*, *to cheat on* and *to manipulate* (34f) as well as by *to blackmail* and *to threaten* (34g), it is arguably also applicable to the OE-psych verb *to frighten* (34b). Moreover, it applies to verbs that also presuppose *autonomous activity* for the object like *to thank* (34e). The test does not apply to the verbs in (34a), (34c) and (34h-l).

(34) **Negation Test: Presupposed perception for O**

x did not $PREDICATE_a$ *y* →$_P$ *y would have been able to perceive the action and to react…*

a. *The woman did not fascinate Peter.*

→P #Peter would have been able to perceive the action and to react.

b. *The woman did not frighten Peter.*

→P ?Peter would have been able to perceive the action and to react.

c. *The woman did not resemble Peter.*

→P #Peter would have been able to perceive the action and to react.

d. *The woman did not call/greet Peter.*

→P Peter would have been able to perceive the action and to react
(= respond to the call/ greet back).

e. *The woman did not thank Peter.*

→P Peter would have been able to perceive the action and to react.

f. *The woman did not bribe/cheat on/manipulate Peter.*

→P Peter would have been able to perceive the action and to react
(= respond to the bribe/ uncover the fraud/ expose the manipulation).

g. *The woman did not blackmail/threaten Peter.*

→P Peter would have been able to perceive the action and to react
(= receive the blackmail/threat and respond to it).

h. *The woman did not help/listen to Peter.*

→P #Peter would have been able to perceive the action and to react.

i. *The woman did not defeat/stop Peter.*

→P #Peter would have been able to perceive the action and to react.

j. *The woman did not cure/injure/insult Peter.*

→P #Peter would have been able to perceive the action and to react.

k. *The woman did not see Peter.*

→P #Peter would have been able to perceive the action and to react.

l. *The woman did not admire/know Peter.*

→P #Peter would have been able to perceive the action and to react.

The sub-property *emotion* is presupposed for the object of verbs denoting negative communicative (and possibly also physical) acts such as 'blackmail', 'threaten' and 'force'. All these verbs presuppose that the object participant has some kind of fear *before* the act is uttered in order to be felicitous, as pointed out by Blume (1998: 279, fn. 17) by the example of 'threaten'. In (35), it is illustrated by the verb *to blackmail* that the presupposition of Peter fearing something before is not only given in an affirmative sentence (35a) but is constant under negation (35b) and in a question (35c):

(35) a. *Mary blackmailed Peter with the knowledge of his tax evasion.*

→P Peter had feared something before.

b. *Mary did not blackmail Peter with the knowledge of his tax evasion.*

→P Peter had feared something before.

c. *Did Mary blackmail Peter with the knowledge of his tax evasion?*

→P Peter had feared something before.

For isolating the presupposed sub-property *emotion*, the negation test has been adapted as follows: *x did not* PREDICATE$_a$ *y* →$_P$ *y had some emotional state…* As can be seen, it only applies to the above-mentioned verbs *to blackmail* and *to threaten* (36f). Thus, to the overwhelming majority of tested verbs (36a-e, g-j) it is not applicable.

(36) **Negation Test: Presupposed emotion for O**

x did not PREDICATE$_a$ *y* →$_P$ *y had some emotional state before…*

a. *The woman did not fascinate/frighten Peter.*

→P #Peter had some emotional state before.

b. *The woman did not resemble Peter.*

→P #Peter had some emotional state before.

c. *The woman did not call/greet/help Peter.*

→P #Peter had some emotional state before.

d. *The woman did not listen to Peter.*

→P #Peter had some emotional state before.

e. *The woman did not bribe/cheat on/manipulate Peter.*

→P #Peter had some emotional state before.

f. *The woman did not threaten/blackmail Peter.*

→P Peter had some emotional state before (= he feared something so that a threat/blackmail would have been possible in the first place).

g. *The woman did not defeat/stop Peter.*

→P #Peter had some emotional state before.

h. *The woman did not cure/injure/insult Peter.*

→P #Peter had some emotional state before.

i. *The woman did not see Peter.*

→P #Peter had some emotional state before.

j. *The woman did not admire/know Peter.*

→P #Peter had some emotional state before.

f. Implicature of sentience for O

Another type of inference repeatedly mentioned in the course of this study is the *conversational implicature of sentience* for the object. As already noted by Dowty (1991: 572, fn. 16) in a sentence like *Mary slapped John*, we typically interpret *John* to be sentient of the slapping although it is not encoded in the verbal semantics of the predicate *to slap* as a proto-agent entailment. This can be proved by the possible selection of inanimate objects that cannot be interpreted as relational potentially sentient like in *Mary slapped the table*. It is thus important "to distinguish entailments of the PREDICATE from what follows from any one sentence as a whole (e.g. entailments that may arise in part from NP meanings. etc.)" (Dowty 1991: 572, fn. 16). Primus (1999a: 58), besides verbal entailments, explicitly integrates conversational implicatures as another way proto-agent properties may be conveyed to an argument. Yet in her account, implicatures are restricted to (animate) subject participants, as for example in *Peter killed the bird* where *Peter* is interpreted to act volitionally though this volitional involvement is cancellable (*Peter unintentionally killed the bird*). Blume (2000: 139) follows Primus and systematically distinguishes between implied and implicated proto-agent properties. Sporadically, she indicates that implicated proto-agent properties may also occur for the object participant (as for the object of 'belong to' and 'kiss').

I have already proposed in section 3.2.1 that for verbs that select an affected object, a conversational implicature of the following type might be formulated: "You may interpret proto-patients that are human as experiencers (in absence of any information to the contrary)":

(37) AFFECTED(y) & HUMAN(y) +> SENTIENT(y)

I assume that the preconditions to derive this implicature are the following:

(1) For a subset of highly transitive predicates (e.g. causative OE-psych verbs like 'frighten' or 'disturb' and verbs like 'awake') the proto-patient property AFFECTED regularly co-occurs with a SENTIENCE entailment for the object.

(2) HUMANS are the core cases of sentient beings; all x ∈ HUMAN qualify as potentially sentient in a predicate SENTIENCE(x).

Thus, in case a HUMAN entity fills an argument slot AFFECTED(y) (e.g. for verbs like 'injure' or 'insult'), the default interpretation is that it is SENTIENT(y) relative to the event, too.

In pragmatic theory, such a default inference that "captures our intuitions about a preferred or normal interpretation" (Levinson 2000: 11) is called a *generalized conversational implicature* (GCI). This notion goes back to Grice (1975) and has been further refined by Levinson (1983, 2000). It can be defined as follows:

(38) An implicature i is *generalized* iff U implicates i *unless* there are unusual specific contextual assumptions that defeat it. (Levinson 2000: 16)

A typical example can be provided by the sentence *Some of the guests are already leaving* where, unless specific contextual cues are given, *some* is interpreted as 'not all' (+> 'Not all of the guests are already leaving'). Just like other types of implicatures, GCIs are based on *the cooperative principle* of human communication (39) and *the maxims of conversation* (40), which themselves are grounded within the rational nature of communication (cf. Grice 1989: 26f.):

(39) **The cooperative principle**

"Make your contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged."

(40) **The maxim of quantity**

1. Make your contribution as informative as is required (for the current purposes of the exchange).

2. Do not make your contribution more informative than is required.

The maxim of quality

Try to make your contribution one that is true.

1. Do not say what you believe is false.

2. Do not say that for which you lack adequate evidence.

The maxim of relation

Be relevant.

The maxim of manner

Be perspicuous.

1. Avoid obscurity of expression.
2. Avoid ambiguity.
3. Be brief (avoid unnecessary prolixity).
4. Be orderly.

Properties Grice (1975: 57f.) attributed to conversational implicatures are their *cancellability* (41a), *non-detachability* (41b), *calculability* (41c) and *non-conventionality* (41d). Further properties are their *reinforcability* (41e) and, less common, their assumed *universality* (41f):

(41) Properties of conversational implicatures (cf. Levinson 2000: 15):

a. **Cancellability (i.e., defeasibility)**: The property of being an inference defeatable by the addition of premises

b. **Non-detachability:** Any expression with the same coded content will tend to carry the same implicatures

c. **Calculability:** The more or less transparent derivation of the inference from premises that include the assumption of rational conversational activity

d. **Non-conventionality:** The noncoded nature of the inferences and their parasitic dependence on what is coded

e. **Reinforcability:** It is often possible to add explicitly what is anyway implicated with less sense of redundancy than would be the case if one repeated the coded content (Sadock 1978, Hom 1991b)

f. **Universality:** Because the inferences are derived ultimately from fundamental considerations of rationality, we expect a strong tendency to universality (unlike with coded meanings, of course); conversational implicatures are motivated, not arbitrary

I will illustrate the properties (41a-e) by applying them to the GCI proposed in (37) above. Let us take the example sentence in (42). As can be seen, it implicates all three sub-properties of *sentience*. The fact that, unlike for the entailment of *sentience*, where *perception* and *emotion* are typically complementary distributed, all three properties coincide may be due to the fact that we are dealing with a physical action. As such, it can both affect the senses (e.g. it can be seen or heard) and – mainly by its result – the psychophysiological sphere.

(42) a. *Mary injured the actor with a knife.*

b. +> the actor knew that he was injured. (+ cognitive state)

c. +> the actor perceived the moment when he was injured. (+ perception)

d. +> the actor felt pain. (+ emotion)

The cancellability of the object's sentience in the above sentence (42a) is illustrated by (43). It can be seen that each sub-property of *sentience* is easily defeasible by creating an appropriate context.

(43) **Cancellability (i.e., defeasibility):**

a. …but he had no idea of it since he was sedated.

b. …but he had no perception of it since he was sedated.

c. …but he had no feeling of it since he was sedated.

Note again that cancellability would not be possible if a verb *entails* sentience for its object, as has been shown by the diagnostics in b. above (e.g. *The woman fascinated/frightened Peter #but he had no feeling of it.*).

In (44), the *non-detachability* of the GCI is shown by rephrasing the sentence. All three alternative expressions to (42a) clearly carry the same implicatures:

(44) **Non-detachability:**

a. *Mary wounded the actor with a knife.*

b. *Mary rammed a knife into the actor's back.*

c. *Mary stabbed the actor with a knife.*

+> the actor knew that he was injured. (+ cognitive state)

+> the actor perceived the moment when he was injured. (+ perception)

+> the actor felt pain. (+ emotion)

The calculability of the implicature from the sentence meaning, the maxims of conversation and the cooperative principle is exemplified in (45) by the implicature of *emotion* (42d).

(45) **Calculability:**

S said that *Mary injured the actor with a knife*. There is no reason to believe that S does not respect the maxims or at least the cooperative principle. For S to be able to say that *Mary injured the actor with a knife*, and thus really respect the maxims or the cooperative principle, S must think that *the actor felt pain*. S believes (and he believes that I believe that he believes) that I can see that the assumption, that he thinks *the actor felt pain*, is necessary. S does nothing to prevent me, the addressee, from thinking that *the actor felt pain*. So S intends, or at least wants to allow me to think that *the actor felt pain*. So by saying *Mary injured the actor with a knife*, S has implicated *the actor felt pain*.

The non-conventionality, meaning that the assumed implicature is not lexically coded, is proved in (46).

(46) **Non-conventionality:**

As shown by the diagnostics in 4.2.1 b. and 4.2.1 e., *sentience* is not encoded as a verbal entailment or a presupposition in the meaning of *to injure*.

The sentences (47a-c) show that it is possible to explicitly add the implicated content without any redundancy. In case of a sentience entailment, by contrast, such reinforcement would rather seem redundant (e.g. *Mary offended the actor #?and he felt (very) hurt*.).

(47) **Reinforcability:**

a. *Mary injured the actor with a knife and he immediately knew what happened to him.*

b. *Mary injured the actor with a knife and he perceived this moment (intensely).*

c. *Mary injured the actor with a knife and he felt (horrible) pain.*

Since the criterion of cancellability is the one that allows the clearest dissociation of GCIs from verbal entailments, I made use of this one in the following diagnostics. It can be seen that, like *to injure*, the verbs *to cure* and *to insult* as used in (48a) and (48b) can be considered core examples of an implicature of sentience for the object. As expected, the test fails to apply to OE-psych verbs such as *to frighten* in (48c) since these verbs convey *sentience* via verbal entailment which cannot be canceled in context. Interestingly, it turns out that some verbs that presuppose the object's agentivity (cf. diagnostics d. and e.) additionally implicate *sentience*: This is the case for verbs like *to cheat on*, *to manipulate*, *to bribe*, *to blackmail*, *to threaten* and *to defeat* (48d-g). Since all these verbs select for an affected object participant, this further supports the initial assumptions about AFFECTED(y) & HUMAN(y) as preconditions for the GCI to be made. Accordingly, the test fails to apply to interaction verbs that do not assign proto-patient properties to the object such as *to call*, *to greet* and *to listen to* (48h-j). It is furthermore not applicable to verbs like *to see*, *to admire* and *to know* which likewise do not select for an affected object (48k-l).

(48) **Implicature of sentience**: *+> sentience & …but y has no idea/perception/feeling of it*

a. *The doctor cured Peter.* [AFFECTED(y) & HUMAN(y)]

+> Peter knows that he is cured.

…but Peter has no idea of it since he has not opened the doctor's letter yet.

b. *The professor insulted Peter.* [AFFECTED(y) & HUMAN(y)]

+> Peter knows that he was insulted.

+> Peter heard/read the insult himself.

+> Peter felt offended by the insult.

…but Peter had no perception, idea or feeling of it since he was absent in the last lecture.

c. *The woman deliberately frightened Peter.* [AFFECTED(y) & HUMAN(y)]

+> Peter felt something.

…#but Peter has no feeling of it since he was absent.

d. *The woman cheated on/manipulated Peter.* [AFFECTED(y) & HUMAN(y)]

?+> Peter knows that was cheated on/manipulated.

…but Peter has no idea of it since he has not recognized her true intentions.

e. *The woman bribed Peter.* [AFFECTED(y) & HUMAN(y)]

+> Peter knows that he was bribed.

…but Peter has no idea of it since he has not read the woman's message yet.

f. *The woman blackmailed/threatened Peter.* [AFFECTED(y) & HUMAN(y)]

+> Peter knows that he was blackmailed/threatened.

…but Peter has no idea of it since he has not read the woman's message yet.

g. *The woman defeated Peter.* [AFFECTED(y) & HUMAN(y)]

+> Peter knows that he was defeated.

…but Peter has no idea of it since he has not checked the final results yet.

h. *The woman called Peter.* [HUMAN(y)]

#+> Peter perceived the call.

i. *The woman greeted Peter.* [HUMAN(y)]

#+> Peter perceived the greeting.

j. *The woman listened to Peter.* [HUMAN(y)]

#+> Peter perceived her listening.

k. *The woman saw Peter.* [HUMAN(y)]

#+> Peter knows that he was seen.

l. *The woman admired/knew Peter.* [HUMAN(y)]

#+> Peter knows that he is admired/known.

The assumption of a GCI of sentience for the object is in line with the intuition, formulated in accounts on transitivity and affectedness, that human objects are perceived as *more affected* than non-human ones, even for the same verb. Hopper/Thompson (1980: 253) for instance state that, comparing the two sentences *I bumped into Charles* and *I bumped into the table*, we would rather expect there to be an effect on *Charles* than on *the table* even though both participants are involved in

the same event. In the same line of argumentation, Næss (2004: 1202, 2007: 111-113), besides *relational* affectedness, assumes a *subjective* dimension of the notion of affectedness:

> More specifically, effects on some types of entities are more salient to human perception and interests than others. An effect on a human participant is more likely to impinge directly on the lives of both the human in question and those surrounding him than an effect on an inanimate object; the situation described by *Peter killed John* is far more likely to have profound effects on all parties involved than, say, *Peter broke the pot*. In short, effects on human or animate entities are perceived as more dramatic, more significant, than effects on inanimates. (Næss 2004: 1202)

So perhaps (subjective) affectedness and relational sentience conveyed via GCI are just different ways to describe the same observation, namely that animate, unlike inanimate, patients are typically interpreted as psychologically affected in the given event. While Næss' explanation is based on the conceptualization of events in which humans are involved, I propose that it is also derivable from the verbal semantics in combination with the nominal property of animacy and concerns the *relation* of the object participant to the event in question.[9] There is, however, one important difference in that the accounts mentioned assume an increase of semantic transitivity in case of an animate patient, whereas I consider the implicated (proto-*agent* property) *sentience* for an object to decrease the transitivity of a predicate.

g. Subject potential of O

A criterion to identify potential (*inherent*) proto-agents in object position provided in the literature is their subject potential in the same event (Primus 2012a: 75ff.; García García 2014: 133f.; Kabatek 2016: 223f.). In the context of DOM, this test is closely connected to the distinguishing function stating that subject-like objects, being highly ranked in the animacy and referentiality scale, are preferably marked since they are potential competitors to the subject role (cf. e.g. Bossong 1991: 162f.). The test is simple and has already been mentioned in section 3.3: If an object has subject potential, it can hypothetically appear as the subject in the same event as illustrated in (49).

(49) a. *Pepe hirió a la actriz en la estación central.*

'Pepe injured the actress at the central station.'

Pepe:	+ auton. movement	*la actriz* :	+ affected
	+> + volition		+ *potential* auton. movement
	+> + sentience		+ *potential* volition
			+ *potentia*l sentience

b. *La actriz hirió a Pepe en la estación central.*

[9] From the general idea, this corresponds to de Swart's (2014) proposal for verbs of contact in Dutch. The crucial difference is, though, that he consideres *sentience* to be a proto-patient and not a proto-agent property (cf. 3.2.1, fn. 7).

'The actress injured Pepe at the central station.'

While in (49a), the object *la actriz* 'the actress' is a proto-patient (+ affected), its inherent properties (+ *potential* autonomous movement, + *potential* volition, + *potential* sentience) qualify it at the same time to function as a potential proto-agent in the same event, as illustrated in (49b). But note that instead of a reversal of the entire predication, as appears to be the case in (49b), we should rather speak of the *subject potential of the object for the same predicate*. This can be justified by the fact that a reversal of the entire predication would run into problems with non-causative (readings of) OE-psych verbs: A sentence like *La tormenta asusta a María* 'The thunderstorm frightens Maria' cannot be reversed *(#María asusta a la tormenta*). Nevertheless, the actual object can function as a potential subject of *asustar* 'to frighten'.

Applying the test to a number of verbs, however, it soon turns out that it is not very selective since it applies to all human objects (50a-h). As for inanimate objects, by contrast, it is only applicable to some predicates, i.e. *to resemble* (51b), *to precede* (51c) and *to defeat* (51f). For some verbs it is completely ruled out since these cannot take inanimate objects (in out-of-the-blue contexts) (51a, e, g).

(50) $\text{PREDICATE}_a(x, y) \rightarrow \text{PREDICATE}_a(y, \ldots)$ (with human objects)

a. *The woman fascinated/frightened Peter.* → Peter fascinated/frightened…

b. *The woman resembled Peter.* → Peter resembled…

c. *The woman called/greeted/helped Peter.* → Peter called/greeted/helped…

d. *The woman defeated Peter.* → Peter defeated…

e. *The woman cured/injured Peter.* → Peter cured/injured…

f. *The woman insulted Peter.* → Peter insulted…

g. *The woman saw Peter.* → Peter saw…

h. *The woman admired/knew Peter.* → Peter admired/knew…

(51) $\text{PREDICATE}_a(x, y) \rightarrow \text{PREDICATE}_a(y, \ldots)$ (with inanimate objects)

a. *The woman fascinated/frightened #the apple.* → n/a

b. *The woman resembled the apple.* → The apple resembled…

c. *Summer precedes autumn.* → Autumn precedes…

d. *The woman called the elevator.* → #The elevator called…

e. *The woman greeted/helped #the apple.* → n/a

f. *The woman defeated loneliness.* → Loneliness defeated…

g. *The woman cured/injured/insulted #the apple.* → n/a

h. *The woman saw the apple.* → #The apple saw…

i. *The woman admired/knew the book.* → #The book admired/knew…

In short, the diagnostic may not be the strongest one but it leads to more fine-grained distinctions than only animacy since it allows us to also identify inanimate objects that qualify as potential subjects of a given predicate (these overlap, at least partly, with the verbs to which the criterion (c) *no co-argument dependency* applies).

h. Summary

Table 2 summarizes the diagnostics and shows that – with the exception of the less selective test (g) – they allow us to differentiate between four levels of communicated agentivity for the object:

- *entailed agentivity* (either *volition*, *sentience* or *no co-argument dependency*),
- *presupposed agentivity* (either *autonomous activity* or *sentience*; additionally for some verbs *conversational implicature of sentience*),
- *potential agentivity* (*conversational implicature of sentience*) and
- *unspecified for agentivity* (no criterion applies).

The table indicates if a test is generally valid (✓), not valid (×) or partially valid (✓/×) for verbs representing the four levels of communicated agentivity. Certainly, the provided tests must be regarded as preliminary and need to be refined and further supplemented by other diagnostics.

Diagnostics	**Communicated agentivity for the object**			
	entailed	**presupposed**	**potential**	**unspecified**
a. Volition entailed	✓/×	×	×	×
b. Sentience entailed	✓/×	×	×	×
c. No co-argument dependency	✓/×	×	×	×
d. Auton. activity presupposed	×	✓/×	×	×
e. Sentience presupposed	✓/×	✓/×	×	×
f. Implicature of sentience	×	✓/×	✓	×
g. Subject potential	✓	✓/×	✓	✓/×
Example verbs	*to fascinate*, *to resemble*, *to frighten*	*to call*, *to greet*, *to help*, *to defeat*	*to cure, to insult*, *to injure*	*to admire*, *to know, to see*

Table 2. Applicability of diagnostics to the four levels of communicated agentivity for the object

4.2.2 Proposal of an agentivity scale

In the previous section, I identified four different types of verbs that can encode agentivity information for their object. These four types correspond to different levels of meaning: First, agentivity can be encoded as a proto-agent entailment in verbal

semantics. Second, it can be presupposed in a preceding sub-"event". Third, the object participant can be potentially agentive, namely by conversational implicature in relation to the event. A fourth class is constituted by verbs that do not pass any of the (selective) diagnostics and are hence classified as being unspecified for the object's agentivity. I propose that the four levels of agentivity, expressing different degrees of communicated agentivity for the object, can be ordered along the following scale:

(52) entailed agentivity for O > presupposed agentivity for O > potential agentivity for O > unspecified agentivity for O

Importantly, the communicated agentivity for the object always interacts with the agentivity resulting from the thematic distinctness of subject and object. While the former specifies the meaning dimension in which agentivity information is conveyed for the object, the latter considers the distribution of proto-properties among the two co-arguments. Both dimensions can be partially cross-classified as shown by the following table:

COMMUNICATED AGENTIVITY FOR O → ASYMMETRY ↓	entailed (e.g. *to fascinate*, *to resemble*, *to frighten*)	> presupposed (e.g. *to call, to greet, to defeat*, *to threaten*)	> potential (e.g. *to injure*, *to insult, to kill*)	> unspecified (e.g. *to admire, to know, to see*)
S *less* agentive than O	+	-	-	-
S as agentive as O	+	+	-	-
S *more* agentive than O	+	+	+	+

Table 3. Cross-classification of agentivity scale (horizontal) and agent-patient asymmetry (vertical)

Note that only the cells with a "+" show existing lexicalizations, while the combinations indicated with "-" are not lexicalized. Thus, entailed agentivity for the object can be represented by non-causative OE-psych verbs like *to fascinate*, for which the subject is less agentive than the object, by verbs with equally agentive co-arguments (e.g. *to resemble*) and by verbs for which the subject is more agentive than the object (causative reading of *to frighten*). For presupposed agentivity, by contrast, solely verbs are attested for which both arguments are at least equally agentive (e.g. *to call*, *to greet*) or for which the subject outranks the object in terms of agentivity (*to defeat*, *to threaten*). As for potential agentivity and unspecified for agentivity, only verbs with the subject being more agentive than the object are attested (see above).

The following section attempts to show that the agentivity scale could be helpful to systematize findings for DOM-sensitive verb classes in different Romance languages.

4.3 Agentivity scale and DOM in Romance languages: A survey of verb classes

The central question for the purposes of this study is whether the levels of agentivity defined phenomenon-independently are reflected in (morphosyntactic) argument realization and, more specifically, if they allow for a more discriminating view on the impact of agentivity on DOM in Romance languages. This section aims to systematize DOM-sensitive verb classes and relevant examples discussed in the literature with regard to the four levels of communicated agentivity for the object. With a focus being on Spanish, I will also present verb classes from several other Romance DOM languages. Each of the four levels of the proposed agentivity scale is dealt with in one subchapter (4.3.1-4.3.4).

4.3.1 Entailed agentivity for the object

I will first deal with the role-semantic configuration at the very left of the agentivity scale where a predicate entails agentivity to its object argument. Cross-classifying *entailed agentivity* with the agent-patient asymmetry, there are three possible configurations between subject and object argument: First, the subject can be *less* agentive than the object. In this case, solely *non-potent* proto-agent properties are assigned to either argument and the object bears *co-argument independency* or *sentience*. Second, the subject can be as agentive as the object. Here, both arguments bear either *volition* and *sentience* or *co-argument independency*. Third, the subject can be *more* agentive than the object. In this configuration, the subject is assigned *potent* proto-agent properties, whereas for the object proto-patient properties and the *non-potent* proto-agent property *sentience* overlap. Table 4 summarizes the three configurations and specifies for each possible verb classes and case frames according to Blume's (2000) linking model and García García's (2007, 2014) generalization of thematic distinctness. It is further indicated which of the three dimensions of argument realization differentiated in section 3.4 can be predicted to display variation for each configuration.

Configuration	Proto-properties of S	Proto-properties of O	Verb classes	Encodings in acc. languages	Dimension of argument realization with variation
S *less* agentive than O	none/ non-potent p-agent	non-potent p-agent	attribution verbs, OE-psych verbs, possession verbs	*/nom/acc*, */nom/dat*	a. Case frame b. Subject vs. DO c. ø DO vs. marked DO
S as agentive as O	potent/non-potent p-agent	potent/non-potent p-agent	(reversible-) symmetrical predicates	*/nom/acc*, */nom/dat*, */nom/PP*	a. Case frame b. Subject vs. DO c. ø DO vs. marked DO
S *more* agentive than O	potent p-agent	p-patient + non-potent p-agent	OE-psych verbs	*/nom/acc*	c. ø DO vs. marked DO

Table 4. Role-semantic configurations for *entailed agentivity*

In accordance with the observations made in section 4.1, the first two role-semantic configurations not only cover the cases where the generalization of thematic distinctness predicts DOM to occur with inanimate objects in Spanish (53a). They also include the cases of limited (or minimal) semantic transitivity of Blume (2000), which license marked case frames. Correspondingly, in Spanish, many predicates of configuration 1 such as *gustar* 'to like' are encoded by an indirect object (53b):

(53) **Configuration 1: S *less* agentive than O**

a. *caracterizar* 'to characterize' [Spanish]

La euforia caracteriza ***??ø/a*** *la situación.* */nom/acc + DOM*

the euphoria characterize.PRS.3SG ø/DOM the situation

'Euphoria characterizes the situation.' (García García 2018a: 228)

b. *gustar* 'to like' [Spanish]

A María le gusta la casa. */nom/dat*

IO PN CL.3SG.DAT like.PRS.3SG the house

'Maria likes the house.'

Besides transitive encodings of reversible-symmetrical predicates in Spanish, which systematically take DOM (54a), many predicates of configuration 2 select a prepositional object (54b) (cf. García García 2014: 148f.). This is also a cross-linguistically attested pattern of reciprocal predicates (cf. König/Kokutani 2006).

(54) **Configuration 2: S as agentive as O**

a. *igualar* 'to equal' [Spanish]

Los productos chinos igualan ***a*** *los productos europeos*

the products Chinese equal.PRS.3PL DOM the products European

en fiabilidad. */nom/acc* + DOM

in reliability

'Chinese products equal European products in reliability.'

(García García 2014: 159)

b. *equivaler a* 'to be equivalent to' [Spanish]

Los productos chinos equivalen ***a*** *los productos*

the products Chinese be equivalent.PRS.3PL to the products

europeos en fiabilidad. */nom/PP*

European in reliability

'Chinese products are equivalent to European products in reliability.'

Beyond Spanish, reversible-symmetrical predicates can be attested with DOM in other Romance languages that have not grammaticalized the marker to the same extent. This is the case for Catalan *seguir* 'to follow' if the object is expressed in preverbal position (55a), while (55b) proves the transitivity of the predicate by the direct object clitic *el*:

(55) *seguir* 'to follow' [Catalan]

a. ***Als*** *exemples de « però » inicial segueixen els exemples*

DOM.the examples of but initial follow.PRS.3PL the examples

de « però » intercalat. */nom/acc* + DOM

of but interspesed

'The examples of initial "but" are followed by the examples of interspersed "but".'

b. *El diumenge, ¿quien dia* ***el*** *segueix? – El dilluns.*

the Sunday which day CL.M.3SG.ACC follow.PRS.3SG the Monday

'Sunday, which day does follow? – Monday.' (Ruaix 1994: 149)

Likewise, Berretta (1991) provides an example from colloquial spoken Italian in which the transitive symmetrical predicate *sposare* 'to marry' (56) is *a*-marked. As in the previous example from Catalan, the position of the direct object in these

examples deviates from its canonical position in that it is placed in the sentence periphery. The left- or right-dislocation of the direct object, as well as its pronominal status, seem to be a general constraint for DOM to appear in Italian (Berretta 1991: 216-220).

(56) *sposare* 'to marry' [colloquial spoken Italian]

Lui	*ama*	*un'altra [...] la*		*vuole*	*sposare:*	***a***	*me*
3SG	love	another	CL.F.3SG.ACC	want.PRS.3SG	marry.INF	DOM	1SG
non	*mi*	*ha*	*mai*	*voluto*	*sposare*		
NEG	CL.1SG	have.PRS.3SG	never	want.PTCP	marry.INF		

'He loves another [...] he wants to marry her: he never wanted to marry me.'

(Berretta 1991: 229)

Moreover, there is a group of perception verbs in Spanish, amongst them *conocer* 'to know' and *encontrar* 'to find', for which two readings are available: a general non-reciprocal reading where the object is not agentive ('to know'; 'to find') and a reciprocal reading where the object's *sentience* is entailed ('to get acquainted with, enter in contact with'; 'to encounter, meet'). While the general meaning can occur without DOM, the specific meaning only arises in combination with the *a*-marker: e.g. *conocer* 'to know' vs. *conocer a* 'to get acquainted with, enter in contact with'.[10] Yet, unlike a comparison of Delbecque (2002: 93f.) might suggest, the difference between the two meanings is not so clear-cut: *conocer a* can also mean 'to know'. However, for my argumentation it is interesting to observe that the symmetrical reading for these verbs correlates with the use of the *a*-marker. Note, however, that in Portuguese and Catalan, where also both readings are available for *conhecer* (Pt.), *conéixer* (Cat.) 'to know' and *encontrar* (Cat., Pt.) 'to find', the reciprocal reading can be perfectly obtained also in the absence of DOM as shown for *encontrar* in the examples below:

(57) *encontrar* 'to find, meet' [Portuguese]

Encontrei	***ø***	*o*	*teu*	*irmão.*
find.PST.1SG	ø	the	POSS.2SG	brother

'I met your brother.'

(58) *encontrar* 'to find, meet' [Standard Catalan]

He	*encontrat*	***ø***	*el*	*teu*	*germà.*
have.PRS.1SG	find.PTCP	ø	the	POSS.2SG	brother

[10] For *conocer*, also the employed tense form plays a role for its meaning in a given context (stative non-agentive vs. inchoative agentive reading but also aspectual coercion possible).

'I met your brother.' (Bossong 1998: 225)

With regard to DOM, the third role-semantic configuration deserves special attention since it comprises overwhelmingly transitively encoded verbs. Examples can be given by the causative OE-psych verb *asustar* 'to frighten' (59a) and the verb *despertar* 'to wake up' (59b) in Spanish, where the object is entailed both the proto-patient property *affected* and the (non-potent) proto-agent property *sentience*:

(59) a. *asustar* 'to frighten' [Spanish]

María asustó ***al*** *niño por diversión.*

PN frighten.PST.3SG DOM.the child for fun

'Maria frightened the child for fun.'

b. *despertar* 'to wake up' [Spanish]

María despertó ***al*** *niño a las siete.*

PN wake up.PST.3SG DOM.the child at the seven

'Maria woke the child up at seven o'clock.'

Consequently, for the levels (a) and (b) of argument realization pointed out in section 3.4, i.e. the selection of case frame and the selection of subject and object, this subclass is not expected to display variation: concerning the realization of subject and object, the argument with potent proto-agent properties is linked to the nominative per default. As for the selection of case frame, the licensing condition for marked valencies (Blume 2000: 200) is not fulfilled by this class since the two arguments show an asymmetric assignment of proto-properties. Hence, it can be predicted that the verbs of this class select for a */nom/acc* case frame. However, the fact that the object is entailed proto-patient and non-potent proto-agent properties at the same time matches with one of the overgeneralizations of the licensing condition for marked valencies put forward by Blume (2000: 211-217): In languages that rank the semantic licensing condition higher than formal constraints (e.g. Finnish or Czech), marked case frames systematically appear with highly transitive verbs showing an asymmetric argument relation under the condition that the object is assigned proto-agent properties at the same time. Blume (2000: 217) labels this argument relation *pseudo-symmetric*.

Besides two-place predicates, the overlap of entailed proto-patient and (non-potent) proto-agent properties is a role-semantic configuration we find for recipients and addressees in ditranstive constructions. It is precisely from this role overlap of recipients that Primus (2012a: 73) starts her argumentation of why animate direct objects get *a*-marked in Spanish, namely because they involve a similar role-overlap: Just as recipients like *el padre* in (60a) are proto-patients in relation to the subject and proto-agents in relation to the direct object, animate direct objects such as *el ladrón* in (60b)

are *actual* proto-patients according to their *entailed* proto-properties but *potential* proto-agents due to their *inherent* meaning.

(60) a. *dar* 'to give' [Spanish]

*La madre le da el niño ***el/al** padre.*

the mother CL.3SG.DAT give.PRS.3SG the child the/IO.the father

'The mother gives the child to the father.' (Primus 2012a: 67)

b. *golpear* 'to hit' [Spanish]

*La actriz golpeó **al** ladrón.*

the actress hit.PST.3SG DOM.the thief

'The actress hit the thief.'

Primus interestingly comments on the relation between entailed and potential proto-agent properties in order to account for the fact that in ditransitives in Standard Spanish, only the indirect but not the direct object is *a*-marked: in her argumentation, *entailed* proto-agent properties for one argument outrank *potential* proto-agent properties:

> An additional explanation is needed for the suppression of DOM in ditransitive clauses, as shown in the Spanish example [(60a)] above. This may be due to a higher-ranking constraint that bans multiple datives in Standard Spanish (as in some other languages). Under these circumstances, two roles compete for the dative: the recipient and the animate patient. In Standard Spanish, the competition is resolved in favour of the proto-agent components that are **entailed by the meaning of the verb**, i.e. subcategorized for by the verb. This means that the dative is used for the recipient and blocked for the animate patient. (Primus 2012a: 79; emphasis: S.M.)

Whether such differentiation can be applied to two-place predicates as well is one central question of this section. Taking into account DOM-sensitive verb classes across Romance languages, it is interesting to note that the class of OE-psych verbs, where the object is entailed *sentience*, frequently turns up. In Spanish, this class has generalized the *a*-marker (Delbecque 2002: 104f.), with example verbs being listed in (61a). In Italian, where DOM is only a marginal phenomenon, it is with OE-psych verbs, as the ones in (61b), the *a*-marker typically occurs (Berretta 1991: 221).

(61) **DOM-sensitive (non-causative and ± causative) OE-psych verbs**

a. Spanish (selection, for complete list cf. Delbecque 2002: 105, fn. 25f.):

abrumar 'to overwhelm', *aburrir* 'to bore', *afectar* 'to affect', *alentar* 'to encourage', *amedrentar* 'to frighten', *aquejar* 'to torment', *asustar* 'to frighten', *atraer* 'to attract', *aturdir* 'to astonish', *azorar* 'to embarrass', *cansar* 'to tire', *cautivar* 'to captivate', *conmover* 'to move', *consolar* 'to console', *contrariar* 'to annoy', *embelesar* 'to charm', *encantar* 'to enchant', *entretener* 'to entertain', *enfadar* 'to make angry', *enojar* 'to annoy', *exasperar* 'to exasperate', *excitar* 'to excite', *fascinar* 'to fascinate', *fastidiar* 'to bother', *irritar* 'to irritate', *molestar* 'to bother', *ofender* 'to offend', *oprimir* 'to oppress', *perturbar* 'to perturb', *preocupar* 'to

preoccupy', *satisfacer* 'to satisfy', *reconfortar* 'to comfort', *satisfacer* 'to satisfy', *seducir* 'to seduce', *serenar* 'to calm', *sosegar* 'to tranquillize', *sorprender* 'to surprise'

b. Italian (cf. Berretta 1991: 221):

affascinare 'to fascinate', *attrarre* 'to attract', *colpire* 'to impress, strike', *confortare* 'to comfort', *consolare* 'to cheer up', *convincere* 'to convince', *deludere* 'to disappoint', *disturbare* 'to disturb', *divertire* 'to entertain', *eccitare* 'to excite', *entusiasmare* 'to excite', *incantare* 'to enchant', *ingannare* 'to deceive', *innervosire* 'to irritate', *invitare* 'to tempt', *mettere (di buon umore)* 'to cheer', *persuadere* 'to persuade', *preoccupare* 'to worry', *rattristare* 'to sadden', *spaventare* 'to frighten'

Interestingly, among the examples of DOM in Standard Catalan provided by Solà (1994: 167f.), there is also a large amount of (non-causative and ± causative) OE-psych verbs such as *irritar* 'to irritate', *satisfer* 'to satisfy', *atreure* 'to attract' and *molestar* 'to bother'. The corresponding example sentences are listed in (62). Note that in all four cases, the direct object occupies a non-canonical position being either pre-posed or left-dislocated.

(62) a. *irritar* 'to irritate' [Standard Catalan]

A *R. Menéndez Pidal, l'* *irritaven* *els* *escrits*

DOM PN CL.M.3SG.ACC irritate.IPRF.3PL the writings

de Las Casas.

of Las Casas

'The writings of Las Casas irritated R. Menéndez Pidal.'

b. *satisfer* 'to satisfy' [Standard Catalan]

¿Satisfà *que* *hi hagi* *un comitè* *de direcció?*

satisfy.PRS.3SG that there is.PRS a committee of steering

A *molts escriptors,* *no.*

DOM many writers no

'Are you satisfied that there is a steering committee? Many writers, no.'

c. *atreure* 'to attract' [Standard Catalan]

Però *a més* *de l'* *erudició,* ***a*** *Galiana* *l'*

but in addition to the erudition DOM PN CL.F.3SG.ACC

atreia *la* *literatura festiva. [...]*

attract.IPRF.3SG the literature festive

'But in addition to his erudition, Galiana was attracted to the festive literature.'

d. *molestar* ‘to bother’ [Standard Catalan]

A	*Mairem,*	*la*	*molestà*	*el*	*suggeriment. [...]*
DOM	PN	CL.F.3SG.ACC	bother.IPRF.3SG	the	suggestion

‘The suggestion bothered Mairem.’ (Solà 1994: 168)

Furthermore, apart from OE-psych verbs, also the verb *svegliare* ‘to wake up’, which entails the object’s sentience, is attested with the *a*-marker in Italian:

(63) *svegliare* ‘to wake up’ [colloquial spoken Italian]

a	*noi*	*c’*	*ha*	*svegliato*	*il*	*telefono, […]*
DOM	1PL	CL.1PL	have.PRS.3SG	wake up.PTCP	the	phone

‘To us, the phone woke us up […].’ (Berretta 1991: 229)

Besides proper verbal semantics, there is also evidence from two constructions strengthening the assumption that entailed proto-agent properties for the direct object strongly favour DOM: *Accusativus cum Infinitivo constructions* (AcI) and *causative constructions*. The two constructions are exemplified for Spanish in the following examples:

(64) a. AcI with the visual perception verb *ver* ‘to see’ [Spanish]

Pepe	*vio*	***a***	*María*	*bajar*	*del*	*autobús.*
PN	see.PST.3SG	DOM	PN	get off.INF	of the	bus

‘Pepe saw Maria get off the bus.’ (García García 2014: 51)

b. Causative construction with the matrix verb *hacer* ‘to make’ [Spanish]

Pepe	*hizo*	*trabajar*	***a***	*Juan.*
PN	make.PST.3SG	work.INF	DOM	PN

‘Pepe made Juan work.’

AcI and causative constructions do not fit into Table 4 above for the following reason: Unlike, for example, OE-psych verbs, the implied agentivity of their object is not due to a proto-agent entailment for the object which is encoded in the lexical semantics of a verb. Instead, the constructions consist of a matrix and an embedded clause with the direct object being the subject of the embedded clause. The proto-agent property is actually a property assigned to the subject of the embedded infinitive in question, such as to the subject of *bajar* ‘to get off’ in (64a) or *trabajar* ‘to work’ in (64b). AcI constructions are reported to show a preference for DOM both synchronically and diachronically (García García 2014, 2018a; Enghels 2007, 2013; Laca 2006). Müller (1971: 500f.) even discusses *a*-marking with the perception verbs *ver* ‘to see’ and *oír* ‘to hear’ as *one* source of the differential object marker in Spanish.

Also, among the examples for DOM in other Romance languages, we often find AcI and causative constructions. This can be illustrated by the example from Standard Catalan in (65). Here, the differential marker serves to dissolve an ambiguity which arises from the fact that *el jutge* 'the judge' could be both interpreted as the subject and as the direct object.

(65) *fer riure* 'to make laugh' [Standard Catalan]

L'	*advocat*	*defensor*	*va*	*fer*	*riure*	***a***	*tothom*
the	lawyer	defense	have.PST.3SG	make.INF	laugh.INF	DOM	everyone

menys	***el/al***	*jutge.*
less	the/DOM.the	judge

'The defense lawyer made everyone laugh, less the judge.' (Solà 1994: 166)

Likewise Italian causative constructions, as the ones listed in (66a), display an affinity for the *a*-marker. As already mentioned above for OE-psych verbs, there are further constraints to be fulfilled: the object must be dislocated and preferably pronominal, as can be illustrated by the example in (66b):

(66) a. CAUSATIVE CONSTRUCTIONS with *fare* 'to make'/ *lasciare* 'to let' + infinitive:

fare dormire 'to make sleep', *fare ingrassare* 'to fatten', *fare morir dal ridere* 'to make die laughing', *fare ridere* 'to make laugh', *fare piangere* 'to make cry', *fare sentir male* 'to make feel bad', *lasciare salire* 'to pick up', *lasciare stare* 'to let be/go' (Berretta 1991: 221)

b. *fare dormire* 'to make sleep' [colloquial spoken Italian]

a	*me*	*il*	*sonnifero*	*ha*	*fatto*	*dormir*	*bene*
DOM	1SG	the	sleeping pill	have.PRS.3SG	make.PTCP	sleep.INF	well

'To me the sleeping pill has made me sleep well.' (Berretta 1991: 226)

The presented constructions and their preference for DOM further support an explanation of the phenomenon in terms of agentivity. Though in the case of AcI and causative constructions, the agentivity of the object arises differently than with OE-psych verbs, which lexically entail agentivity to the object, apparently, the thematic distinctness of subject and direct object is blurred in a similar way.[11] Yet, for the

[11] Interestingly, in many Romance languages, causative constructions show variation between transitive and intransitive encodings (cf. Primus 1999b: 161–163; Ackerman/Moore 1999, 2001: 73–79; Enghels/Roegiest 2012a, b, 2013; for Catalan Solà 1994, for French Lamiroy/Charolles 2011). This variation is mostly shown for the object clitic, with both options being theoretically possible in most of the cases (e.g. *Pedro* ***lo*** *hizo trabajar* vs. *Pedro* ***le*** *hizo trabajar*). The different encodings are typically accompanied by different meanings: while the direct object clitic (*lo*) suggests a patient-like object, the indirect object clitic (*le*) points to an agent-like object (certainly, in Spanish also the phenomenon of *leísmo*/*laísmo* has to be considered). In sum, this is further support for the licensing condition for marked case frames put forward by Blume (2000) predicting that marked case frames may appear when the typical agent-patient asymmetry is absent.

remainder of this study, I will focus on proper verb semantic factors and will not deal with the mentioned verbal constructions.

To sum up, combining entailed proto-agent properties with the thematic asymmetry between subject and object, there are three different role-semantic configurations that can be ordered in a continuum of prototypical semantic transitivity: when the subject is *less* agentive than the object, we find the most considerable deviation from the typical thematic distinctness of subject and object. Also, an equal degree of agentivity of subject and object deviates from the prototypical case, namely the one where the subject is *more* agentive than the object. As for morphosyntactic argument realization, the following observation can be made: the more the deviation from the prototypical subject-object asymmetry, the greater the variation of argument realization pattern seems. This concerns the three dimensions of argument realization differentiated in section 3.4:

(67) a. Case frame (ACC, DAT…)

b. Subject vs. DO

c. ø DO vs. a DO

According to Blume (2000), marked valencies are licensed when a two-place predicate lacks an asymmetric distribution of proto-roles. This can be attested with respect to the configurations, where the subject is either less agentive than the object or where both arguments bear the same level of agentivity. In these two cases, not only the assignment of case frame (67a) but also the selection of subject and object is variable (67b). Moreover, for transitive verbs, as shown for (reversible-) symmetrical and non-causative OE-psych verbs in different Romance languages, these configurations strongly favour or even require DOM (67c). In contrast, when subject and object bear an asymmetric argument relation, marked case frames are not licensed and the selection of subject and direct object can unambiguously take place. However, the fact that the object also shows one proto-agent property (*sentience*), as for causative readings of OE-psych verbs, seems to blur the argument relation and favour DOM (67c).

4.3.2 Presupposed agentivity for the object

It has been shown by Blume (1998, 2000) that not only *entailed* but also *presupposed* proto-agent properties affect morphosyntactic argument realization. Importantly, we must distinguish between the first two levels of argument realization. While for (67a) case selection, presupposed proto-agent properties of the object matter and offer an explanation for the cross-linguistic preference of interaction verbs to take marked valencies, for (67b) the selection of subject and object, presupposed proto-agent properties do not compete with entailed proto-agent properties: the latter are linked to the nominative per default (Blume 1998: 269, 2000: 224). It is plausible that also within the realm of differential morphosyntactic argument realization, such as DOM, presupposed proto-agent properties should be considered, a question that has not been explicitly tackled in previous role-semantic accounts on DOM. This section

wants to show that *presupposed agentivity* helps to subsume several verb classes mentioned for DOM in Spanish and other Romance languages under one category.

For the present class, two role-semantic sub-configurations can be distinguished as indicated in Table 5. The first is one where the subject is as agentive as the object, that is the subject bears one or more entailed proto-agent properties while the object exhibits one or more presupposed proto-agent properties. In the second configuration the subject outranks the object in terms of agentivity, though for the object entailed proto-patient properties overlap with one or more presupposed proto-agent properties.

Configuration	**Proto-properties of S**	**Proto-properties of O**	**Verb classes**	**Encodings in acc. languages**	**Dimension of argument realization with variation**
S as agentive as O	potent p-agent	presupposed p-agent	(1a) verbs of communication and social gesture, (1b) verbs of helping and hindering/auditory perception/movement	*/nom/acc*, */nom/dat*, */nom/PP*	a. Case frame c. ø DO vs. marked DO
S *more* agentive than O	potent p-agent	p-patient + presupposed p-agent	(2a) verbs of threatening/cheating/punishment (2b) verbs of competition (and stopping)	*/nom/acc*	c. ø DO vs. marked DO

Table 5. Role-semantic configurations for *presupposed agentivity*

For each sub-configuration, two main verb classes can be distinguished according to the two types of presupposition worked out in section 4.2.1 (d.-e.). While the verb classes (1a) and (2a) consider a presupposition of *sentience* for the object, the classes (1b) and (2b), in turn, can be ascribed presupposed *autonomous activity* in a prior subevent for the object.

In (68), verbs of communication and social gesture frequently attested with DOM are listed for different Romance languages. As can be seen, these verbs not only appear to have a preference for DOM in languages where the phenomenon is highly grammaticalized, as in Spanish, but also in languages where it is in a medium-level (Catalan) or even initial stage (Italian and French).

(68) **(1a) Verbs of communication and social gesture**

a. **Spanish:**

saludar 'to greet' (Fernández Ramírez 1986: 159; Leonetti 2004: 84; Roegiest 1979: 50), *invitar* 'to invite' (Fernández Ramírez 1986: 159), *bautizar* 'to baptize' (Roegiest 1979: 50), *llamar* 'to call', *obedecer* 'to obey' (García García 2014: 187, 189)

b. **Catalan:**

amoixar 'to flatter' (Wheeler et al. 1999: 462), *convidar* 'to invite', *cridar* 'to call', *votar* 'to vote' (Solà 1994: 165, 167); *invitar* 'to invite', *saludar* 'to greet' (Escandell-Vidal 2009: 838)

c. **Italian:**

chiamare 'to call', *informare* 'to inform', *ringraziare* 'to thank', *salutare* 'to greet' (Berretta 1991: 222)

d. **French:**

interroger 'to question', *prier* 'to ask' (Fagard/Mardale 2014: 158)

Two relevant examples from Catalan and Italian are given below. In both cases, the direct object appears in a dislocated structure: a right-dislocation in (69) and a left-dislocation in (70).

(69) *amoixar* 'to flatter' [Catalan]

Ja sé com l' hai d' amoixar, ***a***

now know.PRS.1SG how CL.F.3SG.ACC have.PRS.1SG to flatter.INF DOM

Na Rosó.

OM PN

'Now I know how to flatter her, (that is,) Rosó.' (Wheeler et al. 1999: 462)

(70) *informare* 'to inform' [colloquial spoken Italian]

Ma questi ***a*** *noi non c' hanno informato per niente*

but these DOM 1SG NEG CL.1PL have.PRS.3PL inform.PTCP at all

'But they haven't informed us at all.' (Berretta 1991: 227)

Insights for a possible impact of *presupposed agentivity* on DOM come from inanimate objects in Spanish that can be ascribed presupposed sentience in the event denoted by the predicate as illustrated in (71):

(71) *llamar* 'to call' [Spanish]

¡Hans, puñeta, llama ***al*** *ascensor!*

PN damn call.PRS.3SG DOM.the elevator

'Hans, damn, call the elevator!'

(García García 2014: 189; García García/Primus/Himmelmann 2018: 30)

Here, the inanimate *a*-marked object can be explained as follows: since an elevator is programmed to perceive a certain signal and to react to it, it can be qualified as sentient, and hence agentive, in the given event (cf. García García/Primus/Himmelmann 2018: 32). But, unlike for human beings, the sentience of elevators is restricted to the programmed stimulus, which is why they cannot context-independently fill the slots of other predicates entailing or presupposing sentience (e.g. 'love', 'cuddle', 'be jealous'), while humans can. Hence, the example in (71) nicely illustrates the relative dimension of presupposed proto-agent properties and its possible impact on DOM, at least for Spanish.

As for the following class, verbs of auditory perception can be emphasized: unlike visual perception verbs, they can be classified as interaction verbs since they presuppose that the object participant "produces an aesthetic or meaningful chain of sounds" (Blume 1998: 273). On the morphosyntactic level this is reflected by their preference for marked valencies such as */nom/dat* or, as suggested by (72), for differentially marked direct objects.

(72) **(1b) Verbs of helping and hindering/ auditory perception/ movement**

a. **Spanish:**

escuchar 'to listen to', *oír* 'to hear' (Enghels 2007: 243), *proteger* 'to protect' (Fernández Ramírez 1986: 159), *temer* 'to fear' (García García 2014: 187)

b. **Catalan:**

escoltar 'to listen to' (Solà 1994: 165f.), *ajudar* 'to help' (Escandell-Vidal 2009: 838, 845)

c. **Italian:**

accompagnare 'to accompany', *aspettare* 'to wait for', *proteggere* 'to protect', *temere* 'to fear' (Berretta 1991: 222)

d. **French:**

aider 'to help', *applaudir* 'to applaud', *écouter* 'to listen to', *empêcher* 'to hinder', *soigner* 'to take care of' (Fagard/Mardale 2014: 158)

As for the Spanish verbs *castigar* 'to punish' and *sobornar* 'to bribe' in (73a), it can be noted that they also take DOM obligatorily with indefinite human object NPs (Leonetti 2004: 84). The verb *amenazar* 'to threaten' is attested also with *a*-marked inanimate objects (García García 2014: 187).

(73) **(2a) verbs of threatening/ cheating/ punishment**

a. **Spanish:**

castigar ‘to punish’, *sobornar* ‘to bribe’ (Leonetti 2004: 84; Torrego 1999: 1788), *amenazar* ‘to threaten’ (Fernández Ramírez 1986: 176, García García 2014: 187)

b. **French:**

escroquer ‘to swindle’ (Fagard/Mardale 2014: 158)

In (74), verbs of competition, which have already been discussed in section 4.1 with regard to their DOM preference in Spanish, are listed.

(74) **(2b) Verbs of competition**

a. **Spanish:**

vencer ‘to win’ (Roegiest 1979: 50; García García 2014: 178), *derrotar* ‘to win, defeat’, *defender* ‘to defend’ (García García 2014: 178)

b. **Catalan:**

guanyar ‘to win’ (Solà 1994: 166)

In sum, the consideration of presupposed proto-agent properties looks promising for a fine-grained agentivity-based differentiation of verb classes with DOM preferences across Romance languages. It is in particular noticeable that also in languages that have not yet grammaticalized DOM, like Italian and French, a range of verbs attested with the marker can be assigned to this agentivity class.

4.3.3 Potential agentivity for the object

As for potential agentivity, there is only one role-semantic configuration, namely only one where the subject outranks the object in terms of agentivity. As will be shown, the major challenge for this class is to disentangle the nominal-based property animacy from the verb-based property agentivity.

Configuration	**Proto-properties of S**	**Proto-properties of O**	**Verb classes**	**Encodings in acc. languages**	**Dimension of argument realization with variation**
S *more* agentive than O	potent p-agent	p-patient + implicated p-agent	verbs of (physical and verbal) injuring, verbs of killing	*/nom/acc*	c. ø DO vs. marked DO

Table 6. Role-semantic configurations for *potential agentivity*

The verbs relevant for this class are often described as verbs that exclusively select for animate (or human) direct objects, i.e. as verbs that have an animacy restriction for the direct object. For Spanish, Torrego (1999: 1787) presents the following example for the verb *insultar* ‘to insult’, which takes the *a*-marker also obligatorily with indefinite human objects:

(75) *Marta insultó* ****ø/a*** *un compañero.*

PN insult.PST.3SG ø/DOM a friend

'Marta insulted a friend.' (Torrego 1999: 1787)

In Torrego's interpretation, DOM is obligatory with the verb *insultar* 'to insult' due to its aspectual property of being telic. Likewise, verbs like *curar* 'to cure' and *emborrachar* 'to make drunk' are subsumed under the same analysis. While there are several problems with the telicity hypothesis (cf. Romero Heredero 2022), which I will not discuss here, it is noticeable that the mentioned verbs also share two other properties, namely their strong preference for human direct objects (García García 2018a: 222f.) as well as their selection of an *affected* object participant.

An analysis in terms of humanness preference for the direct object of certain verbs is further supported by diachronic as well as synchronic findings from different Romance languages. As for Spanish, in a diachronic corpus study consisting of data from the 12th until the 20th century conducted by von Heusinger (2008), it turned out that the verbs *matar* 'to kill' and *herir* 'to hurt, wound' show a significantly higher preference for DOM than the verbs *ver* 'to see' and *hallar* 'to find' which in turn occur significantly more frequent with DOM than *tomar* 'to take' and *poner* 'to put'. Given that these results correspond to the preference for a verb in question to select for a [+ human], [± human] or [(±)/- animate] direct object, von Heusinger (2008: 15) establishes the following scale:

(76) Scale of verb classes according to selectional preferences

[+ human] >	[± human] >	[(±)/- animate]
Class 1	Class 2	Class 3
matar / herir	*ver / hallar*	*tomar / poner*
'kill' / 'hurt'	'see' / 'find'	'take' / 'put' (von Heusinger 2008: 15)

Interestingly, also for Old Portuguese (Delille 1970: 63), *matar* 'to kill' and *ferir*[12] 'to hurt, wound' are among the verbs that show a preference for DOM in the 16th century. Likewise, for Modern Sicilian the verb *ammazzari* 'to kill' is characterized as a verb that "regge oggetti preposizionali con altissima frequenza" (Guardiano 2000: 29). Also, for Modern French, which under certain syntactic and semantic conditions shows rare instances of DOM, it has been observed that only verbs of the classes 1 [+ human] and 2 [± human] take the marker (Fagard/Mardale 2014: 155). Among the verbs that are frequently attested with DOM we find again *blesser*[13] 'to hurt, injure', *insulter* 'to insult' and *tuer* 'to kill' (Fagard/Mardale 2014: 158). In

[12] Please note that *ferir* 'to hurt, wound' can also have a second meaning which refers to the psychological sphere, namely 'to offend'. In this case, it could be interpreted as an OE-psych verb.
[13] Note that also French *blesser* 'to hurt, injure' can have a second, psychological reading 'to offend'.

order to account for such preference of the mentioned Spanish verbs in (76), von Heusinger (2008) states:

> Frequency shows the preference of one type of verb for a certain type of animacy: the higher the preference for human direct objects in general, the higher the preference to mark such a human direct object. Note that for this contrast it is not the animacy of the direct object, but the preference of the verb for a human direct object that matters. This observation needs more investigation, in particular we need to know what property of the verb might correlate with its preference towards human objects. (von Heusinger 2008: 29)

Certainly, one could also assume a selectional restriction (in a broader sense) for the object here, namely [+ animate], that leads to a higher frequency of the marker. However, such an explanation would be not restrictive enough and not coherent in the present role-semantic account. Since a great variety of the verbs with *entailed agentivity* (e.g. OE-psych verbs) and *presupposed agentivity* (interaction verbs) most typically select for an animate – mostly human – direct object, too, a selectional restriction [+ animate] would not allow us to differentiate between the three classes (cf. also García García/Primus/Himmelmann 2018: 35). Instead, we have already seen in the diagnostics in 4.2 that a consideration of the *relational* property agentivity and the types it is encoded in the verbal semantics can lead to a more fine-grained classification of predicate classes. Here, it can be noted that verbs with *potential agentivity* do not pass the diagnostics (a.-e.), while *entailed* and *presupposed agentivity*, respectively, do. Instead, they all convey the generalized conversational implicature of sentience for the object (cf. diagnostic f.). Since the affectedness of the object, given for verbs like Spanish *matar* 'to kill' or *herir* 'to hurt, wound', is a precondition for the implicature to arise, this analysis is also compatible with the finding of von Heusinger/Kaiser (2011) that the verbal factor affectedness (partly) correlates with the diachronically attested DOM preference of certain verb classes in Spanish.

Note that the notion of *relational* potential agentivity employed here must be distinguished from *inherent* potential agentivity as used in the following licensing condition for DOM in Spanish:

(77) A differential object marker is licensed by an object qualifying as a minimal or **potential proto-agent** in a given event. The proto-agent properties may either be subcategorized by the verb or assigned according to **intrinsic properties** of the object referent. (García García/ Primus/Himmelmann 2018: 27; emphasis: S.M.)

In (77), the starting point is the (cognitive-) ontological status of an animate object referent which can be ascribed *intrinsic* (or *inherent*) proto-agent properties. From this perspective, there are three proto-agent properties that unilaterally imply the animacy of the participant in question, namely *control*, *possession* and *experience* (in the terminology of Primus' 1999a, b proto-role model):

(78) Unilateral implications from proto-agentivity to animacy (García García 2014: 131):

a. CTRL(x, …) → HUMAN(x, …) → ANIMATE(x, …)

b. POSSESS(x, …) → HUMAN(x, …) → ANIMATE(x, …)

c. EXPER(x, …) → ANIMATE(x, …)

The inference from animacy to potential agentivity is then explained by abductive reasoning, i.e. the reversal of the implications in (78a-c) above. In this way, *independent* of their actual assignment of proto-properties in a given event, animate – especially human – object referents are *a priori* interpreted as potential controllers, possessors or experiencers:

(79) Reversed implications (by abductive reasoning, cf. Primus 2012a: 85):

a. HUMAN(x, …) → CTRL(x, …)

b. HUMAN(x, …) → POSSESS(x, …)

c. ANIMATE(x, …) → EXPER(x, …)

As argued by Primus, animacy-driven DOM can be role-semantically explained as a "role overlap" for the object participant that "occurs whenever a noun phrase subcategorized for patient properties acquires potential agent properties due to its intrinsic meaning" (2012a: 73). As a consequence, a clear-cut asymmetric co-argument dependency in which the proto-patient is only defined by its dependency from a proto-agent, expressed by the role hierarchy proto-agent $>_{dep}$ proto-patient, is blurred. In contrast, relational potential agentivity is *verb-dependent* and considers the – *empirically justified* – implicated sentience of the object participant relative to the event in question (for further discussion cf. section 6.2.1).

To conclude, what characterizes the verbs of this class are the following two points: (i) they all select for an animate – typically human – direct object, and (ii) they entail the object's affectedness. As has been argued, based on these two preconditions, the object participant is interpreted as being sentient relative to the event, too.

4.3.4 Object unspecified for agentivity

The level *unspecified for agentivity*, located at the low end of the agentivity scale, can be characterized as very heterogeneous. The common denominator of all verb classes that can be subsumed under it is the role-semantic configuration that the subject is *more* agentive than the object, with the object not conveying any type of proto-agent properties. As a closer look at verb classes and DOM in the literature reveals, it seems convenient to make a subdivision into three sub-configurations which are listed in the following table:

Configuration	Proto-properties of S	Proto-properties of O	Verb classes	Encodings in acc. languages	Dimension of argument realization with variation
S *more* agentive than O	potent p-agent	p-patient	destruction/ consumption/ creation verbs, contact verbs	*/nom/acc*	c. ø DO vs. marked DO
	potent p-agent	none	controlled perception verbs	*/nom/acc*, */nom/PP*	a. Case frame c. ø DO vs. marked DO
	non-potent p-agent	none	SE-psych verbs, knowledge/ perception verbs, possession verbs, verbs with cognate objects/property arguments/light verb constructions	*/nom/acc*	a. Case frame b. Subject vs. DO c. ø DO vs. marked DO

Table 7. Role-semantic configurations for *unspecified for agentivity*

The core case of the first sub-configuration are maximally transitive verbs where the subject clearly outranks the direct object in terms of agentivity. These comprise in particular destruction, consumption and creation verbs. As shown by García García (2014: 194-201) as a proof of his generalization of thematic distinctness, DOM is ungrammatical for these verb classes:

(80) **Highly transitive verbs with maximally affected (or effected) objects [Spanish]**

a. **Destruction verbs:**

derrumbar ***ø/*a*** *la casa, cancelar* ***ø/*a*** *la reunión*

'to collapse the house, cancel the meeting'

b. **Consumption verbs:**

comer(se) ***ø/*a*** *la manzana, tomar(se)* ***el/*al*** *café*

'to eat the apple, to drink the coffee'

c. **Creation verbs:**

escribir ***ø/*a*** *la carta*, *causar* ***ø/*a*** *la enfermedad*

'to write the letter, to cause the disease' (García García 2014: 194)

There are two further verb classes of highly transitive verbs that should be mentioned here. The first are Spanish verbs like *seguir* 'to follow, continue', *sustituir* 'to replace', *suplantar* 'to forge' or *superar* 'to overcome', which have a second reversible reading denoting a minimally transitive event (cf. García García 2014: 147-170). While in the latter reading, DOM is obligatory (e.g. *El otoño sigue *el/al verano*

'Autumn follows summer'), in the given non-reversible reading the *a*-marker is absent, as exemplified below for *suplantar* 'to forge' (81a) and *seguir* 'to follow, continue' (82b).

(81) **Highly transitive verbs (with a second reversible reading) [Spanish]**

a. *suplantar* 'to forge'

El seguro ha suplantado ***ø/*a*** *la póliza.*

the insurance have.PRS.3SG forge.PTCP ø/DOM the policy

'The insurance has forged the policy.'

b. *seguir* 'to follow, continue'

Sigan ***ø*** *esta calle hasta el final [...]*

follow.IMP.3PL ø this street to the end

'Follow this street to the end […]' (García García 2014: 152; 165)

Note that, in case these predicates have a reversible reading, they would shift from class 4 (*unspecified for agentivity*) to class 1 (*entailed agentivity*) of the agentivity scale (cf. 4.3.1). The second class can be broadly classified as contact verbs[14] including the Spanish verbs listed in (82).

(82) **Contact verbs [Spanish]**

abrazar 'to embrace', *acariciar* 'to caress', *agarrar* 'to grab', *besar* 'to kiss', *golpear* 'to beat', *tocar* 'to touch'

These verbs bear a role-semantic particularity: While with an inanimate object, the participant is interpreted as only physically affected, with an animate – especially human – direct object, besides physical affectedness *sentience* is conveyed via generalized conversational implicature. In the former case, the marker is ungrammatical (83a) while it is compulsory in the latter case (83b). The presence of an implicature is proved by its cancellability in (83c).

(83) *besar* 'to kiss' [Spanish]

a. *El sacerdote besó* ***el/*al*** *altar.*

the priest kiss.PST.3SG the/DOM.the altar

'The priest kissed the altar.'

b. *Pepe besó* ****ø/a*** *su novio.*

[14] As a general orientation, I would include here Levin's (1993) classes 18 (*contact by impact*), 20 (*verbs of contact*, *touch verbs*) as well as 41.1 (*verbs of Caring for the Whole Body*), all defined for English.

PN kiss.PST.3SG ø/DOM POSS.3SG boyfriend

'Pepe kissed his boyfriend.'

+> his boyfriend was sentient of it (perception + cognitive state)

c. *Pepe besó **a** su novio pero éste no lo percibió/sabía porque ya estaba dormido.*

'Pepe kissed his boyfriend but he did not perceive/know it because he was already asleep.'

Thus, it can be assumed that once a contact verb of this class selects for an animate object, unless explicitly cancelled, a role-semantic shift from class 4 (*unspecified for agentivity*) to class 3 (*potential agentivity*) takes place.

The second sub-configuration captures verbs that have a subject bearing potent proto-agent properties and an object without any proto-properties. This configuration can be found with controlled perception verbs such as Spanish *mirar* 'to look at', *contemplar* 'to contemplate' or *buscar* 'to look for'. The DOM preference of these verbs is not consistent, for *buscar* for instance it typically correlates with the specificity of the object referent (Leonetti 2004: 80).

The third sub-configuration is one where the subject bears non-potent proto-agent properties and the object is not assigned any proto-properties and thus concerns predicates of restricted or minimal transitivity. A core class are SE-psych verbs, which at least in Spanish, generally also take DOM with human indefinite NPs as illustrated in (84).

(84) **SE-psych verbs [Spanish]**

{*odiar/ admirar/ despreciar/ amar/ aborrecer*} ****ø/a** una persona*

'to hate/ admire/ despise/ love/ detest…a person' (Leonetti 2004: 99)

Moreover, uncontrolled perception verbs like *ver* 'to see', *conocer* 'know' and *encontrar* 'find' can be subsumed under this configuration. Both types of verb classes mentioned can be also attested with DOM if a language has not grammaticalized the phenomenon yet, as in Italian (*conoscere* 'to know', *vedere* 'to see', cf. Berretta 1991: 222) and French (*aimer* 'to love', *voir* 'to see', cf. Fagard/Mardale 2014: 158). With regard to Spanish, there is one further remark to be made. As has been already pointed out in section 4.3.1, verbs like *conocer* 'to know' and *encontrar* 'to find', besides their general meaning, also have a more specific meaning, which involves a symmetrical argument relation between subject and object, as shown in (85).

(85) **Knowledge/perception verbs with a second symmetrical reading [Spanish]**

a. *conocer* ø DO 'to identify, to know' vs.

conocer a DO 'to get acquainted with, to enter in contact with'

b. *encontrar* ø DO 'to find' vs.

encontrar a DO 'to encounter, to meet' (Delbecque 2002: 93f.)

While the general meaning (rather) occurs without DOM, the specific meaning only arises in combination with the *a*-marker.[15] García García (2005) explains such cases within an approach on DOM and informativeness in Spanish. He argues that in case a speaker uses the alternate without *a*, a *quantity and quality based implicature* (qq-implicature) conveys that there is a more informative meaning available that is not meant. Note that the more informative meaning also requires more semantic truth-conditions: in the case of *conocer a* and *encontrar a*, the object also exhibits the proto-agent property of *sentience*. Importantly, also here a shift from class 4 (*unspecified for agentivity*) to 1 (*entailed agentivity*) on the agentivity scale takes place.

A further class covered by the third configuration are verbs with direct objects that do not have argument status. The fact that DOM is ruled out in these cases confirms the generalization of thematic distinctness: if there is only *one* argument, namely the subject, maximal distinctness is given and there is no need to distinguish it from another participant (García García 2014: 207). This is the case for verbs with cognate objects (86a), verbs with property arguments (86b) as well as light verb constructions (86c).

(86) **Direct objects without argument status [Spanish]**

a. **Verbs with cognate objects:**

vivir ***ø/*a*** *una vida escandalosa*, *morir* ***ø/*a*** *una muerte dolorosa*

'live a scandalous life, die a painful death'

b. **Verbs with property arguments:**

Ana es ***ø/*a*** *profesora. / La película dura* ***ø/*a*** *dos horas.*

'Ana is a teacher.' / 'The film lasts two hours.'

c. **Light verb constructions:**

tomar ***ø/*a*** *una decisión, prestar* ***ø/*a*** *ayuda*

'to make a decision, provide help' (García García 2014: 202-213)

In conclusion, two points can be made with regard to verb classes with an object *unspecified for agentivity*. First, verbs of this class are less attested with DOM than with the three higher ranked classes. This is also coherent taking into account the thematic distinctness of subject and object, which is higher when the object does not compete with the subject since it does not bear any type of proto-agent properties. Second, the reviewed verb classes allow for a number of meaning shifts or pragmatic enrichments including a shift of the role-semantic configuration. Importantly, these shifts make the verb enter in a higher-ranked class of the agentivity scale: i.e. contact

[15] A similar observation is made by Delille (1970: 63-65) for Old Portuguese.

verbs with an animate object would enter into class 3 (*potential agentivity*), reversible readings of highly transitive predicates and symmetrical readings of knowledge/perception verbs would enter into class 1 (*entailed agentivity*).[16]

4.4 Working hypothesis

The newly defined role-semantic dimension, namely the communicated agentivity for O, as displayed by the agentivity scale in (87a), serves as the basis to formulate my working hypothesis on the impact of agentivity on DOM in Romance. Crucially, as has been pointed out throughout this chapter, the communicated agentivity for O is related to the agent-patient asymmetry between S and O (87b) as another role-semantic dimension.

(87) **Role-semantic dimensions:**

a. **Communicated agentivity for O:**

entailed agentivity > presupposed agentivity > potential agentivity > unspecified for agentivity

b. **Agent-patient asymmetry (or thematic distinctness) between S and O:**

S is *less* agentive than O > S is as agentive as O > S is *more* agentive than O

As for (87b), it is stated by the generalization of thematic distinctness (García García 2007: 71, 2014: 145) that DOM occurs when we find a deviation from the prototypical clear-cut agent-patient asymmetry of a sentence. This deviation could lie in the fact that (i) subject and object bear the same number of proto-agent properties or that (ii) the object outranks the subject in terms of agentivity. In addition to that, I assume a further role-semantic dimension that affects DOM, namely the communicated agentivity for the object (87a). In this way, a further type of non-prototypical role-semantic configuration can be captured: the overlap of proto-patient and (*entailed*, *presupposed* or *conversationally implicated*) proto-agent properties for the object.

As for the impact of the communicated agentivity for the object on DOM in Romance languages, I derive the following working hypothesis:

(88) **Working hypothesis: Communicated agentivity for the object and DOM in Romance languages**

a. The *more* explicit agentivity information of the DO is communicated by the verbal predicate, the *more* likely DOM is to occur.

[16] This confirms the cross-linguistic observation that meaning shifts occur with verbs that are *less profiled* for one argument in the causal chain (Levin/Rappaport Hovav 2005: 214-216). This explanation from event structure can be transferred to the role-semantic dimension: Once a verb (i) does not entail *any* proto-properties to its object in its main reading or (ii) entails proto-patient properties but no change of state (thus has no complex, but a simple event structure), it may generally allow for more profiling options.

→ DOM is more likely to be found with verbs classes that *entail* the object's agentivity (e.g. (reversible-)symmetrical predicates and OE-psych verbs) than with verbs of the classes placed more to the right on the scale.

b. The *less* explicit agentivity information of the DO is communicated by the verbal predicate, the *less* likely DOM is to occur.

→ DOM is less likely to be found with verbs classes whose object is *unspecified for agentivity* (e.g. SE-psych verbs and perception verbs) than with verbs of the classes placed more to the left on the scale.

It is furthermore assumed that we deal insofar with an implicational scale, that I suppose any DOM occurrence with a verb class *x* on the scale to imply the possibility of DOM with any verb class to the left (*x-1*).[17]

In the following, I will prove the presented hypothesis in a systematic way for two Romance languages, namely Sicilian and Catalan.

4.5 Summary

This chapter started from the observation in 4.1 that Blume's (2000: 200) licensing condition for marked valencies and García García's (2007: 71, 2014: 145) generalization of thematic distinctness for DOM with inanimate objects in Spanish make similar predictions. Both morphosyntactic phenomena are licensed under a similar role-semantic configuration of two co-arguments, namely the *non-asymmetric* distribution of proto-roles. Moreover, both approaches suggest that the object's agentivity can be communicated in different ways: it can be entailed, presupposed or only potential. In 4.2, I presented a range of semantic diagnostics that can serve to identify the type of communicated agentivity for a given verb. It has been shown that the diagnostics allow for a differentiation of four role-semantic classes of the object's agentivity: *entailed agentivity*, *presupposed agentivity*, *potential (conversationally implicated) agentivity* and *unspecified for agentivity*. I have further argued that these four types can be arranged on a scale that ranks them from the strongest form of lexical encoding, i.e. verbal entailment of agentivity, to the absence of any relevant lexically encoded information, i.e. *unspecified for agentivity*. Cross-classifying the agentivity scale with the agent-patient asymmetry allows for a fine-grained distinction of role-semantic configurations that, roughly speaking, range from non-prototypical transitive scenarios to more prototypical ones. Section 4.3 then applies the four-point scale, cross-classified with thematic distinctness, to DOM in Romance. It is suggested that the four agentivity levels serve to systematize DOM-sensitive verb classes across Romance languages. On this basis, the working hypothesis developed in 4.4 states that the more explicitly the direct object's agentivity is communicated

[17] From the working hypothesis in (88) a further (diachronic) implication can be derived which concerns the grammaticalization of DOM in Romance: Languages can be assumed to grammaticalize DOM along the agentivity scale, i.e. the first verbs grammaticalizing the marker are those entailing the object's agentivity (these can also be predicated to be the first verbs to lose the *differentiality* of the marker in a later stage of grammaticalization). I will come back to these implications of the more general working hypothesis in sections 6.1.3 and 6.2.3.

by the verbal semantics, the more likely DOM is to appear. This goes along with the prediction that in grammaticalization, it is verbs to the left of the scale (*entailed agentivity*) that take the marker first while verbs at the right of the scale (*unspecified for agentivity*) are the last to grammaticalize the marker. If these predictions are confirmed by empirical tendencies of DOM in Romance needs to be tested in a systematic way.

5. DOM and communicated agentivity in Sicilian and Catalan

In this chapter, I will test the working hypothesis elaborated in section 4.4 for two Romance languages, Sicilian and Catalan. Since DOM has not been grammaticalized to the same extent in both languages as it is in Spanish, possible effects of agentivity are expected to be more easily detectable. According to the agentivity scale deduced in the previous chapter, four verb classes have been defined for each language. These verb classes have been tested in an acceptability judgement task, which was carried out for one variety of Western Sicilian and two subvarieties of Central Catalan. Acceptability judgement tasks have been proven to be a particular suitable methodological tool in recent research on the phenomenon of DOM, being conducted for different Spanish varieties (cf. Montrul 2013, 2014; Montrul/Bhatt/Girju 2015; Caro Reina/García García/von Heusinger 2021). A main advantage of gradient acceptability judgements for investigating DOM is the fact that it offers a direct comparison of two competing forms for each sentence, that is one with DOM and one without DOM.

5.1 Acceptability judgement study on Sicilian

In the following, an acceptability judgement study conducted on a Sicilian variety will be presented. Beforehand, it must be emphasized that although I speak of Sicilian as a (regional) *language*, its classification among Romance languages is a matter of debate. Lacking standardization and official status, Sicilian can by no means be classified as an *ausbau* language in the sense of Kloss (1967, 1987). Rather, it consists of a "closely related set of linguistic varieties on a number of levels" (Pountain 2016: 638). Yet there is disagreement about the question whether Sicilian meets the definition of an *abstand* language relative to Italian, that is whether it is sufficiently structurally distinguishable from Italian to be referred to as a proper language rather than a dialect of the former. Bossong (2008: 28), on the basis of the possible four combinations of *ausbau* and *abstand* for an idiom, provides the four labels *language* [+ ausbau, + abstand], *dialect* [- ausbau, - abstand], *culture dialect* [+ ausbau, - abstand] and *abstand language* [- ausbau, + abstand]. Sicilian, in his classification of Romance languages based on these criteria, is only a *dialect* lacking both *ausbau* and *abstand* to Italian. That this categorization is debatable can be seen, for instance, by the platform *Ethnologu*e (SIL International), which describes Sicilian as "distinct enough from standard Italian […] to be considered a separate language" (https://www.ethnologue.com/language/scn, last accessed 30 December 2020). It provides the following list of Sicilian varieties:

(1) Varieties of Sicilian (according to *Ethnologue*):

- Western Sicilian (Central-Western Agrigentino, Palermo, Trapani),
- Central Metafonetica,
- Southeast Metafonetica,
- Eastern Nonmetafonetica,
- Messinese,

- Isole Eolie,
- Pantesco,
- Southern Calabro

As for the sociolinguistic situation in Sicily, some brief notes are in order. Nowadays, nearly all speakers of a Sicilian variety are also native speakers of Italian, they hence know Standard (written) Italian with its regional differences. There is general agreement on the view that the situation of Sicilian and Italian in everyday life is not of *diglossia* in its traditional sense (Alfonzetti 1998: 181f., cf. references therein): While a diglossic situation is characterized by a functional differentiation between a low and a high variety, with the low variety being used in everyday communication and the high variety being limited to a set of more formal situations, the situation in Sicily shows a large overlap of domains in which both varieties are used. Hence, Italian is used just as Sicilian in informal communicative situations, for instance in family or friendship domains. This large overlap of domains manifests itself in a widespread use of code-switching and code-mixing of both varieties. In a characterization of this pattern, Alfonzetti (1998: 182) refers to the distinction of code-switching types made by Myers-Scotton (1993) and identifies "the prevailing pattern in Sicily as 'overall switching as the unmarked choice', i.e. a type of code-switching which does not necessarily correlate with a negotiation of interpersonal relationships or with other gross changes in the situation".

As outlined in section 2.2.2, the phenomenon of DOM is attested across Sicilian varieties. For the acceptability judgement study conducted, the Alcamo variety of Western Sicilian was chosen. The following section explains this choice and gives a brief overview of the DOM pattern of Alcamo Sicilian. In particular, morphosyntactic characteristics that must be considered for the investigation of DOM are pointed out.

5.1.1 Variety of Alcamo (province of Trapani, Western Sicilian)

The Alcamo variety can be geographically limited to the city of Alcamo with its 45,000 inhabitants, located in the Province of Trapani. The choice of this variety is primarily due to experiences from a previous pilot study carried out in 2017 for Palermitan Sicilian, another variety of Western Sicilian, to gain first insights on the impact of verbal semantics and nominal parameters on DOM. First, the city of Palermo (670,000 inhabitants) displays a lot of variation inside the city itself. This variation concerns phonological factors in particular, which makes it difficult to create target sentences for an acceptability judgement task that are generally acceptable for participants from different parts of the city. In a smaller city, by contrast, we can expect to find a more "homogeneous" variety in the sense that it is not split up in several internal subvarieties. Second, Palermitan Sicilian shows a morphosyntactic feature that makes it particularly difficult to detect and test DOM, namely the fact that the definite article consists of a single vowel: the reduced forms *u* (masculine singular, full form *lu*) and *a* (feminine singular, full form *la*), respectively. In this case, the definite article frequently fusions with the *a*-marker. Hence,

in the case of the presence of a definite article, the difference between DOM and no-DOM is only very subtle in Palermitan Sicilian. It is especially difficult to identify when it combines with the feminine article, which is also *a*.[1] In contrast, in the variety of Alcamo, DOM does not fuse with the definite article being more readily detectable with NPs bearing the masculine and feminine definite article. Table 1 illustrates the differences of both varieties regarding the forms of the definite article and its combination with the *a*-marker:

	M.SG	**+ DOM**	**F.SG**	**+ DOM**	**M.F.PL**	**+ DOM**
Palermo	u	a + u > o[2]	a + a > a	a	i	a + i > e
Alcamo	lu	a + lu	la	a + la	li	a + li

Table 1. Forms of the definite article in the dialects of Palermo and Alcamo (Western Sicilian)

Written as well as oral sources of the Sicilian of Alcamo are rare. There is, however, a compilation of traditional chants entitled *Alcamo canta* (Rispoli 1996), collected between 1970 and 1974 in the city of Alcamo, from which a few examples are provided in order to get initial insights about the use of DOM. Example (2) confirms one expectation arising from the nominal factors for DOM in Sicilian presented in section 2.1: while DOM is absent with the inanimate objects *'sti cosi* 'these things', *la robba* 'the stuff' and *li casi* 'the houses', it occurs with the singular kinship term *so' figghia* 'his daughter'; in both cases the verb *vuliri* 'to want' is used.

(2) *Jé cci rissi: "Nun vogghiu 'sti cosi*

1SG CL.3SG.DAT say.PST.1SG NEG want.PRS.1SG these things

*vogghiu **a** so' figghia, la robba e li casi".*

want.PRS.1SG DOM his daughter the stuff and the houses

'I said to him: "I don't want these things, I want his daughter, the stuff and the houses."'

(Rispoli 1996: 23; emphasis: S.M.)

[1] Yet, it is mentioned by Guardiano (2000: 22) that there is a considerable phonetic difference between the differential object marker *a* (in her terminology "preposition"), the definite article *a* and the fusion of both *a*-marker and definite article: "La pronuncia delle tre forme e tuttavia assai diversa: ognuna di esse e distinguibile dalle altre. La preposizione semplice ha una pronuncia indipendente, seguita da un breve stacco prima del sostantivo che la segue; la pronuncia dell'articolo determinativo non comporta nessuno stacco dalla parola che segue, ma e rapida; infine la preposizione articolata prevede una pronuncia continua ma una durata decisamente maggiore rispetto all' articolo." ("The pronunciation of the three forms is nevertheless very different: each of them is distinguishable from the others. The simple preposition has an independent pronunciation, followed by a brief detachment before the noun that follows it; the pronunciation of the definite article does not involve any detachment from the following word, but is rapid; finally, the articulated preposition provides a continuous pronunciation but a much longer duration than the article"; translation: S.M.).

[2] Sometimes in written Sicilian the fusion of the masculine definite article and the *a*-marker is expressed by *ô* and, accordingly, by *ê* for the feminine definite article.

The following example is of interest because it consists of an enumeration of different types of object NPs with the verb *ajutari* 'to help'. As can be seen, DOM occurs with the 1st and 2nd person personal pronouns *tia* 'you' and *mia* 'me', with the definite human NP *lu nostru cumpagnu* 'our friend' as well as with universal quantifier *qualunqui pirsuna sia* 'whichever person'.

(3) *ni lu nomu di Gisù Giuseppi e Maria*

in the name of Jesus Joseph and Maria

ch' avi a ajutari ***a*** *tia e* ***a*** *mia*

who have.PRS.3SG to help.INF DOM 2SG and DOM 1SG

a *lu nostrum cumpagnu*

DOM the POSS.1PL friend

e ***a*** *qualunqui pirsuna sia.*

and DOM any person being

'In the name of Jesus, Joseph and Maria, who have to help you and me and our friend and any person being.'

(Rispoli 1996: 47; emphasis: S.M.)

Interestingly, in a similar sentence with the same verb several pages further in the chants, DOM is used with the personal pronouns *mia* 'me' and *tia* 'you' and with the human NP *tutti li pirsuni* 'all people' marked by the universal quantifier *tutti* 'all' and the definite article, but lacks with the definite human NP *lu nostru cumpagnu* 'our friend':

(4) *ni lu nomi di Gisù, S. Anna, Jachinu e Maria*

in the name of Jesus S. Anna Joachim and Maria

chi a 'stu munnu e a 'ddu munnu

who in this world and in that world

santi pirsuni fôru

holy persons be.PST.3PL

e hannu a ajutari ***a*** *mia e* ***a*** *tia*

and have.PRS.3PL to help.INF DOM 1SG and DOM 2SG

e lu nostru cumpagnu e ***a*** *tutti li pirsuni.*

and the POSS.1PL friend and DOM all the persons

‘In the name of Jesus, S. Anna, Joachim and Maria, who have been saints in this world and in that world, and have to help you and me and our friend and all the people.’

(Rispoli 1996: 50; emphasis: S.M.)

As the next example illustrates, the use of DOM is not stable with proper names either. In one and the same sentence, the deity names *santu Vituzzu* and *santa Criscenza* lack the marker while it is present with *santu Roccu*. This is particularly interesting since in a hierarchy of proper name classes, deity names are ranked highest and are hence expected to be the first class to grammaticalize the marker after personal pronouns (Caro Reina 2020: 238f.).

(5) *E chiamu santu Vituzzu*

and call.PRS.1SG saint Vituzzu

chi é a lu Capu

who be.PRS.3SG at the Cape

*e **a** santu Roccu buccadoru*

and DOM saint Roccu know-it-all

e santa Criscenza so’ soru

and saint Criscenza POSS.3SG sister

‘And I call Saint Vituzzu, who is at the Cape, and Saint Roccu know-it-all and his sister Saint Criscenza.’

(Rispoli 1996: 48; emphasis: S.M.)

In summary, although no systematic analysis of all relevant examples of *Alcamo canta* was made, the presented examples suggest two things: first, they confirm that at least as traditional chants are concerned, DOM seems to be a common phenomenon of the variety. Second, given that DOM is subject to variation, they let us assume that at least with definite human NPs and deity names, the marker has rather optional than obligatory status. Note, however, that also text type-specific characteristics, such as the rhythm or metric of a chant, could have influenced the decision to insert DOM or not.

In order to get more insights into the cut-off point of DOM on the animacy and referentiality scale, which is necessary for the later formulation of appropriate test items, I conducted simple grammaticality judgements from a male 52-year-old speaker of Alcamo Sicilian. As the sentences in (6) illustrate, the only categories in which DOM is considered compulsory are personal pronouns (6a) and personal names (6b).[3] With singular kinship terms (6c) and human definite NPs (6d), DOM

[3] The other subclasses of proper names, such as deity names, animal names, place names, distinguished in Caro Reina (2020) have not been tested.

is preferred while it is optional with human indefinite NPs (6e). As expected, inanimate NPs usually do not allow the *a*-marker (6f).

(6) a. *Peppinu talìa* ****ø/a*** *mia/ tia/ idda /iddu.* [pronoun]
PN watch.PRS.3SG ø/DOM 1SG/ 2SG/ 3SG
'Peppino looks at me/you/her/him.'

b. *Peppinu talìa* ****ø/a*** *Maria.* [personal name]
PN watch.PRS.3SG ø/DOM PN
'Peppino looks at Maria.'

c. *Peppinu talìa* ***??ø/a*** *so' soru.* [singular kinship term]
PN watch.PRS.3SG ø/DOM POSS.3SG sister
'Peppino looks at his sister.'

d. *Peppinu talìa* ***?ø/a*** *la fimmina.* [human definite NP]
PN watch.PRS.3SG ø/DOM the woman
'Peppino looks at the woman.'

e. *Peppinu talìa* ***ø/a*** *'na fimmina.* [human indefinite NP]
PN watch.PRS.3SG ø/DOM a woman
'Peppino looks at a woman.'

f. *Peppinu talìa* ***ø/*a*** *la casa.* [inanimate definite NP]
PN watch.PRS.3SG ø/DOM the house
'Peppino looks at the house.'

These grammaticality judgements broadly correspond to what is reported in the literature on the cut-off point of DOM in Sicilian (cf. section 2.1): while the marker is compulsory with pronouns and personal names, it is preferred but not obligatory with definite human NPs and fully optional with indefinite human NPs. These insights are of importance for the parameter of referentiality, which, along with agentivity, will be implemented as a variable in the acceptability judgement test. In the following, the hypotheses of the study are formulated.

5.1.2 Hypotheses

Based on the assumed impact of the degree of communicated agentivity for the object on DOM and on what is generally reported in the literature on the impact of the referentiality (or definiteness) scale on DOM, three main hypotheses (H1-H3) have been formulated. For each hypothesis, testable predictions are made.

The acceptability judgement study operates with three independent variables:

- Agentivity (four values in accordance with the agentivity scale: *entailed agentivity*, *presupposed agentivity*, *potential agentivity*, *unspecified for agentivity*)
- Referentiality (binary values: *definite NP* vs. *indefinite NP*)
- DOM (binary values: *DOM* vs. *no-DOM*)

The dependent variable is the *acceptability rating*. This might seem irritating at first since the working hypothesis (section 4.4) postulates that the communicated agentivity for the direct object has an impact on DOM. Accordingly, DOM should be modelled as the dependent and not as an independent variable. Note, however, that we are dealing with an acceptability judgement task. Such types of data can never make statements about the actual use of DOM. Rather, the detailed insights on the perceived grammaticality of a linguistic form only allow for indirect conclusions on the impact on agentivity on DOM (or no-DOM) by interaction effects of both variables. In corpus studies or production tests, by contrast, DOM can be modelled as dependent variable.

Hypothesis 1 (H1) concerns the interaction between agentivity and DOM. Predictions 1a (P1a) and 1b (P1b) have been formulated based on the agentivity scale put forward in the working hypothesis in section 4.4. Note that in this case, ">" means 'is more acceptable/rated better than'.

(7) **H1 (Agentivity & DOM) and predictions for the acceptability judgement task:**

H1a. The more explicit agentivity information of the DO is communicated by the verbal predicate, the more acceptable DOM is.

H1b. The less explicit agentivity information of the DO is communicated by the verbal predicate, the more acceptable no-DOM is.

P1a: In the DOM condition, we predict the following scale of acceptability for the four verb classes ranging from entailed agentivity for the direct object to unspecified for the object's agentivity:

entailed agentivity & DOM > presupposed agentivity & DOM > potential agentivity & DOM > unspecified for agentivity & DOM

P1b: In the no-DOM condition, we predict the following scale of acceptability for the four verb classes ranging from unspecified for the direct object's agentivity to entailed agentivity for the object:

unspecified for agentivity & no-DOM > potential agentivity & no-DOM > presupposed agentivity & no-DOM > entailed agentivity & no-DOM

Hypothesis 2 (H2) relates to the interaction between DOM and referentiality. It is premised on the cross-linguistic generalization, also validated for Romance languages, that DOM expands along the referentiality (or definiteness) scale (cf. e.g. Bossong 1991; Aissen 2003; von Heusinger 2008). In a simplified way, it is thus predicted that DOM is more acceptable with definite NPs than with indefinite NPs

(P2a) and, correspondingly, that the absence of DOM is more acceptable with indefinite NPs than with definite NPs (P2b).

(8) **H2 (Referentiality & DOM) and predictions for the acceptability judgement task:**

H2a. DOM is more acceptable with definite than with indefinite NPs.

H2b. No-DOM is more acceptable with indefinite than with definite NPs.

P2a: In the DOM condition, I predict the following scale of acceptability for definite and indefinite NPs:

definite NP & DOM > indefinite NP & DOM

P2b: In the no-DOM condition, I predict the following scale of acceptability for definite and indefinite NPs:

indefinite NP & no-DOM > definite NP & no-DOM

Hypothesis 3 (H3) concerns the interaction between all three variables, i.e. agentivity, referentiality and DOM. Here I expect that the agentivity scale interacts in the same way with both definite and indefinite object NPs. Predictions 3a (P3a) and 3c (P3c) state the same cline of agentivity classes for definite and indefinite NPs in the DOM condition. In the no-DOM condition, the reverse cline of acceptability of the four agentivity classes is predicted both for definite NPs (P3b) and indefinite NPs (P3d).

(9) **H3 (Agentivity & Referentiality & DOM) and predictions for the acceptability judgement task:**

H3a. The more explicit agentivity information of the DO is communicated by the verbal predicate, the more acceptable DOM is. It is assumed that the scale of acceptability for the four verb classes does not differ for definite and indefinite object NPs.

H3b. The less explicit agentivity information of the DO is communicated by the verbal predicate, the more acceptable no-DOM is. It is assumed that the scale of acceptability for the four verb classes does not differ for definite and indefinite object NPs.

P3a: In the definite NP & DOM condition, I predict the following scale of acceptability for the four verb classes ranging from entailed agentivity for the direct object to unspecified for the object's agentivity:

entailed agentivity & def. NP & DOM > presupposed agentivity & def. NP & DOM > potential agentivity & def. NP DOM > unspecified for agentivity & def. NP & DOM

P3b: In the definite NP & no-DOM condition, I predict the following scale of acceptability for the four verb classes ranging from unspecified for the direct object's agentivity to entailed agentivity for the object:

unspecified for agentivity & def. NP & no-DOM > potential agentivity & def. NP & no-DOM > presupposed agentivity & def. NP & no-DOM > entailed agentivity & def. NP & no-DOM

P3c: In the indefinite NP & DOM condition, I predict the following scale of acceptability for the four verb classes ranging from entailed agentivity for the direct object to unspecified for the object's agentivity:

entailed agentivity & indef. NP & DOM > presupposed agentivity & indef. NP & DOM > potential agentivity & indef. NP & DOM > unspecified for agentivity & indef. NP & DOM

P3d: In the indefinite NP & no-DOM condition, I predict the following scale of acceptability for the four verb classes ranging from unspecified for the direct object's agentivity to entailed agentivity for the object:

unspecified for agentivity & indef. NP & no-DOM > potential agentivity & indef. NP & no-DOM > presupposed agentivity & indef. NP & no-DOM > entailed agentivity & indef. NP & no-DOM

5.1.3 Study design

The creation of the test items required the consideration of further variety-specific characteristics, which will be described in more detail in the following. In order to test the variables agentivity and referentiality, both verb type and object NP type have been manipulated. As for the verb type, for each of the four agentivity classes five lexicalizations have been specified (cf. (10)). For the type of object NP, the lexicalizations for definite and indefinite NPs are listed in (11). In order to keep the animacy of the object constant, only lexicalizations that point to human referents have been defined.

(10) **Verb type: agentivity encoding of the direct object**

a. entailed agentivity for the object:

5 OE-psych verbs (*affènniri* 'to offend', *cunsulari* 'to console', *fari scantari* 'to frighten', *fari siddiari* 'to provoke', *nuciri* 'to bother')

b. presupposed agentivity for the object:

2 interaction verbs (*aiutari* 'to help' and *sentiri* 'to hear'), 3 highly transitive verbs (*acchiappari* 'to catch', *arristari* 'to stop' and *sarvari* 'to save')

c. potential agentivity for the object:

5 highly transitive verbs (*ammazzari* 'to kill', *curari* 'to cure', *firiri* 'to injure', *pigghiari a paroli* 'to insult', *struppiari* 'to cripple')

d. object unspecified for agentivity:

2 visual perception verbs (*taliari* 'to watch' and *viriri* 'to see'), 1 other perception verb (*truvari* 'to find'), 1 SE-psych verb (*'un putiri viriri* 'to hate')[4] and 1 knowledge verb (*canusciri* 'to know')

[4] Actually more SE-psych verbs should have been tested but it turned out that most of them are not used transitively: their object is either (i) an indirect object or (ii) a prepositional object:
(i) e.g. 'to adore/estimate/like' – *piaciri a*
(ii) e.g. 'to fear' – *scantarisi di*, 'to respect' – *purtari rispittu a*

(11) **Object NP type: human definite or human indefinite NP**

a. definite NP (SG, marked by the definite article):

l'atturi 'the actor', *lu so' capu* 'his/her boss', *lu latru* 'the thief' (2x), *l'omu* 'the man', *lu parrinu* 'the priest', *lu so' patruni di casa* 'his/her landlord', *lu picciut-teddu* 'the little boy', *lu prisidenti* 'the president', *lu prufissuri* 'the teacher', *lu so' prufissuri* 'his/her teacher', *lu sbirru* 'the policeman' (2x), *lu sinnacu* 'the deputy', *la vecchia* 'the old woman', *lu vecchiu* 'the old man', *la so' vicina* 'his/her neighbour', *lu so' vicinu* 'his/her neighbour', *la so' zita* 'his/her girl-friend', *lu so' zitu* 'his/her boyfriend'

b. indefinite NP (SG, marked by the indefinite article):

n'amicu 'a friend' (3x), *un avvucatu* 'a judge', *'na cumpagna* 'a friend', *un cumpagnu* 'a friend', *'na fimmina* 'a woman', *un malatu* 'an ill man', *un maliutu* 'a criminal' (2x), *un 'mmriacuni* 'a drunk', *un ospiti* 'a guest', *'na picciridda* 'a girl', *un picciriddu* 'a child' (2x), *un politicu* 'a politician', *un scarparu* 'a shoe-maker', *un vicinu* 'a neighbour', *un vinnituri* 'a market seller'

In order to make sure that the verbs in (10) are syntactically transitive, it has been proved with a native speaker that the object of each verb can be substituted by the direct object clitic *lu* (M.SG) or *la* (F.SG) (vs. the indirect object clitic *ci*), as exemplified for the verb *aiutari* 'to help' in (12):[5]

(12) a. *Pietru aiuta* ***ø/a*** *la fimmina.*

PN help.PRS.3SG ø/DOM the woman

'Pietro helps the woman.'

b. *Pietru* ***la/*ci*** *aiuta.*

PN CL.F.3SG.ACC/CL.3SG.DAT help.PRS.3SG

'Pietro helps her.'

The lexicalizations of the direct object NP in (11) were balanced between *sortal* and *relational* nouns as differentiated by Löbner (1985: 292f.): while sortal nouns such as *lu latru* 'the thief' or *'na picciridda* 'a little girl' classify objects and are logically one-place predicates, relational nouns such as *la so' vicina* 'his/her neighbour' or *un amicu* 'a friend' are conceptually related to another object and can thus be described as two-place predicates. For the sake of sounding more natural, most of the definite relational nouns are accompanied by a possessive.

The subject of the target sentences always consisted of a human personal name. Here common Sicilian first names and abbreviations were used in order to make the sentences sound most natural. A further argument for constantly using animate, more specifically human, NPs in subject position is to prevent possible effects arising from non-canonical animacy asymmetries, i.e. with an inanimate subject and an

[5] A passivization test, also typically used to test the direct object status, is not available since Sicilian does not have a periphrastic passive.

animate direct object. Furthermore, two variety-specific properties had to be taken into account: First, the selected tense was the past perfect (*trapassato prossimo*) (13a), since for verbs ending in *-ari* in the variety of Alcamo both the present tense (*presente*) (13b) and the preterite (*passato remoto*) (13c) have third-person singular endings in *-a* which make DOM phonetically hard to detect.

(13) a. Past perfect (*trapassato prossimo*):

Dinu 'nmenzu la strata avìa arristatu ***ø/a***

PN in the middle of the street have.IPRF.3SG stop.PTCP ø/DOM

lu latru.

the thief

'Dino (had) stopped the thief in the middle of the street.'

b. Present tense (*presente*):

Dinu 'nmenzu la strata arrista ***ø/a*** *lu latru.*

PN in the middle of the street stop.PRS.3SG ø/DOM the thief

'Dino stops the thief in the middle of the street.'

c. Preterite (*passato remoto*):

Dinu 'nmenzu la strata arristà ***ø/a*** *lu latru.*

PN in the middle of the street stop.PST.3SG ø/DOM the thief

'Dino stopped the thief in the middle of the street.'

An exception was made for the verbs *canusciri* 'to know' and *'un putiri viriri* 'to hate': For the former, this was due to the fact that a past perfect form of *canusciri* would rather lead to the meaning 'get to know'. Note that in this case, we would deal with a symmetrical predicate in which the object participant is entailed *sentience*. Hence, for guaranteeing the desired meaning with a direct object unspecified for agentivity, the present tense was chosen. For the latter, a past perfect reading was rejected due to the stative nature of the predicate *'un putiri viriri* 'to hate'. This is why the present tense was also chosen here. Note that since for both exceptions we do not deal with verbs ending in *-ari*, a possible detection of DOM was guaranteed. Second, in order to make a sentence sound more natural, an adverbial phrase was added. This adverbial phrase was either of local, temporal, manner or causal type. Testing different sentence structures with a native speaker, it turned out that the adverbial phrase was most natural in second position, as in the sentences in (13). [6] Note that the direct object always appeared in clause-internal postverbal position

[6] The preferred clause structure can thus be described as follows: [[$_{\text{PRECORE SLOT}}$ *Dinu*] [$_{\text{PERIPHERAL}}$ *'nmenzu la strata*] [$_{\text{CORE}}$ *avia arristatu ø/a lu latru*]]. The subject NP is placed in the precore slot followed by a peripheral element, the adverbial phrase, which is in turn followed by the core clause.

controlling for the fact that left- or right-dislocated direct objects are more readily *a*-marked in Sicilian than those occurring in canonical position (cf. e.g. Iemmolo 2010a).

Crossing the two variables agentivity (*entailed agentivity*, *presupposed agentivity*, *potential agentivity* and *unspecified for agentivity*) and referentiality (*definite object NP* vs. *indefinite object NP*) with the variable DOM (*DOM* vs. *no-DOM*), I came up with the following 16 (8 × 2) critical conditions as each item had two variants, one with and one without DOM.

(14) **Critical conditions**

a. entailed agentivity + definite object NP + [± DOM]

Vicenzu cu li so' 'nciurii avìa fattu siddiari ***ø/a***

PN with his insults have.IPRF.3SG provoke.PTCP ø/DOM

lu sbirru.

the policeman

'Vicenzo (had) provoked the policeman with his insults.'

b. entailed agentivity + indefinite object NP[7] + [± DOM]

Sarvaturi cu lu so' tintu fari avìa fattu siddiari

PN with his bad behaviour have.IPRF.3SG provoke.PTCP

ø/a *'na fimmina.*

At first sight, this syntactic structure seems to rule against what Bentley (2008: 280) states, namely that in Sicilian, "the position which follows the Pre-Core Slot can be filled by an argument but not easily by an adverbial". However, this seems to vary depending on the information-structural function of the argument occupying the precore slot: while Bentley refers to constrastive focal elements (which in the above example would be *?DINU 'nmenzu la strata avia arristatu ø/a lu latru(, nun Marcu)* 'It was Dino who had stopped the thief in the middle of the street, not Marco'), in my sentences, the subject can be classified as a topic. In this case, it can be easily followed by an adverbial. Based on Cruschina (2012), I suppose that the information given by the adverbial could be interpreted as new information focus and is hence preferred in fronted position. Note that Sicilian, like Sardinian, but unlike most other Romance languages, typically makes use of focus fronting not only for contrastive focus but also to express new information focus. As shown by Cruschina (2012: 62), the expression of new information focus in the left periphery of the sentence "can involve constituents of any category: direct objects, indirect objects, nominal and adjectival predicates" as well as adjuncts (Cruschina 2012: 62, fn. 30), depending on the relevance of the given information in context. However, given the absence of a proper context in the presented target sentences, it is difficult to exactly determine its information-structural value. Note that the offered explanation would furthermore be compatible with the syntactic analysis provided by Alexiadou/Anagnostopoulou (1998: 501ff.) stating that preverbal subjects in null subject languages, like Sicilian, involve left-dislocation of the subject.

[7] Please note that we can assume *referential specificity* of the indefinite object NP. Given the lack of contextual cues, though, the *epistemic specificity* of the NPs in question cannot be determined (for an overview of different types of specificity, cf. von Heusinger 2011).

ø/DOM a woman

'Sarvaturi (had) provoked a woman with his bad behaviour.'

c. presupposed agentivity + definite object NP + [± DOM]

Marcu fora la putia avìa acchiappatu ***ø/a*** *lu latru.*

PN outside the store have.IPRF.3SG catch.PTCP ø/DOM the thief

'Marco (had) caught the thief outside the store.'

d. presupposed agentivity + indefinite object NP + [± DOM]

Pinu a lu mercatu avìa acchiappatu ***ø/a*** *un maliutu.*

PN on the market have.IPRF.3SG catch.PTCP ø/DOM a criminal

'Pino (had) caught a criminal on the market.'

e. potential agentivity + definite object NP + [± DOM]

Maria in du' simani avìa curatu ***ø/a***

PN in two weeks have.IPRF.3SG cure.PTCP ø/DOM

lu picciutteddu.

the little boy

'Maria (had) cured the little boy in two weeks.'

f. potential agentivity + indefinite object NP + [± DOM]

Titidda cu la so' midicina avìa curatu ***ø/a*** *un malatu.*

PN with her medicine have.IPRF.3SG cure.PTCP ø/DOM a patient

'Titidda (had) cured a patient with her medicine.'

g. unspecified for agentivity + definite object NP + [± DOM]

Brizzita a la staziuni avìa vistu ***ø/a*** *lu prisidenti.*

PN at the train station have.IPRF.3SG see.PTCP ø/DOM the president

'Brizzita saw (/ had seen) the president at the train station.'

h. unspecified for agentivity + indefinite object NP + [± DOM]

Pitrina a lu mercatu avìa vistu ***ø/a*** *un politicu.*

PN on the market have.IPRF.3SG see.PTCP ø/DOM a politician

'Pitrina saw (/ had seen) a politician on the market.'

Five sets of the 16 (8 × 2) critical conditions were constructed, with an outcome of 80 sentences in total. Each of the 20 verbs was used with two different object

lexicalizations. In order to prevent participants from detecting the phenomenon tested, 20 fillers were integrated. The fillers can be divided into 10 grammatical and 10 ungrammatical sentences. The former consisted of transitive verbs (with inanimate objects) or intransitive verbs. The latter included incorrect prepositions, clitics and choice of tense. An example of both filler types is given in (15):

(15) a. grammatical filler [transitive verb + inanimate DO]

Teresa 'un àvi ancora capùtu la questiòni.

PN NEG have.PRS.3SG still understand.PTCP the problem

'Teresa still does not understand the problem.'

b. ungrammatical filler [intransitive verb + incorrect clitic]

A Ninu 'un **lu piaci lu gattu nivuru.*[8]

IO PN NEG CL.M.SG.ACC like.PRS.3SG the cat black

'Nino does not like black cats.'

While (15a) displays a grammatical transitive sentence with an inanimate direct object, the sentence in (15b) is ungrammatical due to the insertion of the direct object clitic *lu* (instead of the indirect object clitic *ci*) with the intransitive verb *piaciri* 'to like'.

All of the material was pseudo-randomized and distributed according to Latin square to two lists A and B. Hence, each list consisted of 40 test items (20 with DOM and 20 without DOM) and 20 fillers.[9] In order to not sound repetitive to the participants, I varied all lexicalizations except for the verb. One shortcoming of this is that possible effects might not be unambiguously attributed to verbal semantics but might theoretically be due to other elements of the sentence (personal name, adverbial, lexicalization of object NP), too. The set of test items and fillers was recorded with a 78-year-old male native speaker from Alcamo at a normal speaking pace.[10] The judgements are given on a five-point scale (ranging from 1 – *not acceptable* to 5 – *totally acceptable*).

5.1.4 Participants

25 native speakers of the Sicilian variety of Alcamo volunteered for the questionnaire study. Participants were recruited in the city of Alcamo in November and December 2018. All of them were born in Alcamo and lived there at the time of data

[8] The correct version of the sentence contains the indirect object clitic *ci*:
*A Ninu 'un **ci** piaci lu gattu nivuru.*
IO PN NEG CL.SG.DAT like.PRS.3SG the cat black
'Nino does not like black cats.'

[9] For a list of the test items, cf. Appendix 1.1. The audio versions A and B of the questionnaire are available here: https://uni-koeln.sciebo.de/s/kmQyHw8DHzUsg0t.

[10] Because of sound problems, a few sentences had to be re-recorded with a 60-year-old male native speaker from Alcamo.

collection. As a pretest revealed, although the test items are presented orally, a basic level of literacy is needed to properly fulfil the judgement task. This is why literacy was adopted as a further criterion for participation.[11] Participants were informed about the academic purpose of the study and gave their verbal consent. As for the research topic, they were told that the study aimed at investigating the variety of Alcamo in general. After the analysis of fillers, one participant of version A and one participant of version B were eliminated since their ratings of the filler sentences sharply deviated from what was expected (more than 37.5 % , i.e. 6/20 "false" answers).[12] The remaining 23 questionnaires that were used for the analysis are comprised of 11 for version A and 12 for version B. The following table offers an overview of the collected metadata:

Question	**Number of participants (total n=23)**
Age	Mean age 57.2 years, range 20-80 years
Sex	Female (5), male (18)
Level of study	Middle school (7), High school (12), University education (4)
Place of birth (municipality/province)	Alcamo (Trapani) (23)
Place of residence (municipality/province)	Alcamo (Trapani) (23)
Mother tongue	Italian (22), Sicilian (1)
Second language(s)	English (6), German (1), Italian (1), Sicilian (3)
Language(s) spoken at home	Italian (5), Sicilian (3), Both (15)
Preferred language: Italian or Sicilian (or both)	Italian (7), Sicilian (5), Both (11)

Table 2. Participants of the study

The high proportion of male participants (18 male vs. 5 female) was chosen deliberately since a number of sociolinguistic studies revealed that men use lower status variants more than women, while the latter tend to use higher status, more prestigious variants (cf. e.g. Labov 1966, 1972; Trudgill 1974; Macaulay 1978).[13]

[11] The illiteracy rate for people in Sicily older than 65 years has been 10.4 % in 2001. Compared to Northern Italian regions, this number is relatively high, cf. e.g. 2.3 % in Toscana and 1.2 % in Lombardia. Among Southern Italian regions, however, we still find higher percentages for the regions of Puglia (11.4 %), Campania (12.0 %), Basilicata (16.9 %) and Calabria (19.7 %) (cf. Vincenzo D'Aprile 2005).

[12] The filler analysis consisted of two steps: In a first step, I excluded all fillers that did not work the way they should. For the grammatical ("good") fillers, I decided to exclude all fillers that got a mean value under 4.0 (3 fillers). For the ungrammatical ("bad") fillers, I excluded one filler whose mean value was over 3.0. The mean rating for good fillers was 4.344 and 2.896 for bad fillers. Please note that these mean values show a considerable range of variation, something that normally is not predicted for fully grammatical and ungrammatical filler sentences since they should normally lead participants to use the outer edges of the scale. I had to take into account this variation in a second step: I counted the ratings for each filler of each participant. Due to the great range of variation, a rating was considered "false" if a good filler was rated by 1 or if a bad filler was rated by 5. As a consequence, two participants were excluded who gave more than 6 (37.5 %) "wrong" answers.

[13] The mentioned studies consisted of production experiments conducted for varieties of American or British English mainly focussing on phonetical variants. Another finding was that lower social classes (especially working class) used more non-standard, i.e. less prestigious forms than middle

The relatively high mean age (57.2 years) is due to the preferred recruitment of older speakers of Sicilian. This is because of the experience that older speakers are more dialect preserving while younger speakers tend to be highly influenced by modern globalized media.

5.1.5 Procedure and data analysis

The two lists A and B of the questionnaire were presented to the participants in oral format, as MP3 audios. Participants were randomly assigned to one of the two lists. They were orally instructed to imagine themselves in an informal conversation with people they usually talk to in Sicilian (e.g. family or friends) of which the presented sentences should be imagined to be part of. Participants were asked to rate each sentence on a 5-point scale ranging from 1 (*not acceptable*) to 5 (*totally acceptable*). They were explicitly told that there were no right or wrong answers but that they should rather follow their intuition. The ratings were given on a piece of paper. Each participant was tested individually, always taking into account the individual tempo. Sentences were presented one at a time, with a short pause after each sentence. Participants were allowed to listen to a sentence more than one time if necessary. After having finished the questionnaire, they were asked to fill in a form with their sociodemographic data (cf. Appendix 1.3). Participation took 15 minutes on average. All data were collected anonymously. The meta language between the investigator and the subjects was Italian.

For the statistical analysis, I decided to use a mixed-effects regression model. Such a model has the advantage that it includes not only fixed effects but also random effects. The consideration of random effects allows for the control of speaker-specific and item-specific idiosyncrasies to prevent that the regression coefficients are distorted by these effects (Gries 2015: 97). Since I conducted 40 enquiries for one and the same participant and several enquiries for one and the same item across participants (11 in version A, and 12 in version B), it was probable that some of the collected data points were related. In order to consider these idiosyncrasies in the model, I integrated (i) participants and (ii) items as random effects. The fixed effects referred to the three independent variables, namely (i) agentivity, (ii) referentiality and (iii) DOM. In order to test the assumed four-level ranking of agentivity classes formulated in P1a-b and P3a-d, I performed comparisons of the mean values in each condition as indicated in (16):

(16) **Performed comparisons of agentivity levels in statistical modelling:**

---

and upper social classes. Though the preference for male Sicilian speakers is justified with the results of these studies, one must be well aware of two points: first, the presented view is actually much more complex (for a comprehensive overview cf. Kotthoff/Nübling 2018: 243–271). Besides a simple correlation of sex (and social class) with linguistic variants, other factors such as language-related requirements of working environments have to be taken into account (cf. Kotthoff/Nübling 2018: 250f.). Second, from a methodological point of view, in contrast to the presented studies, I did not conduct a production study but an acceptability judgement task. Yet it cannot be said that people who tend to *produce* lower status variants also *prefer* the very same forms in an acceptability judgement task (cf. e.g. Adli 2015).

a. entailed agentivity vs. presupposed agentivity

b. entailed agentivity vs. potential agentivity

c. entailed agentivity vs. unspecified for agentivity

d. presupposed agentivity vs. potential agentivity

e. presupposed agentivity vs. unspecified for agentivity

f. potential agentivity vs. unspecified for agentivity

The data analysis was carried out in R (version 3.6.1: R Core Team 2017) with the packages lme4[14] (Bates et al. 2015) and lmerTest[15] (Kuznetsova et al. 2017). Effects were considered significant if $p < .05$.

5.1.6 Results

This section summarizes the results according to H1, H2 and H3 formulated in section 5.1.2.[16] All hypotheses concerned the interaction between at least two of the three independent variables.

Main effects of agentivity, referentiality and DOM were not ascertained.

5.1.6.1 Hypothesis 1 (Agentivity & DOM)

H1 which concerns the interaction between DOM and agentivity with the predictions P1a and P1b repeated in (17), was not confirmed.

(17) **H1: Agentivity & DOM**

P1a.
entailed agentivity & DOM > presupposed agentivity & DOM > potential agentivity & DOM > unspecified for agentivity & DOM

P1b.
unspecified for agentivity & no-DOM > potential agentivity & no-DOM > presupposed agentivity & no-DOM > entailed agentivity & no-DOM

Figure 1 illustrates the distribution of acceptability values in condition 1a and 1b. As displayed by the boxplots, the dispersion of the data is relatively low with 50 to 75% of the answers being clustered between 4 and 5 on the scale. It is solely for *potential agentivity & DOM* and for *unspecified for agentivity & no-DOM* that the range of acceptability values covers the entire scale with 25% of the answers being between 1 and 3. **Figure 2** and **Table 3** show the estimated marginal means (EMM) of acceptability values with 95% confidence intervals in condition 1a (*Agentivity & DOM*) and condition 1b (*Agentivity & no-DOM*). As for condition 1a, we can see that the acceptability of *entailed agentivity & DOM* and *presupposed agentivity &*

[14] Linear Mixed-Effects Models using 'Eigen' and S4.
[15] Tests in Linear Mixed Effects Models.
[16] I am very grateful to Maximilian Hörl for his support with the statistical analysis of my data.

DOM was nearly equal and only slightly better than for *potential agentivity & DOM*. The agentivity level, which was expected to rank lowest, namely *unspecified for agentivity & DOM*, gets the highest acceptability score. However, in the model none of the numeric differences between the four means reaches statistical significance (cf. **Table 4**). In condition 1b (*Agentivity & no-DOM*), against P1b, the numeric results exactly mirror the opposite of what was expected: *entailed agentivity & no-DOM* is accepted best and *unspecified for agentivity & no-DOM* gets the worst ranking (cf. **Figure 2** and **Table 3**). Yet, also here the model revealed that none of the measured effects was statistically significant (cf. **Table 5**).

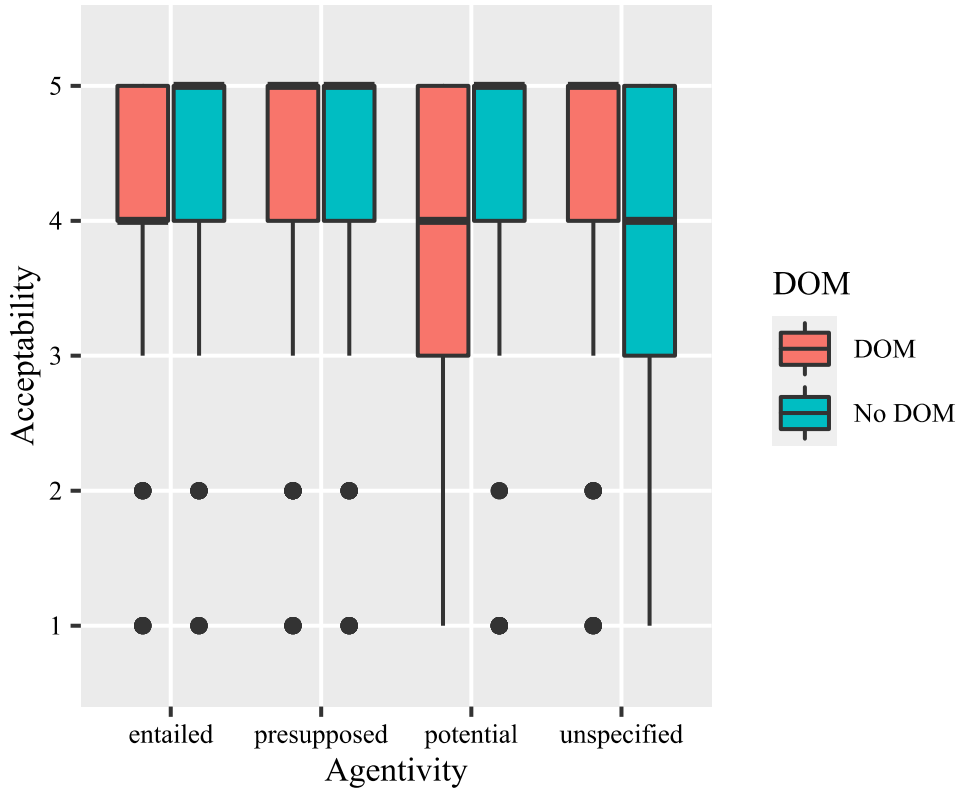


Figure 1. Boxplots of acceptability values in condition 1a (Agentivity & DOM) and 1b (Agentivity & no-DOM) with median values and outliers

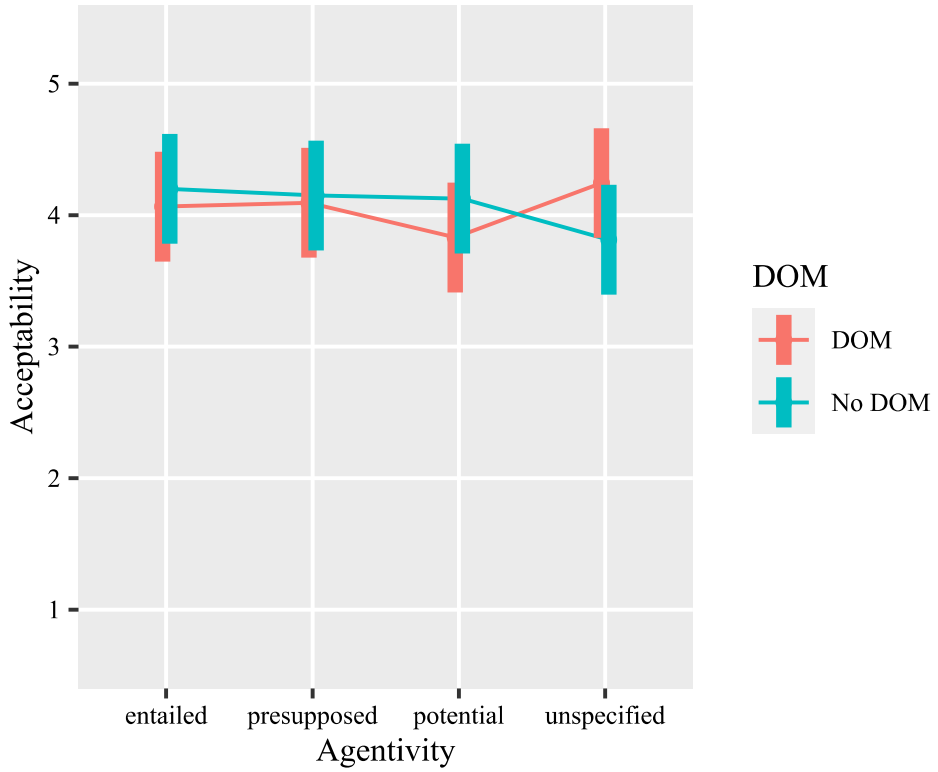


Figure 2. Estimated marginal means (EMM) of acceptability values in condition 1a (Agentivity & DOM) and 1b (Agentivity & no-DOM), confidence level used: 0.95

Agentivity	**EMM**	**SE**	**df**	**lower CL**	**upper CL**
(i) entailed					
1a (DOM)	4.07	0.21	82.6	3.65	4.48
1b (no-DOM)	4.20	0.21	82.6	3.78	4.62
(ii) presupposed					
1a (DOM)	4.09	0.21	82.6	3.68	4.51
1b (no-DOM)	4.15	0.21	82.6	3.73	4.57
(iii) potential					
1a (DOM)	3.83	0.21	82.6	3.41	4.25
1b (no-DOM)	4.13	0.21	82.6	3.71	4.54
(iv) unspecified					
1a (DOM)	4.24	0.21	82.6	3.83	4.66
1b (no-DOM)	3.81	0.21	82.6	3.40	4.23

Table 3. Estimated marginal means (EMM) of acceptability values in condition 1a (Agentivity & DOM) and 1b (Agentivity & no-DOM), degrees-of-freedom method: kenward-roger, confidence level used: 0.95

Condition 1a (Agentivity & DOM)

Fixed effects	Estimate	Standard error	*t*-value	Pr(>|t|)
Diff1: entailed vs. presupp. agentivity & DOM	0.02896	0.24227	0.120	0.905
Diff2: entailed vs. potential agentivity & DOM	-0.23483	0.24227	-0.969	0.336
Diff3: entailed vs. unspec. for agentivity & DOM	0.17929	0.24227	0.740	0.462
Diff4: presupp. vs. potential agentivity & DOM	-0.26379	0.24227	-1.089	0.280
Diff5: presupp. vs. unspec. for agentivity & DOM	0.15033	0.24227	0.621	0.537
Diff6: potential vs. unspec. for agentivity & DOM	0.4141	0.2423	1.709	0.0917

Table 4. Parameter estimates for the fixed effects in condition 1a (Agentivity & DOM)

Condition 1b (Agentivity & no-DOM)

Fixed effects	Estimate	Standard error	*t*-value	Pr(>|t|)
Diff1: entailed vs. presupp. agentivity & no-DOM	-0.05064	0.24227	-0.209	0.835
Diff2: entailed vs. potential agentivity & no-DOM	-0.07427	0.24227	-0.307	0.760
Diff3: entailed vs. unspec. for agentivity & no-DOM	-0.38720	0.24227	-1.598	0.114
Diff4: presupp. vs. potential agentivity & no-DOM	-0.02363	0.24227	-0.098	0.923
Diff5: presupp. vs. unspec. for agentivity & no-DOM	-0.33656	0.24227	-1.389	0.169
Diff6: potential vs. unspec. for agentivity & no-DOM	-0.31293	0.24227	-1.292	0.2006

Table 5. Parameter estimates for the fixed effects in condition 1b (Agentivity & no-DOM)

5.1.6.2 Hypothesis 2 (Referentiality & DOM)

H2, which proved the interaction between referentiality and DOM with the predictions repeated in (18), was not confirmed either.

(18) **H2: Referentiality & DOM**

P2a. definite NP & DOM > indefinite NP & DOM

P2b. indefinite NP & no-DOM > definite NP & no-DOM

Figure 3 shows the distribution of acceptability values in condition 2a and 2b. As for *definite NP & DOM*, *indefinite NP & DOM* and *indefinite NP & no-DOM*, 75% of the acceptability values are clustered between 4 and 5 on the scale. The remaining 25% are distributed between 3 and 4. A greater range of variation is given for *definite NP & no-DOM* where the boxplot indicates that 25% of the answers are between 1 and 3. As can be seen from **Figure 4** and **Table 6**, *definite NP* and *indefinite NP* show approximately the same EMM in both the DOM (2a) and no-DOM (2b) condition. Since there were only very slight differences between the acceptability values in both conditions, the model revealed no significant effects (cf. **Table 7**).

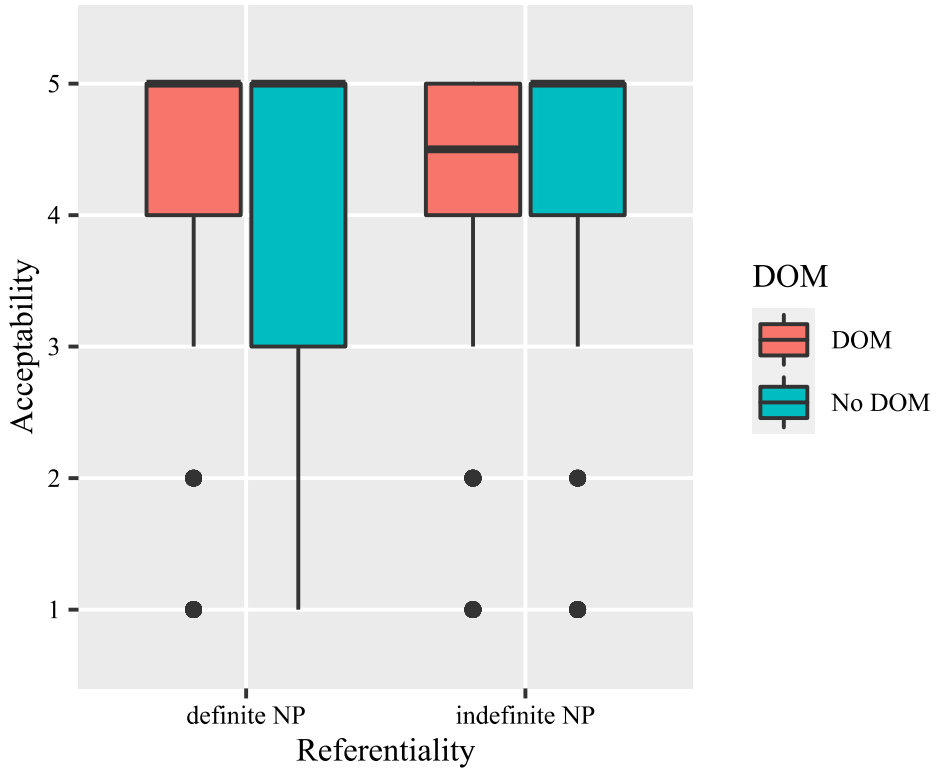


Figure 3. Boxplots of acceptability values in condition 2a (Referentiality & DOM) and 2b (Referentiality & no-DOM) with median values and outliers

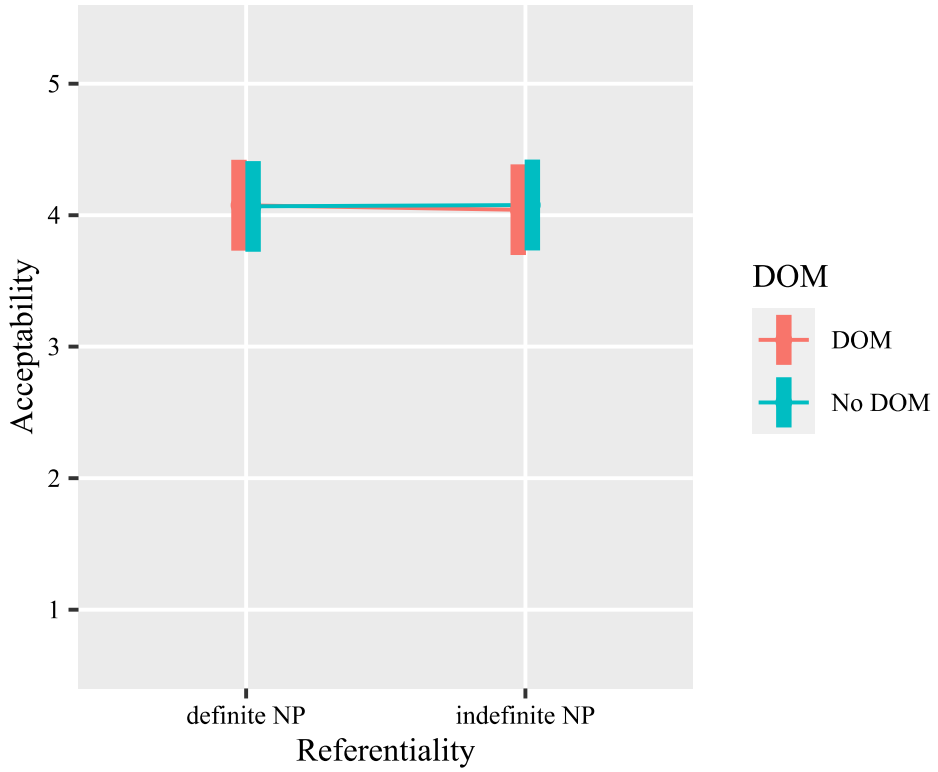


Figure 4. Estimated marginal means (EMM) of acceptability values in condition 2a (Referentiality & DOM) and 2b (Referentiality & no-DOM), confidence level used: 0.95

Referentiality	**EMM**	**SE**	**df**	**lower CL**	**upper CL**
(i) definite NP					
2a (DOM)	4.08	0.173	61.5	3.73	4.42
2b (no-DOM)	4.07	0.173	61.5	3.72	4.41
(ii) indefinite NP					
2a (DOM)	4.04	0.173	61.5	3.70	4.39
2b (no-DOM)	4.08	0.173	61.5	3.73	4.42

Table 6. Estimated marginal means (EMM) of acceptability values in condition 2a (Referentiality & DOM) and 2b (Referentiality & no-DOM), degrees-of-freedom method: kenward-roger, confidence level used: 0.95

Conditions 2a and 2b

	Fixed effects	Estimate	Standard error	*t*-value	Pr(>|t|)
2a	definite NP vs. indefinite NP & DOM	-0.033849	0.173833	-0.195	0.846
2b	definite NP vs. indefinite NP & no-DOM	0.010327	0.173833	0.059	0.953

Table 7. Parameter estimates for the fixed effects in condition 2a (Referentiality & DOM) and 2b (Referentiality & no-DOM)

5.1.6.3 Hypothesis 3 (Agentivity & Referentiality & DOM)

As for H3, the predicted interaction between agentivity, referentiality and DOM, shown again in (19), was partly confirmed, namely for condition 3a. In the other three conditions 3b-3d, the predictions were not confirmed.

(19) **H3: Agentivity & Referentiality & DOM**

P3a.
entailed agentivity & def. NP & DOM > presupposed agentivity & def. NP & DOM > potential agentivity & def. NP & DOM > unspecified for agentivity & def. NP & DOM

P3b.
unspecified for agentivity & def. NP & no-DOM > potential agentivity & def. NP & no-DOM > presupposed agentivity & def. NP & no-DOM > entailed agentivity & def. NP & no-DOM

P3c.
entailed agentivity & indef. NP & DOM > presupposed agentivity & indef. NP & DOM > potential agentivity & indef. NP & DOM > unspecified for agentivity & indef. NP & DOM

P3d.
unspecified for agentivity & indef. NP & no-DOM > potential agentivity & indef. NP & no-DOM > presupposed agentivity & indef. NP & no-DOM > entailed agentivity & indef. NP & no-DOM

Figure 5 and **Figure 7** illustrate the distribution of acceptability values in the conditions 3a+b and 3c+d, respectively. As a very general observation, it can be stated that the dispersion of the data is considerably higher in condition 3c and 3d (*Agentivity & Indefinite NP & DOM*) than in condition 3a and 3b (*Agentivity & Definite NP & DOM*). The EMM of the four verb classes with 95% confidence intervals are illustrated in **Figure 6** (condition 3a and 3b) and **Figure 8** (condition 3c and 3d) and listed in **Table 8**. In condition 3a, an effect of *entailed agentivity & definite NP & DOM* and *presupposed agentivity & definite NP & DOM* is found: both classes are clearly better rated than *potential agentivity & definite NP & DOM*. The two differences are statistically significant (cf. **Table 10**). No similar effect was found in the reverse no-DOM condition 3b. As for condition 3c, *presupposed agentivity*

& indefinite NP & DOM and *potential agentivity & indefinite NP & DOM* get approximately equal ratings, *entailed agentivity & indefinite NP & DOM* is rated slightly worse, while *unspecified for agentivity & indefinite NP & DOM* obtains the best rating. The reverse tendency is visible in condition 3d, where *entailed agentivity & indefinite NP & no-DOM* is ranked best and *unspecified for agentivity & indefinite NP & no-DOM* is ranked lowest. None of the effects in condition 3c and 3d reaches statistical significance.[17]

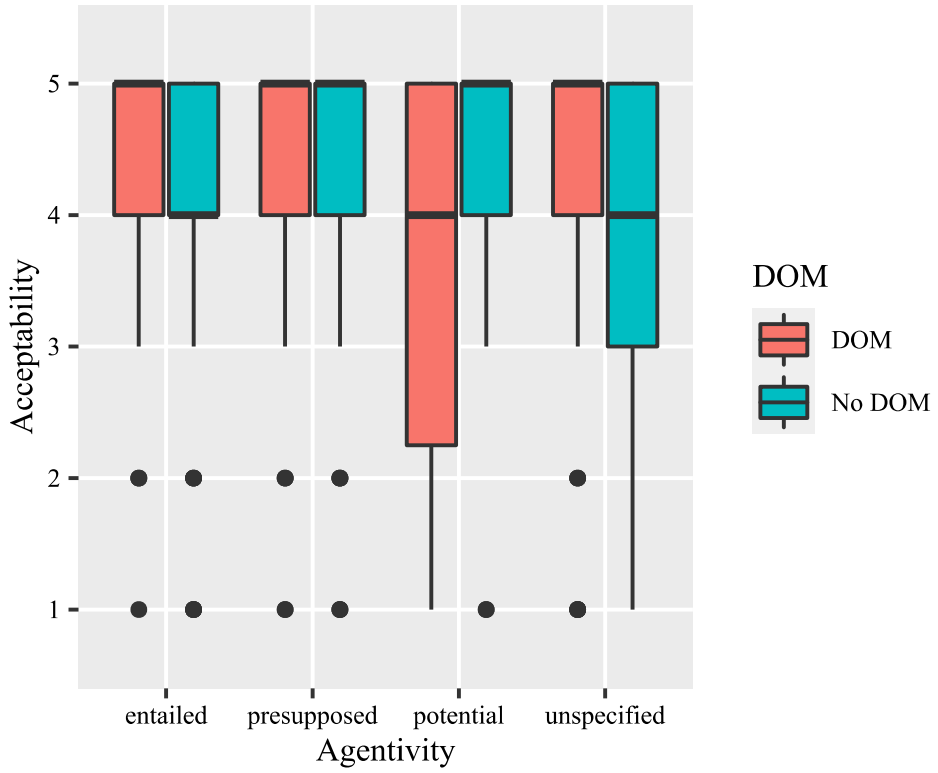


Figure 5. Boxplots of acceptability values in condition 3a (Agentivity & definite NP & DOM) and 3b (Agentivity & definite NP & no-DOM) with median values and outliers

[17] For the mean acceptability values of the individual verbs in the conditions 3a-3d, cf. Appendix 1.4.

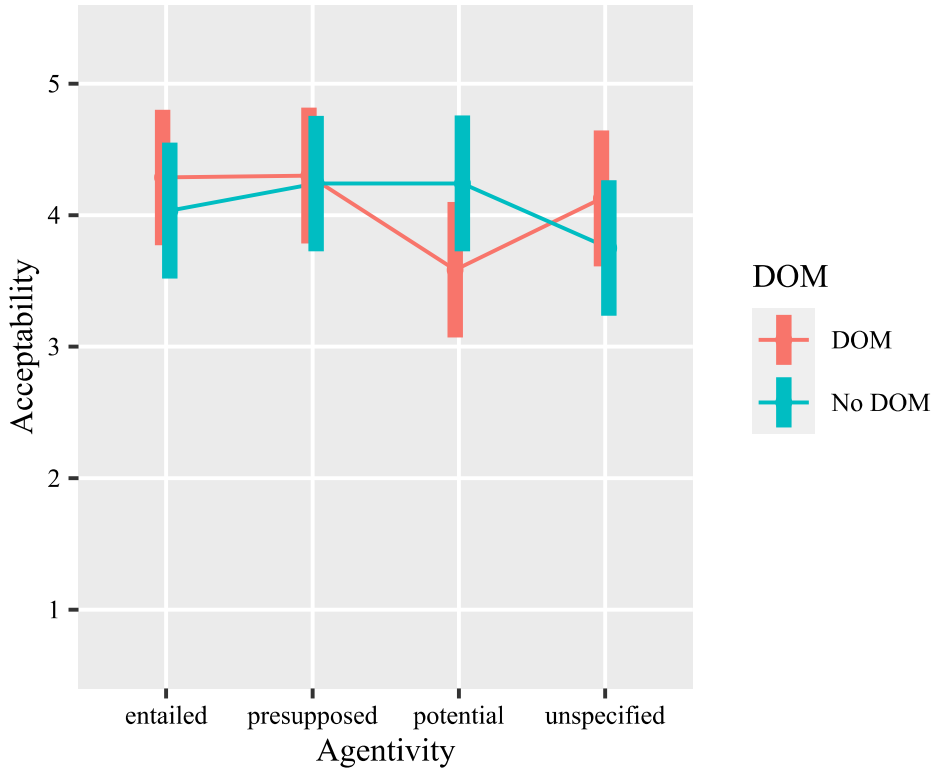


Figure 6. Estimated marginal means (EMM) of acceptability values in condition 3a (Agentivity & definite NP & DOM) and 3b (Agentivity & definite NP & no-DOM), confidence level used: 0.95

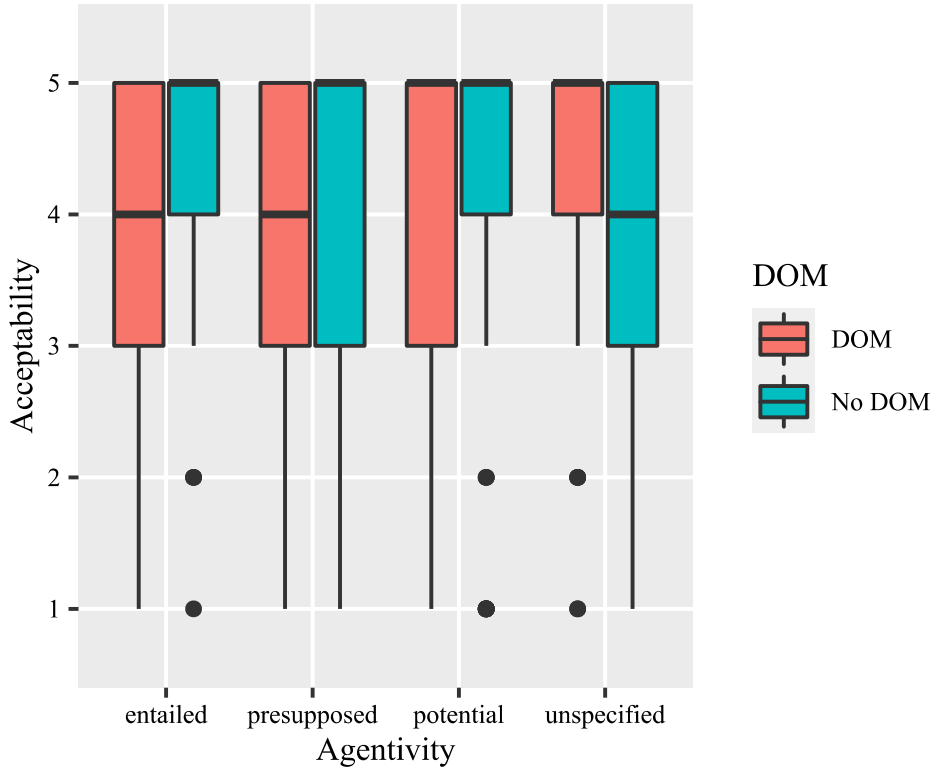


Figure 7. Boxplots of acceptability values in condition 3c (Agentivity & indefinite NP & DOM) and 3d (Agentivity & indefinite NP & no-DOM) with median values and outliers

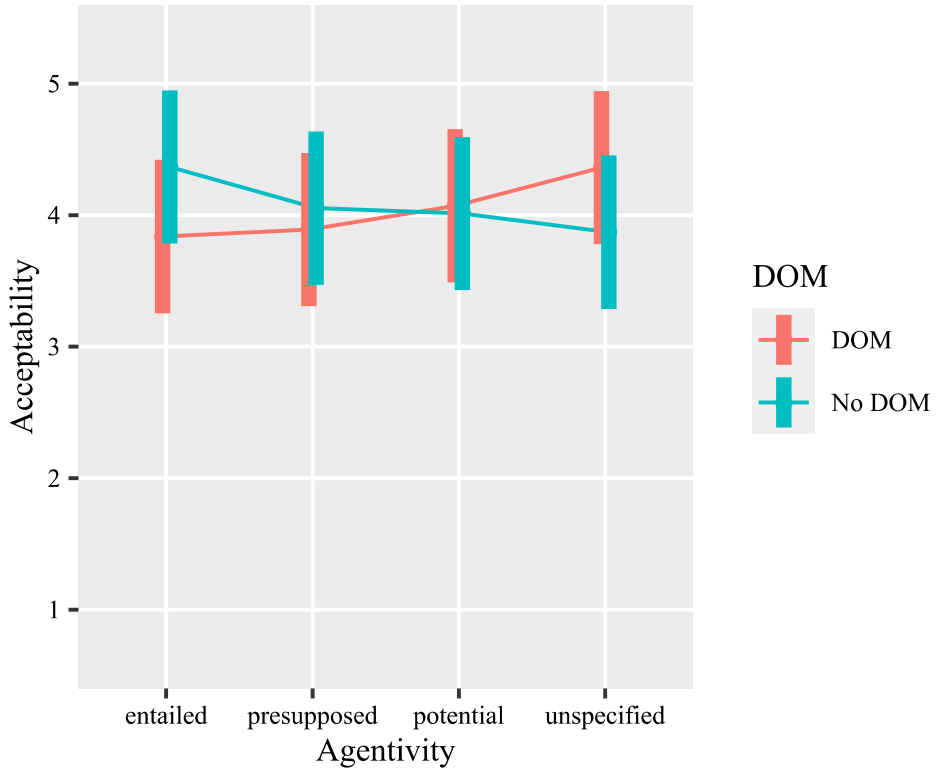


Figure 8. Estimated marginal means (EMM) of acceptability values in condition 3c (Agentivity & indefinite NP & DOM) and 3d (Agentivity & indefinite NP & no-DOM), confidence level used: 0.95

Agentivity & definite NP	**EMM**	**SE**	**df**	**lower CL**	**upper CL**
(i) entailed & definite NP					
3a (DOM)	4.29	0.256	43.9	3.77	4.80
3b (no-DOM)	4.03	0.256	44.2	3.52	4.55
(ii) presupposed & definite NP					
3a (DOM)	4.30	0.256	44.2	3.78	4.82
3b (no-DOM)	4.24	0.256	43.9	3.73	4.76
(iii) potential & definite NP					
3a (DOM)	3.58	0.256	43.9	3.07	4.10
3b (no-DOM)	4.24	0.256	44.2	3.72	4.76
(iv) unspecified & definite NP					
3a (DOM)	4.13	0.256	44.2	3.61	4.64
3b (no-DOM)	3.75	0.256	43.9	3.24	4.27

Table 8. Estimated marginal means (EMM) of acceptability values in condition 3a (Agentivity & definite NP & DOM) and 3b (Agentivity & definite NP & no-DOM), degrees-of-freedom method: kenward-roger, confidence level used: 0.95

Agentivity & indefinite NP	EMM	SE	df	lower CL	upper CL
(i) entailed & indefinite NP					
3c (DOM)	3.84	0.289	42.4	3.25	4.42
3d (no-DOM)	4.37	0.288	42.1	3.78	4.95
(ii) presupposed & indefinite NP					
3c (DOM)	3.89	0.288	42.1	3.31	4.47
3d (no-DOM)	4.05	0.289	42.4	3.47	4.64
(iii) potential & indefinite NP					
3c (DOM)	4.07	0.289	42.4	3.49	4.65
3d (no-DOM)	4.01	0.288	42.1	3.43	4.59
(iv) unspecified & indefinite NP					
3c (DOM)	4.36	0.288	42.1	3.78	4.94
3d (no-DOM)	3.87	0.289	42.4	3.29	4.45

Table 9. Estimated marginal means (EMM) of acceptability values in condition 3c (Agentivity & indefinite NP & DOM) and 3d (Agentivity & indefinite NP & no-DOM), degrees-of-freedom method: kenward-roger, confidence level used: 0.95

Condition 3a

Fixed effects	Estimate	Standard error	*t*-value	Pr(>|t|)
Diff1: entailed vs. presupp. agentivity & def. NP & DOM	0.01426	0.32264	0.044	0.9650
Diff2: entailed vs. potential agentivity & def. NP & DOM	-0.70235	0.32105	-2.188	**0.0363***
Diff3: entailed vs. unspec. for agentivity & def. NP & DOM	-0.15938	0.32264	-0.494	0.6247
Diff4: presupp. vs. potential agentivity & def. NP & DOM	-0.71660	0.32264	-2.221	**0.0337***
Diff5: presupp. vs. unspec. for agentivity & def. NP & DOM	-0.17364	0.32200	-0.539	0.5935
Diff6: potential vs. unspec. for agentivity & def. NP & DOM	0.5430	0.3226	1.683	0.1023

Table 10. Parameter estimates for the fixed effects in condition 3a (Agentivity & definite NP & DOM)

Condition 3b

Fixed effects	Estimate	Standard error	*t*-value	Pr(>|t|)
Diff1: entailed vs. presupp. agentivity & def. NP & no-DOM	0.2055	0.3226	0.637	0.529
Diff2: entailed vs. potential agentivity & def. NP & no-DOM	0.2067	0.3220	0.642	0.526
Diff3: entailed vs. unspec. for agentivity & def. NP & no-DOM	-0.2845	0.3226	-0.882	0.385
Diff4: presupp. vs. potential agentivity & def. NP & no-DOM	0.001165	0.322643	0.004	0.997
Diff5: presupp. vs. unspec. for agentivity & def. NP & no-DOM	-0.490007	0.321050	-1.526	0.137
Diff6: potential vs. unspec. for agentivity & def. NP & no-DOM	-0.491172	0.322643	-1.522	0.138

Table 11. Parameter estimates for the fixed effects in condition 3b (Agentivity & definite NP & no-DOM)

Condition 3c

Fixed effects	Estimate	Standard error	*t*-value	Pr(>|t|)
Diff1: entailed vs. presupp. agentivity & indef. NP & DOM	0.05275	0.37220	0.142	0.888
Diff2: entailed vs. potential agentivity & indef. NP & DOM	0.23422	0.37143	0.631	0.533
Diff3: entailed vs. unspec. for agentivity & indef. NP & DOM	0.52470	0.37220	1.410	0.169
Diff4: presupp. vs. potential agentivity & indef. NP & DOM	0.18147	0.37220	0.488	0.629
Diff5: presupp. vs. unspec. for agentivity & indef. NP & DOM	0.47195	0.37060	1.274	0.212
Diff6: potential vs. unspec. for agentivity & indef. NP & DOM	0.29049	0.37220	0.780	0.441

Table 12. Parameter estimates for the fixed effects in condition 3c (Agentivity & indefinite NP & DOM)

Condition 3d

Fixed effects	Estimate	Standard error	*t*-value	Pr(>|t|)
Diff1: entailed vs. presupp. agentivity & indef. NP & no-DOM	-0.3130	0.3722	-0.841	0.407
Diff2: entailed vs. potential agentivity & indef. NP & no-DOM	-0.3542	0.3706	-0.956	0.3465
Diff3: entailed vs. unspec. for agentivity & indef. NP & no-DOM	-0.4954	0.3722	-1.331	0.1928
Diff4: presupp. vs. potential agentivity & indef. NP & no-DOM	-0.04125	0.37220	-0.111	0.912
Diff5: presupp. vs. unspec. for agentivity & indef. NP & no-DOM	-0.18241	0.37143	-0.491	0.627
Diff6: potential vs. unspec. for agentivity & indef. NP & no-DOM	-0.14117	0.37220	-0.379	0.707

Table 13. Parameter estimates for the fixed effects in condition 3d (Agentivity & indefinite NP & no-DOM)

To summarize, the results reveal the following pattern of acceptability in Sicilian: H1 (Agentivity & DOM) and H2 (Referentiality & DOM) were not confirmed. For H3 (Agentivity & Referentiality & DOM) the following significant effects in condition 3a (*Agentivity & definite NP & DOM*) were found:

(20) a. entailed agentivity & def. NP & DOM > potential agentivity & def. NP & DOM

b. presupposed agentivity & def. NP & DOM > potential agentivity & def. NP & DOM

As shown in (20a), for *a*-marked definite objects, verbs entailing agentivity for its direct object are ranked better than verbs conveying potential agentivity for the object. A similar effect was found for verbs presupposing agentivity for the object, which are rated significantly better than verbs with a potentially agentive object (cf. (20b)). Yet the attested significant interaction was not confirmed in condition 3b (*Agentivity & definite NP & no-DOM*). Also, for indefinite NPs (conditions 3c and 3d), no such effect could be found. Rather the reverse tendency seems to be true: In condition 3c (*Agentivity & indefinite NP & DOM*) *entailed agentivity* is ranked worst and *unspecified for agentivity* is accepted best, while the other two classes lie inbetween. This is mirrored in condition 3d (*Agentivity & indefinite NP & no-DOM*) where *entailed agentivity* is rated best and *unspecified for agentivity* gets the lowest acceptability ratings. In the following, I will discuss the results obtained.

5.1.7 Discussion

As for H1 (Agentivity & DOM) and H2 (Referentiality & DOM), the predictions were not confirmed. I did not detect an interaction effect of the single independent

variables agentivity and referentiality with DOM, neither in the DOM nor in the no-DOM condition. Concerning H2, it was striking that the four mean values of definite and indefinite NPs in condition 2a (*Referentiality & DOM*) and 2b (*Referentiality & no-DOM*) were almost equal. With regard to the DOM condition, the lack of difference between definite and indefinite NPs corresponds to the results of acceptability judgement tasks conducted for different Spanish varieties: In a study of Montrul/Bhatt/Girju (2015: 582), we see a similar distribution for two different groups of Spanish heritage speakers in the US, a group of Mexican immigrants in the US as well as for two groups of Mexican Spanish native speakers living in Mexico. Likewise, in a recent study of Caro Reina/García García/von Heusinger (2021), similar results were reached for *a*-marked definite and indefinite objects in an acceptability judgement test on Cuban and European Spanish. It can be noted, though, that while in the present study *a*-marked definite and indefinite objects got mean ratings of approximately 4.0 (out of 5.0), in the studies on Spanish participants assigned relatively higher scores in these cases, namely nearly the maximum of the given scales of acceptability (i.e. close to 6.0 out of 6.0, cf. Caro Reina/García García/von Heusinger: 358). Yet, in contrast to the results for Sicilian, in the studies on Spanish mentioned we do see a difference in the no-DOM condition where the absence of DOM is clearly better accepted with indefinite than with definite NPs. In the present study, by contrast, the acceptability ratings for definite NPs and indefinite NPs in condition 2b (*Referentiality & no-DOM*) do not show any difference. A plausible explanation for this finding could be the lack of grammaticalization of DOM in Sicilian: The absence of DOM does not lead to ungrammaticality, neither with human definite nor with human indefinite direct objects. Rather, the findings suggest that Sicilian displays full optionality for DOM with both NP types. This matches what is reported in the literature about the distribution of DOM in Sicilian, namely that while pronouns and personal names are obligatorily marked, with definite NPs (especially those marked by the definite article), we observe more variation (cf. e.g. Iemmolo 2010b, Guardiano 2000).

With regard to H3, the predictions were partially confirmed. In condition 3a (*Agentivity & definite NP & DOM*), two triple interaction effects were found. The verb classes *entailed agentivity* and *presupposed agentivity* are rated significantly better than *potential agentivity*, while *entailed agentivity* and *presupposed agentivity* among themselves do not differ. Both *entailed agentivity* and *presupposed agentivity* do not differ in comparison to *unspecified for agentivity*. Moreover, there was no significant difference between *potential agentivity* and *unspecified for agentivity*. The confirmed ranking of verb classes – in interaction with *definite NP* and *DOM* – is summarized in (21):

(21) obtained scale of acceptability in condition 3a:

entailed agentivity & def. NP & DOM = presupposed agentivity & def. NP & DOM > potential agentivity & def. NP & DOM

The measured interaction effects can be interpreted as follows: If agentivity is encoded in the verbal semantics, as it is for both *entailed agentivity* and *presupposed agentivity*, DOM is more acceptable than in cases where the object is only potentially agentive via implicature, as with *potential agentivity*. At the same time, the direct object must be highly ranked in the definiteness scale, so it must be highly prominent in terms of referentiality. As we have seen for H1, where the interaction between agentivity and DOM was tested independently of referentiality and no effect was measured, this latter factor is indispensable.

Going more into detail, it is important to take a look at the homogeneity of acceptability ratings for the individual items within a given class. In the class *potential agentivity*, there was one item that was rated extremely well relative to the four others in condition 3a, obtaining a mean acceptability score of 4.833:

(22) *Ninetta fora la scola avìa pigghiatu a paroli a*
PN in front of the school have.IPRF.3SG insult.PTCP DOM
lu so' prufissuri.
the POSS.3SG teacher
'Ninetta (had) insulted her teacher in front of the school.'

Unlike the other four verbs of the same class, for *pigghiari a paroli* 'to insult', the object is not physically but only verbally affected. It can be only hypothesized that, since the verb meaning comes very close to the OE-psych verb *affènniri* 'to offend', the object might be interpreted as being implied *sentience* and thus patterns with the class *entailed agentivity*. Yet, with the insertion of the locative adverbial *fora la scola* 'in front of the school' I tried to exclude this latter reading in that the external circumstances of the act of verbal violence and not its manner is stressed.

With regard to the verb class *unspecified for agentivity*, there is also a class-internal outlier in condition 3a: Acceptability ratings for the item containing the verb *canusciri* 'to know' (23) sharply deviate from the other four items in that they are extremely high and nearly reach the highest mean value possible (4.917 out of 5.0). This is at the same time the overall best rated item.

(23) *Stefanu da tantu tempu canusci a lu parrinu.*
PN for a long time know.PRS.3SG DOM the priest
'Stefanu knows the priest for a long time.'

Choosing the verb *canusciri* 'to know', I tried to exclude its second reading 'to get to know' by selecting the present tense in contrast to the past perfect. Furthermore, the temporal adverbial chosen should force a non-symmetrical reading in which only the subject is ascribed the proto-agent property *sentience*. Yet it is possible that participants inferred the object participant to be sentient too, supposing that the

priest in question knows Stefanu as well. In this case, *canusciri* like *pigghiari a paroli* above, should be arguably classified as belonging to the class *entailed agentivity*.

The interaction effect of agentivity, referentiality and DOM shown for condition 3a was not confirmed for the reverse condition 3b (*Agentivity & definite NP & no-DOM*). Numerically, *presupposed agentivity* was ranked best, followed by *potential agentivity*, *entailed agentivity* and *unspecified for agentivity*. In comparison to condition 3a, we can furthermore state that for *potential agentivity*, acceptability was better in the no-DOM than in the DOM condition, while it was better in the DOM than in the no-DOM condition for *entailed agentivity* and *unspecified for agentivity* and quasi equal in both conditions for *presupposed agentivity*. Note that none of the reported differences reaches statistical significance. I assume that, like for condition 2b (*Referentiality & no-DOM*), there were no clear differences in acceptability compared to condition 3a because DOM has not fully grammaticalized in Sicilian. Therefore it did not lose its optionality, not even for verb classes such as *entailed agentivity* and *presupposed agentivity*, which in the reverse condition 3a are rated significantly better with DOM than the verb class *potential agentivity*.

The results for condition 3c and 3d suggest that the impact of agentivity also differs with respect to referentiality. Starting with condition 3c (*Agentivity & indefinite NP & DOM*), it can be stated that acceptability was best with *unspecified for agentivity* and *potential agentivity*, followed by *presupposed agentivity* and worst with *entailed agentivity*. Correspondingly, in condition 3d (*Agentivity & indefinite NP & no-DOM*) the absence of DOM was rated best with *entailed agentivity* and worst with *unspecified for agentivity*. Further support for the opposing behaviour of verb classes depending on the referentiality of the object NP (condition 3a+3b vs. 3c+3d) comes from the behaviour of individual verbs (cf. Appendix 1.4): If one contrasts the mean ratings for DOM vs. no-DOM for definite NPs and indefinite NPs, respectively, the difference depending on referentiality becomes especially striking: for *entailed agentivity*, DOM with definite NPs (condition 3a) is generally better rated than no-DOM (condition 3b), for indefinite NPs it is rather the converse, in that no-DOM (condition 3d) is generally rated better than DOM (condition 3c). In contrast, for *unspecified for agentivity* (with the exception of *canusciri* 'to know'), regarding definite NPs the difference of acceptability between DOM (condition 3a) and no-DOM (condition 3b) is only marginal whereas for indefinite NPs, DOM (condition 3c) is strikingly better rated than no-DOM (condition 3d). For the classes *presupposed agentivity* and *potential agentivity*, individual verbs show a similar behaviour, of which *aiutari* 'to help', *ammazzari* 'to kill' and *curari* 'to cure' can be especially mentioned.

The two class-internal outliers of condition 3a (*pigghiari a paroli* 'to insult' and *canusciri* 'to know'), deviate from the other verbs of their respective class and pattern with *entailed agentivity*, too: both verbs are rated clearly better without DOM (condition 3d) than with DOM (condition 3c).

In summary, we are confronted with the puzzling result that while in condition 3a the acceptability for DOM is highest for *entailed agentivity & definite NP* and *presupposed agentivity & definite NP*, in condition 3c DOM was rated best for *unspecified for agentivity & indefinite NP* and worst for *entailed agentivity & indefinite NP*. Although the latter effect was not significant, it is a strong tendency that is also reflected in the individual verbs. How can we account for the seemingly contradictory behaviour of the verb classes in question?

One plausible interpretation would be that the acceptability of DOM shows different – maybe competing – patterns for definite and indefinite object NPs in Sicilian. As was specified in the study design (5.1.3), the target sentences contain an adverbial phrase. Yet it was not controlled for the type of adverbial, so there were adverbials expressing temporal and local circumstances as well as adverbials of manner and purpose. The following two tables provide an overview of the adverbials used for the individual verbs of *entailed agentivity* (Table 14) and *unspecified for agentivity* (Table 15):

Verb	**adverbial def. NP**	**type of adverbial**	**adverbial indef. NP**	**type of adverbial**
affènniri 'to offend'	*cu lu so' malu lamentu* 'with his evil accusations'	manner	*cu paroli tinti* 'with bad words'	manner
nuciri 'to bother'	*cu li so' vuci* 'with his screaming'	manner	*arre' nni la bibliuteca* 'once again in the library'	location
cunsulari 'to console'	*cu paroli duci* 'with sweet words'	manner	*cu amuri* 'with love'	manner
fari siddiari 'to provoke'	*cu li so' 'nciurii* 'with his insults'	manner	*cu lu so' tintu fari* 'with his bad behaviour'	manner
fari scantari 'to frighten'	*p'addivèrtisi* 'for fun'	purpose	*pi ririri* 'to laugh'	purpose

Table 14. Adverbials for *entailed agentivity*

Verb	**adverbial def. NP**	**type of adverbial**	**adverbial indef. NP**	**type of adverbial**
viriri 'to see'	*a la staziuni* 'at the station'	location	*a lu mercatu* 'on the market'	location
taliari 'to watch'	*tutta priata* 'with admiration'	manner	*cu 'nteressi* 'with interest'	manner
truvari 'to find'	*nni lu varveri* 'at the barber's'	location	*a lu cafè* 'at the bar'	location
'un putiri viriri 'to hate'	*di lu primu mumentu*	temporal	*di lu primu iornu*	temporal

	'from the first moment on'		'from the first day on'	
canusciri 'to know'	*da tantu tempu* 'for a long time'	temporal	*da deci anni* 'for ten years'	temporal

Table 15. Adverbials for *unspecified for agentivity*

As can be seen, with the exception of *nuciri* 'to bother', the type of adverbial was identical for each individual verb across conditions. Concerning OE-psych verbs, it has been shown in the literature that these verbs usually have an implicit causality bias towards the subject (Kehler et al. 2008). Accordingly, a likely continuation of a sentence containing an OE-psych verb is expected to be a cause with the former subject being expressed as the subject of the explanation (e.g. *Peter offended Mary because **he** is a mean guy*). However, it was found that if the implicit causality bias is pre-empted by the insertion of a simple cause in the initial sentence (i.e. in the former case, *Peter offended Mary by his mean behaviour*), continuations giving explanations are clearly less attested (Solstad/Bott 2017: 25). Looking at the data for OE-psych verbs with indefinite NPs, we see that when a cause is given the absence of DOM was rated clearly better than the occurrence of DOM. In the case of *nuciri* 'to bother', where we have an adverbial indicating a location, DOM is rated slightly better than the absence of DOM. In this line of argumentation, the acceptability of DOM with indefinite NPs would at least partly depend on the expected continuation in the given discourse. For the class *unspecified for agentivity* and some verbs of *potential agentivity* (*ammazzari* 'to kill', *curari* 'to cure'), the high ratings for DOM with indefinite NPs (exception *canusciri* 'to know') in comparison to no-DOM would be then accounted for by an expected continuation.

Certainly, such explanation is highly speculative and cannot be proved by this data since I only presented out-of-the-blue sentences. As for future research, two kinds of studies are needed: Firstly, a slightly modified replication of my study should be conducted in which the type of adverbial is kept constant across verb classes. In this way, we would see if the low acceptability of *a*-marked indefinite objects for *entailed agentivity* is upheld or if it was instead a side effect of the inserted causal adverbial. Secondly, and beyond the scope of the present research question, an analysis of discourse data would have to prove if in Sicilian, the insertion of DOM can be also related to the status of the *a*-marked participant in the following discourse.[18]

5.2 Acceptability judgement study on Catalan

The second language for which the acceptability of DOM is tested in this work is Catalan. For Catalan, and especially treating the phenomenon of DOM, it is

[18] Such function has been confirmed for DOM in Romanian by Chiriacescu/von Heusinger (2010, 2011): As they show, *pe*-marked definite and indefinite direct objects are referentially more persistent than unmarked ones and more likely to become topics. A similar analysis has been given by Weissenrieder (1990) for DOM with animate (non-human) objects in Spanish where DOM is optional. Note that the mentioned studies do not take into account the verbal semantics.

necessary to distinguish between different varieties of the language as was specified in section 2.1. While in Standard Catalan, DOM is restricted to strong pronouns only, the marker is more widespread in the regional spoken varieties of Central Catalan, Valencian and Balearic Catalan.[19] Table 16 exemplifies the cut-off points for the varieties mentioned:

	personal pronoun	**proper name**	**definite** human NP	inani-mate NP	**indefinite** human NP	inanimate NP
Standard	a	ø	ø	ø	ø	ø
Central	a	a/ø	a/ø	ø	??a/ø	ø
Valencian	a	a	a	ø	?a/ø	ø
Balearic non-detached	a	ø	ø	ø	ø	ø
detached	a	a	a	a	a/ø	a/ø

Table 16. Cut-off points for DOM in different Catalan varieties

This study is limited to two subvarieties of Central Catalan: the varieties of Barcelona and Girona. The choice of Central Catalan was made due to the expected variation for DOM with definite human NPs: in cases where DOM is optional, other factors, such as role semantics, can be expected to surface more clearly than when DOM shows a higher degree of grammaticalization. The latter is the case for Valencian where the expansion of DOM is comparable to the situation in Modern Spanish (Sancho Cremades 1995: 199f.). Balearic Catalan was not taken into consideration because DOM underlies primarily information structural constraints and therefore does not occur in canonical word order with the exception of strong pronouns (cf. Escandell-Vidal 2009). As the design should be kept as close to the Sicilian study as possible presenting single SVO sentences without context, Balearic Catalan would not have proved a favourable variety for comparison.

5.2.1 Varieties of Barcelona and Girona (Central Catalan)

As mentioned above, for the present study two varieties of Central Catalan were selected, namely those spoken in Barcelona and Girona. Note that I considered the entire provinces of Barcelona and Girona, though expecting that the majority of participants would come from the cities of Barcelona and Girona. Although (nearly) all speakers of Catalan are bilingual and also native speakers of Spanish, one can assume that the influence of Spanish is higher in the metropolis of Barcelona than in Girona (Cortés et al. 2009: 185f.). A statistical survey of the IDESCAT (Institut d'Estadística de Catalunya 2018) shows that the *Àmbit Metropolità*, i.e. the territory of Barcelona, displays the lowest rate of Catalan as habitual language (27.5%) vs.

[19] For the other three varieties of Catalan (Northwestern Catalan, Northern Catalan and Alghero Catalan), the extension of DOM has been less described in the literature. I therefore do not include these varieties in my description above.

54.1% in the *Àmbit de les Comarques Gironines*, i.e. the territory of Girona.[20] For this reason, the phenomenon of DOM can be expected to be more widespread in the former than in the latter variety of Central Catalan (Pineda 2013: 99f.).

Two characteristics of Central Catalan are of importance for the study design: the first is a phonological one, manifested by the pronounciation [əl] of the masculine definite article *el* vs. [ɛl] in Valencian. The second is a morphosyntactic one: in Catalan in general, it is usual to mark personal names with either an onymic marker or the definite article (Caro Reina 2014; Alsina 2016: 376). The type of determination differs according to the variety in question. Table 17 shows the general pattern for the five main varieties:

Catalan variety	**M**	**F**	**M.F (with initial vowel)**
Balearic	*en Pere*	*na Maria*	*n'Andreu*
North-Western Catalan	*el/lo Pere*	*la Maria*	*l'Andreu*
Central Catalan, Roussillon Catalan	*en Pere*	*la Maria*	*l'Andreu*
Valencian, Alghero Catalan	*ø Pere*	*ø Maria*	*ø Andreu*

Table 17. Use of the onymic marker and definite article with personal names in Catalan varieties (Caro Reina 2014: 183)

However, taking a closer look at Central Catalan, we see another difference between the varieties of Girona and Barcelona: while the dialect of Girona follows the general Central Catalan pattern, in the dialect of Barcelona the onymic marker *en* has been replaced by the masculine definite article *el*:

Central Catalan variety	**M**	**F**	**M.F (with initial vowel)**
Barcelona	*el Pere*	*la Maria*	*l'Andreu*
Girona	*en Pere*	*la Maria*	*l'Andreu*

Table 18. Use of the onymic marker and definite article with personal names in the two Central Catalan varieties of Barcelona and Girona

The characteristics mentioned must be considered for the study design and are further specified in section 5.2.3. Note that whenever I speak of Catalan in the following without further specification, I mean the variety of Central Catalan.

5.2.2 Hypotheses

Unlike the study on Sicilian presented, the acceptability judgement study on Catalan solely operates with the following two independent variables:

[20] Cf. https://www.idescat.cat/pub/?id=eulp&lang=en (last accessed 30 June 2023).

- Agentivity (four values in accordance with the agentivity scale: *entailed agentivity*, *presupposed agentivity*, *potential agentivity*, *unspecified for agentivity*)
- DOM (binary values: *DOM* vs. *no-DOM*)

The main reason for not including referentiality (*definite NP* vs. *indefinite NP*) as a further variable is that, unlike in Sicilian, DOM with indefinite human NPs is said to be rather ungrammatical in Catalan.

The dependent variable is the *acceptability rating*.

Hypothesis 1 (H1) is formulated as an undirected hypothesis stating that a main effect of DOM is expected to appear:

(24) **H1 (DOM vs. no-DOM) and prediction for the acceptability judgement task:**

The ratings for direct objects exhibiting DOM differ from those that do not exhibit DOM:

P1: DOM ≠ no-DOM

Hypothesis 2 (H2) concerns the interaction between agentivity and DOM. Predictions 2a (P2a) and 2b (P2b) were formulated based on the agentivity scale. Note that ">" means 'is more acceptable/rated better than'.

(25) **H2 (Agentivity & DOM) and predictions for the acceptability judgement task:**

H2a. The more explicit agentivity information of the DO is communicated by the verbal predicate, the more acceptable DOM is.

H2b. The less explicit agentivity information of the DO is communicated by the verbal predicate, the more acceptable no-DOM is.

P2a: In the DOM condition, I predict the following scale of acceptability for the four verb classes ranging from entailed agentivity for the direct object to unspecified for the object's agentivity:

entailed agentivity & DOM > presupposed agentivity & DOM > potential agentivity & DOM > unspecified for agentivity & DOM

P2b: In the no-DOM condition, I predict the following scale of acceptability for the four verb classes ranging unspecified for the direct object's agentivity to entailed agentivity for the object:

unspecified for agentivity & no-DOM > potential agentivity & no-DOM > presupposed agentivity & no-DOM > entailed agentivity & no-DOM

In the following, the study design for the two versions of the questionnaire is addressed.

5.2.3 Study design

For testing agentivity, verb type has been manipulated. According to the study on Sicilian, for each of the four agentivity classes, five lexicalizations were defined:

(26) **Verb type: agentivity encoding of the direct object**

a. entailed agentivity for the object:

5 OE-psych verbs (*consolar* ‘to console’, *espantar* ‘to frighten’, *molestar* ‘to bother’, *ofendre* ‘to offend’, *provocar* ‘to provoke’)

b. presupposed agentivity for the object:

2 interaction verbs (*ajudar* ‘to help’ and *sentir* ‘to hear’), 3 highly transitive verbs (*aturar* ‘to stop’, *retener* ‘to catch’ and *salvar* ‘to save’)

c. potential agentivity for the object:

5 highly transitive verbs (*curar* ‘to cure’, *ferir* ‘to injure’, *insultar* ‘to insult’, *lesionar* ‘to cripple’, *matar* ‘to kill’)

d. object unspecified for agentivity:

2 visual perception verbs (*contemplar* ‘to watch’ and *veure* ‘to see’), 1 other perception verb (*trobar* ‘to find’), 1 SE-psych verb (*odiar* ‘to hate’) and 1 knowledge verb (*conéixer* ‘to know’)

The syntactic transitivity of the selected verbs was checked with a native speaker as exemplified for the verb *espantar* ‘to frighten’ in (27): the object NP of (27a) can be substituted by the direct object clitic *la* (F.SG) (vs. the indirect object clitic *li*) (27b) and passivization of the sentence is possible (27c).

(27) a. *El/En Pere va espantar* ***ø/a*** *la dona.*

the/OM PN have.PST.3SG frighten.INF ø/DOM the woman

‘Peter frightened the woman.’

b. *El/En Pere* ***la/*li*** *va espantar.*

the/OM PN CL.F.3SG.ACC/CL.3SG.DAT have.PST.3SG frighten.INF

‘Peter frightened her.’

c. *La dona va ser espantat per el/en Pere.*

the woman have.PST.3SG be.INF frighten.PTCP by the/OM PN

‘The woman was frightened by Peter.’

The object NP type was kept constant in that only human definite NPs were chosen. Just as for Sicilian, the lexicalizations of the direct object NP, listed in (28), were balanced between relational and sortal nouns (e.g. relational *la veïna* ‘the neighbour’ vs. sortal *la delinqüent* ‘the criminal’). The restriction to feminine NPs and NPs accompanied by the abbreviated masculine definite article *l’* is due to a phonological property of Eastern Catalan varieties taken up further below.

(28) **Object NP type (stable): human definite NP**

l'adversària 'the opponent', *l'agent de policia* 'the police officer' (2x), *la ballerina* 'the dancer', *la companya* 'the friend' (2x), *la companya de treball* 'the colleague', *la criminal* 'the criminal', *la delinqüent* 'the criminal', *la dona* 'the woman', *la escriptora* 'the writer', *la jardinera* 'the gardener', *la lladre* 'the thief', *la nena* 'the girl' (2x), *la pacient* 'the patient', *la patrona* 'the boss', *la presidenta* 'the president', *la veïna* 'the neighbour' (2x)

As for the subject NP, which always consisted of a personal name, common Catalan first names were chosen. In order to make the sentences sound as natural as possible, a postverbal adverbial phrase of either local, temporal, manner or causal type was added for each sentence. Sentence word order is left constant (SVO) controlling for the fact that left- or right-dislocated direct objects are more readily *a*-marked than those occurring in canonical word order (cf. e.g. Iemmolo 2010a, in preparation; Solà 1994).

There is one phonological characteristic of Eastern Catalan varieties, such as Central Catalan, which must be considered for the design of test items: Unlike Western Catalan, Central Catalan exhibits centralization of /a ɛ e/ > [ə] in unstressed syllables (cf. Caro Reina 2019: 97). As can be seen in example (29a), /a/ is pronounced as [ə] in unstressed syllables, while in North-Western Catalan the process does not apply and /a/ is pronounced as [a] (29b):

(29) a. Eastern Catalan (here Central Catalan):

*pa**a**lla* [ˈpa**ə**ʎə] vs. *pa**a**ll-et-a* [pə**ə**ˈʎɛtə]

'straw' 'straw-DIM-F'

b. Western Catalan (here North-Western Catalan):

*pa**a**lla* [ˈpa**a**ʎa] vs. *pa**a**ll-et-a* [pa**a**ˈʎeta]

'straw' 'straw-DIM-F'

(Caro Reina 2019: 97; emphasis: S.M.)

For my purposes, it is important that the differential object marker *a* is an unstressed vowel pronounced as [ə] in Central Catalan. When two unstressed vowels involving schwa are adjacent, vowel deletion takes place and one single schwa sound is produced (cf. e.g. Caro Reina 2019: 121f.; Vallverdú Albornà 2002: 144f.). Thus, verbs ending in a final [ə] such as *estim-**a*** [əsˈtimə] 'love-PRS.3SG' that are followed by the masculine definite article *el* [əl] (30a) or by DOM fused with *el* [əl] (30b), both lead to the deletion of one schwa. Crucially, the morphosyntactic difference, that is either the lack or the occurrence of DOM, is not detectable in the spoken language (cf. Escandell-Vidal 2009: 842, fn. 4).

(30) a. *Estim**a*** ***el*** *Pere.* [əsˈtiməlˈpɛrə]

love.PRS.3SG the PN

'S/he loves Peter.'

b. *Estim**a*** ***al*** *Pere.* [əsˈtiməlˈpɛrə]

love.PRS.3SG DOM.the PN

'S/he loves Peter.' (Escandell-Vidal 2009: 842, fn. 4)

The phonological processes described (centralization and vowel deletion) have a two-fold impact on the design: First, only NPs with the feminine definite article *la* (or *l'*) (31a) and NPs accompanied by the abbreviated masculine definite article *l'* (31b) can be used as direct objects. Object NPs with the full masculine definite article *el* (31c), in contrast, could not be employed since they do not allow us to hear a difference between the no-DOM and DOM condition. The phonetic sequences of no-DOM and DOM in combination with the definite article are indicated in square brackets.

(31) a. *La Maria ha curat* ***ø/a la [lə/ələ]*** *pacient*

the PN have.PRS.3SG cure.PTCP ø/DOM the patient

en dues setmanes.

in two weeks

'Maria has cured the patient in two weeks.'

b. *El Vicent ha lesionat* ***ø/a l' [l/əl]***

the PN have.PRS.3SG cripple.PTCP ø/DOM the

agent de policia amb un ganivet.

policeman with a knife

'Vicent has crippled the policeman with a knife.'

c. *La Maria ha curat* ***ø/a el [əl/əl]*** *pacient*

the PN have.PRS.3SG cure.PTCP ø/DOM the patient

en dues setmanes.

in two weeks

'Maria has cured the patient in two weeks.'

Second, the tense chosen was the present perfect (*pretèrit indefinit*) (32a), since for verbs endings in *-ar* [a][21] in Central Catalan both the present tense (*present d'indicatiu*) (32b) and the periphrastic preterite (*pretèrit perfet*) (32c) would lead to vowel deletion in the third person singular, which makes DOM phonetically

[21] The deletion of word-final *r* is a further phonological characteristic of Catalan. With the exception of Valencian, it applies to all varieties (cf. Caro Reina 2019: 136ff.).

impossible to distinguish from the absence of DOM. The phonetic sequences of the verb-final sound in combination with no-DOM and DOM are indicated in square brackets.

(32) a. Present perfect (*pretèrit indefinit*)

El/En Marc ha aturat ***ø/a*** ***[t/tə]*** *la lladre*

the/OM PN have.PRS.3SG stop.PTCP ø/DOM the thief

davant de la botiga.

in front of the store

'Marc has stopped the thief outside the store.'

b. Present tense (*present d'indicatiu*)

El/En Marc atura ***ø/a*** ***[ə/ə]*** *la lladre davant de*

the/OM PN stop.PRS.3SG ø/DOM the thief in front of

la botiga.

the store

'Marc stops the thief outside the store.'

c. Periphrastic Preterite (*pretèrit perfet*)

El/En Marc va atura ***ø/a*** ***[a/a]***[22] *la lladre*

the/OM PN have.PST.3SG stop.INF ø/DOM the thief

davant de la botiga.

in front of the store

'Marc stopped the thief outside the store.'

Just as for Sicilian, an exception has been made for the verb *conéixer* 'to know': In order to guarantee a stative reading ('to know') and prevent an inchoative, symmetrical reading with entailed *sentience* for the direct object ('to get to know'), I used the present tense form of this verb (PRS.3SG: *coneix*). Since the form does not end in a vowel but in a consonant [ʃ], the difference between DOM and no-DOM is clearly detectable.

As mentioned in 5.2.1, the dialects of Barcelona and Girona can be distinguished by one main morphosyntactic characteristic: the marking of male personal names. While in Barcelona they are marked by the masculine definite article *el* (abbreviated

[22] Here vowel deletion follows from the adjacency of the stressed verb-final vowel [a] of the infinitive [ə.tuˈra] and the unstressed *a*-marker [ə] with the latter being deleted (cf. Caro Reina 2019: 121f.).

l'), in Girona a more archaic form was preserved, namely *en* (abbreviated *n'*). Female personal names are marked by the feminine definite article *la* (or its abbreviated form *l'*) in both varieties. Respecting this difference, (33) displays the 8 critical conditions of the study:

(33) **Critical conditions**

a. entailed agentivity + [± DOM]

El/En Jordi ha provocat ***ø/a*** *l' agent de policia*

the/OM PN have.PRS.3SG provoke.PTCP ø/DOM the policeman

amb insults.

with insults

'Jordi has provoked the policeman with (his) insults.'

b. presupposed agentivity + [± DOM]

El/En Marc ha aturat ***ø/a*** *la delinqüent*

the/OM PN have.PRS.3SG stop.PTCP ø/DOM the criminal

davant de la botiga.

in front of the store

'Marc has stopped the criminal in front of the store.'

c. potential agentivity + [± DOM]

La Maria ha curat ***ø/a*** *la pacient en dues*

the PN have.PRS.3SG cure.PTCP ø/DOM the patient in two

setmanes.

weeks

'Maria has cured the patient in two weeks.'

d. unspecified for agentivity + [± DOM]

La Marta ha vist ***ø/a*** *la presidenta a l' aeroport.*

the PN have.PRS.3SG see.PTCP ø/DOM the president at the airport

'Marta has seen the president at the airport.'

Five sets of the 8 (4 × 2) critical conditions were constructed. The 40 sentences were distributed to two lists A and B according to Latin square, each containing 20 sentences. Additionally, 20 fillers were integrated in each list. They were identical for the two lists and comprised of 10 grammatical and 10 ungrammatical fillers (with incorrect prepositions, wrong clitics and wrong choice of tense). Thus, in

comparison to the Sicilian study (2/3 items vs. 1/3 fillers), there was an equal number of fillers and test items (half-half). This was considered particularly important since the normative pressure to reject DOM in Catalan is very high, which makes distracting the participants more important. In a further step, the entire material was pseudo-randomized.[23] The set of test items and fillers was recorded with a 23-year-old male native speaker for Barcelona[24] and with a 25-year-old female native speaker for Girona.[25] The judgements are given on a five-point scale (ranging from 1 – *not acceptable* to 5 – *totally acceptable*).

5.2.4 Participants

Participants were recruited online in April and May 2019 using the LimeSurvey tool. Before starting, they were informed of the academic purpose of the study. The research topic was not mentioned; instead participants were told that the study aimed to investigate the Catalan varieties of Barcelona and Girona, respectively, in general.

As for the Barcelona version, 67 Catalan speakers originating from the province of Barcelona volunteered in the study. As a consequence of the filler analysis, one participant of version A and one participant of version B were eliminated since their mean ratings of the ungrammatical fillers considerably deviated from what was expected (mean value better than 3.0).[26] Furthermore, two participants of version A and one participant of version B were excluded because they indicated that Catalan was not their mother tongue. As a result, 62 questionnaires were included in the analysis (36 of version A and 26 of version B). Table 19 gives an overview of the main sociodemographic data:

Barcelona

Question	**Number of participants (total n=62)**
Age	Mean age = 34.8 years, range 20-62 years
Sex	Female (44), male (18)
Level of study	Secondary school (1), High school (6), Apprenticeship (1) University education (54)

[23] For a list of the test items, cf. Appendix 2.1 and 2.2. The audio versions A and B of the questionnaire are available here: https://uni-koeln.sciebo.de/s/kmQyHw8DHzUsg0t.
[24] The speaker grew up in Germany and is thus a so-called *heritage speaker* of Catalan. He is bilingual Catalan-German.
[25] The speaker is from Girona and is bilingual Catalan-Spanish.
[26] The filler analysis consisted of the following steps: First, it was checked whether the individual filler sentences worked the way they should. More precisely, a grammatical filler should show a mean value not lower than 4.0, whereas an ungrammatical filler should not get a mean rating better than 2.0. The results for all twenty filler sentences met the requirements, so that no filler had to be excluded. Second, the mean acceptability ratings for grammatical and ungrammatical fillers were analyzed for each participant. Participants who displayed a mean rating lower than 3.5 for grammatical and better than 3.0 for ungrammatical fillers were eliminated from the data analysis. After all relevant participants were excluded (also those whose mother tongue was not Catalan), the overall mean rating was 4.6 [CI 4.46; 4.73] for grammatical and 1.4 [CI 1.27; 1.54] for ungrammatical fillers.

Mother tongue	Catalan (62), Spanish (21), other (5)
Second language(s)	English (52), German (17), Italian (6), Spanish (6), French (38), Portuguese (1), Japanese (1), Dutch (1), Russian (1), Catalan Sign Language (2)
Language(s) spoken at home	Only Catalan (45), only Spanish (4), both Catalan and Spanish (8), other (6)
Preferred language: Catalan or Spanish (or no preference)	Catalan (43), Spanish (3), no preference (16)

Table 19. Participants of the study (Barcelona)

As regards the Girona version, 65 Catalan speakers originating from the province of Girona volunteered in the study. The filler analysis led to the exclusion of one participant of version B since the mean rating for ungrammatical fillers sharply deviated from the expected rating (mean value better than 3.0).[27] Additionally, three participants of version A and B, respectively, were eliminated because Catalan was not their mother tongue. Hence, 58 questionnaires were considered for the analysis (27 of version A and 31 of version B). Table 20 summarizes the main sociodemographic data:

Girona

Question	**Number of participants (total n=58)**
Age	Mean age = 39.3 years, range 19-71 years
Sex	Female (37), male (21)
Level of study	High school (4), Apprenticeship (3), University education (51)
Mother tongue	Catalan (58), Spanish (14), other (0)
Second language(s)	English (49), German (6), Italian (11), Spanish (3), French (31), Portuguese (2), Chinese (1), Occitan (1)
Language(s) spoken at home	Only Catalan (54), only Spanish (4), both Catalan and Spanish (0), other (0)
Preferred language: Catalan or Spanish (or no preference)	Catalan (49), Spanish (1), no preference (8)

Table 20. Participants of the study (Girona)

5.2.5 Procedure and data analysis

As for Sicilian, the two lists of the questionnaire for the two Central Catalan varieties were presented to the participants in oral format, that is as MP3 audio recordings each consisting of one sentence. Participants were randomly assigned to one

[27] Just as for the Barcelona version, in a first step of the filler analysis it was checked if individual filler sentences deviated from the expected rating, which was not the case: No grammatical filler had a mean value lower than 4.0 and no ungrammatical filler displayed a mean rating better than 2.0. Hence, no filler had to be excluded. In a second step, the mean acceptability ratings for grammatical and ungrammatical fillers were analyzed for each participant. Participants showing a mean rating lower than 3.5 for grammatical and better than 3.0 for ungrammatical fillers were eliminated from the data analysis. After the exclusion of all relevant participants (including those whose mother tongue was not Catalan), the overall mean rating was 4.64 [CI 4.51; 4.78] for grammatical and 1.28 [CI 1.14; 1.41] for ungrammatical fillers.

of the two lists in the LimeSurvey tool. Before starting the questionnaire, they were asked to fill in a form with their sociodemographic data.[28] In a next step, they had to read a short instruction in which they were told to imagine themselves in an informal conversation with people they usually talk to in Catalan (e.g. family or friends). They were asked to imagine the sentences occurring in such a context. Participants were asked to rate each of the forty sentences (20 test items and 20 fillers) on a 5-point scale ranging from 1 (*not acceptable*) to 5 (*totally acceptable*). Sentences were presented one-by-one on the screen, each MP3 could be listened to as often as desired. All data were collected anonymously. The language of the questionnaire was Standard Catalan.

Regarding the statistical analysis, I decided to use a mixed-effects regression model. Just like in the analysis of the Sicilian data, I defined (i) participants and (ii) items as random effects. The fixed effects referred to the independent variables, which where only two in contrast to the Sicilian study, namely (i) DOM and (ii) agentivity. Recall that both independent variables were categorical variables: while (i) DOM has binary values (*DOM* vs. *no-DOM*), (ii) agentivity consisted of four values, namely *entailed agentivity*, *presupposed agentivity*, *potential agentivity* and *unspecified for agentivity*. According to my working hypothesis (section 4.4), I assumed the following ranking of the four agentivity levels:

(34) entailed agentivity > presupposed agentivity > potential agentivity > unspecified for agentivity

The formulated predictions P2a and P2b of H2 (Agentivity & DOM) implemented this ranking. As argued in 5.1.5, for the statistic modelling I decided to calculate the mean value for each agentivity level in each condition and to compare them in the following way:

(35) **Performed comparisons of agentivity levels in statistical modelling:**

a. entailed agentivity vs. presupposed agentivity

b. entailed agentivity vs. potential agentivity

c. entailed agentivity vs. unspecified for agentivity

d. presupposed agentivity vs. potential agentivity

e. presupposed agentivity vs. unspecified for agentivity

f. potential agentivity vs. unspecified for agentivity

[28] An example of the form is provided in Appendix 2.4.

The data analysis was again carried out in R (version 3.6.1: R Core Team 2017) with the packages lme4[29] (Bates et al. 2015) and lmerTest[30] (Kuznetsova et al. 2017). Effects were considered significant if $p < .05$.

5.2.6 Results

In the following, the results are summarized according to the two hypotheses worked out in section 5.2.2.[31] According to the hypotheses, both a main effect of DOM (H1) and an interaction effect of DOM and agentivity (H2) were taken into account. Note again that unlike the study on Sicilian, the factor referentiality was kept stable in the present study in that solely definite human object NPs were considered.

As a prior check for outliers among the test items in both datasets revealed, two pairs of items had to be disregarded for the analysis of the Barcelona data (verbs *ajudar* 'to help' and *insultar* 'to insult') and one pair of items had to be eliminated from the analysis of the Girona data (verb *ajudar* 'to help'). A pair of items was considered an *outlier* if both the DOM and the no-DOM version considerably deviated from the mean values of the other pairs of items, i.e. more specifically if both versions displayed a mean acceptability value lower than 3.0.

5.2.6.1 Hypothesis 1 (DOM vs. no-DOM)

The undirected H1 predicting a main effect of DOM on sentence acceptability, as repeated in P1 (36), was confirmed.

(36) **H1: DOM vs. no-DOM**

P1. DOM ≠ no-DOM

Figure 9 and **Figure 10** illustrate the distribution of acceptability values for DOM vs. no-DOM for Barcelona and Girona, respectively. As indicated by the boxplots, the dispersion of the data is considerably higher for DOM than for no-DOM in both datasets. While for DOM, the range of acceptability values covers the entire scale with 25% of the answers being lower than 2 for Barcelona and lower than 3 for Girona, for no-DOM 75% of the acceptability values lie between 4 and 5 for both varieties. **Figure 11** and **Figure 12** display the estimated marginal means (EMM) of acceptability values for DOM vs. no-DOM with 95% confidence intervals for Barcelona and Girona, respectively. The values are listed in **Table 21** below. As can be seen, in both varieties the absence of DOM is rated clearly better than the occurrence of DOM: While for Barcelona, the acceptability value for DOM reaches an EMM of 3.33 (95% CI [3.13; 3.53]), the EMM for Girona was slightly higher, reaching 3.83 (95% CI [3.65; 4.01]). Correspondingly, the absence of DOM was

[29] Linear Mixed-Effects Models using 'Eigen' and S4.
[30] Tests in Linear Mixed Effects Models.
[31] Also here I would like to thank Maximilian Hörl for his help with the statistical analysis of my data.

rated 4.24 (95% CI [4.04; 4.44]) for Barcelona and – slightly better – namely 4.43 (95% CI [4.25; 4.61]) for Girona. In both cases, the difference between DOM and no-DOM is highly significant with $p < .0001$ (cf. **Table 22**).

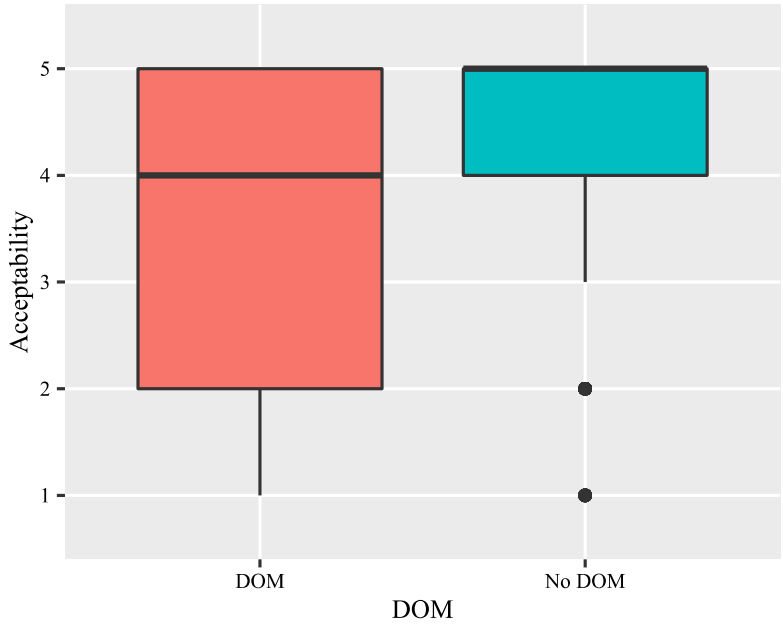


Figure 9. Boxplots of acceptability values for DOM vs. no-DOM (H1) with definite human NPs with median values and outliers for Barcelona

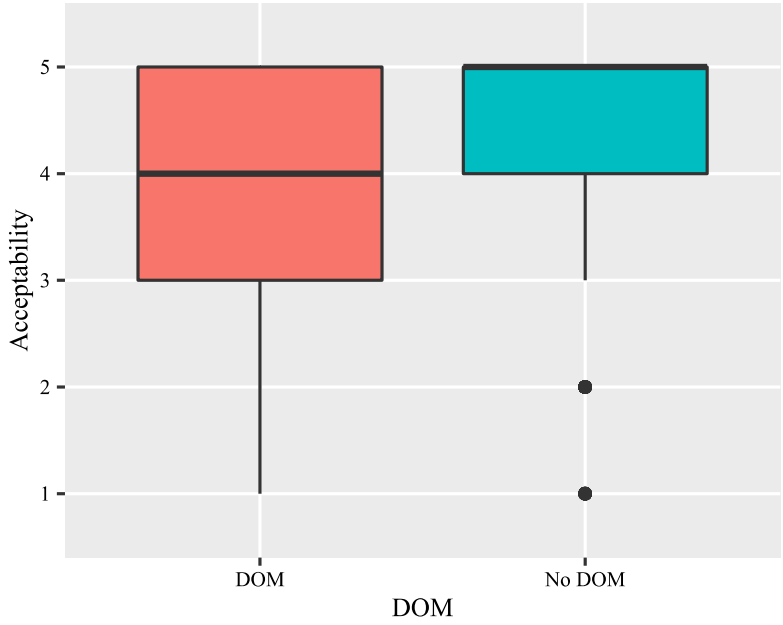


Figure 10. Boxplots of acceptability values for DOM vs. no-DOM (H1) with definite human NPs with median values and outliers for Girona

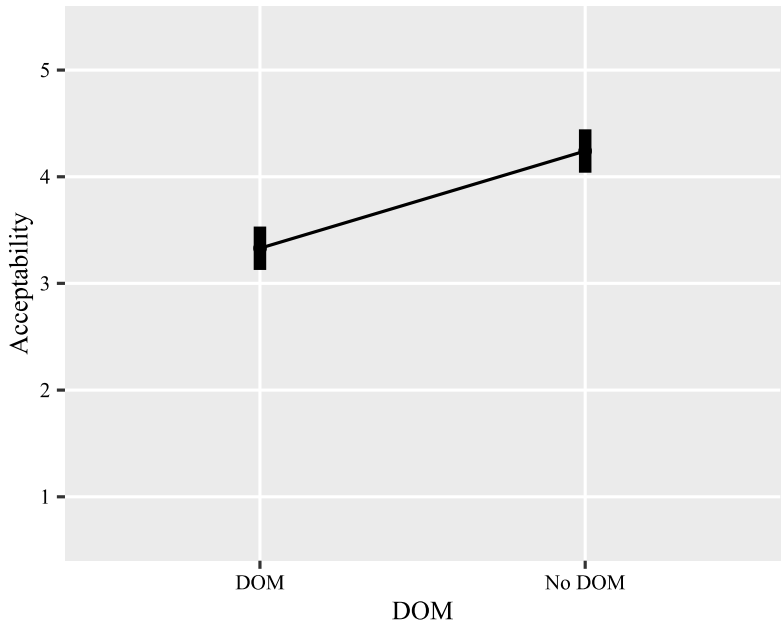


Figure 11. Estimated marginal means (EMM) for DOM vs. no-DOM (H1) with definite human NPs for Barcelona, confidence level used: 0.95

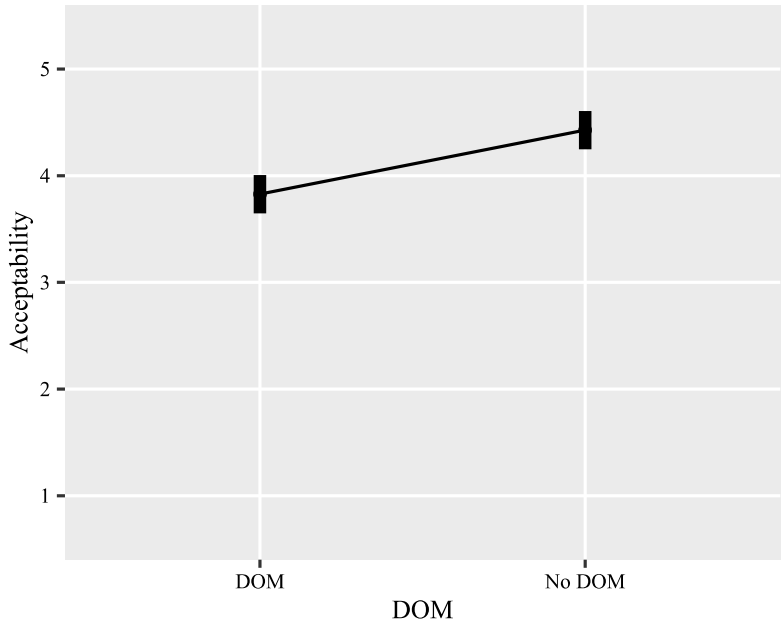


Figure 12. Estimated marginal means (EMM) for DOM vs. no-DOM (H1) with definite human NPs for Girona, confidence level used: 0.95

DOM	EMM	SE	df	lower CL	upper CL
Barcelona					
DOM	3.33	0.102	74.8	3.13	3.53
no-DOM	4.24	0.102	74.8	4.04	4.44
Girona					
DOM	3.83	0.0900	73.3	3.65	4.01
no-DOM	4.43	0.0902	73.4	4.25	4.61

Table 21. Estimated marginal means (EMM) of acceptability values for DOM vs. no-DOM (H1) with definite human NPs for Barcelona and Girona, degrees-of-freedom method: kenward-roger, confidence level used: 0.95

Contrast	estimate	SE	df	t.ratio	p.value
Barcelona					
DOM vs. no-DOM	-0.911	0.086	32.6	-10.588	**< .0001**
Girona					
DOM vs. no-DOM	-0.6	0.0768	35	-7.817	**< .0001**

Table 22. Parameter estimates for DOM vs. no-DOM (H1) with definite human NPs for Barcelona and Girona, degrees-of-freedom method: kenward-roger, t-test

5.2.6.2 Hypothesis 2 (Agentivity & DOM)

A main effect of agentivity was not found in either of the two datasets.

H2, which concerns the interaction between DOM and agentivity with the predictions P2a and P2b taken up again in (37), was partly confirmed for both Barcelona and Girona, namely for condition 2b. For condition 2a, it was rejected.

(37) **H2: Agentivity & DOM**

P2a.
entailed agentivity & DOM > presupposed agentivity & DOM > potential agentivity & DOM > unspecified for agentivity & DOM

P2b.
unspecified for agentivity & no-DOM > potential agentivity & no-DOM > presupposed agentivity & no-DOM > entailed agentivity & no-DOM

The results are presented separately for each variety.

(1) Barcelona

Figure 13 shows the distribution of acceptability values in condition 2a (*Agentivity & DOM*) and 2b (*Agentivity & no-DOM*). It can be seen by the boxplots that the dispersion of the data is considerably higher for DOM (2a) than for no-DOM (2b)

across agentivity levels: In the DOM condition, acceptability values for all four agentivity classes range over the entire scale with 25% of the answers being lower than 2. The median lies at 4 for *entailed agentivity*, *presupposed agentivity* and *unspecified for agentivity*, while it is at 3 for *potential agentivity*. In the no-DOM condition, by contrast, for *presupposed agentivity*, *potential agentivity* and *unspecified for agentivity*, acceptability values only range from 3 to 5 with 75% of the acceptability values being between 4 and 5 (with a few outliers, indicated by the dots). The median is at 5 for the three classes. The class *entailed agentivity* deviates from the other three in that it displays a higher dispersion of acceptability values: while 75% of the values are distributed from 3 to 5, the remaining 25% range from 1 to 3 (median value = 4). **Figure 14** shows the EMM of acceptability values in both conditions with 95% confidence intervals. The values are listed in **Table 23** below. In condition 2a, contrary to prediction, *unspecified for agentivity* obtains the highest ratings with an EMM of 3.46 (95% CI [3.19; 3.73]), followed by *entailed agentivity* (3.36, 95% CI [3.09; 3.63]), *presupposed agentivity* (3.28, 95% CI [2.99; 3.57]) and *potential agentivity* (3.17, 95% CI [2.88; 3.46]):

(38) Obtained scale of acceptability in condition 2a (*Agentivity & DOM*):

unspecified for agentivity & DOM > entailed agentivity & DOM > presupposed agentivity & DOM > potential agentivity & DOM

In condition 2b, in accordance with P2b, *unspecified for agentivity* (4.41, 95% CI [4.13; 4.68]) is rated highest and *entailed agentivity* is least acceptable exhibiting an EMM of 4.04 (95% CI [3.77; 4.31]). Yet, the higher ratings for *presupposed agentivity* (4.40, 95% CI [4.11; 4.69]) than for *potential agentivity* (4.14, 95% CI [3.85; 4.43]) oppose the predictions.

(39) Obtained scale of acceptability in condition 2b (*Agentivity & no-DOM*):

unspecified for agentivity & no-DOM > presupposed agentivity & no-DOM > potential agentivity & no-DOM > entailed agentivity & no-DOM

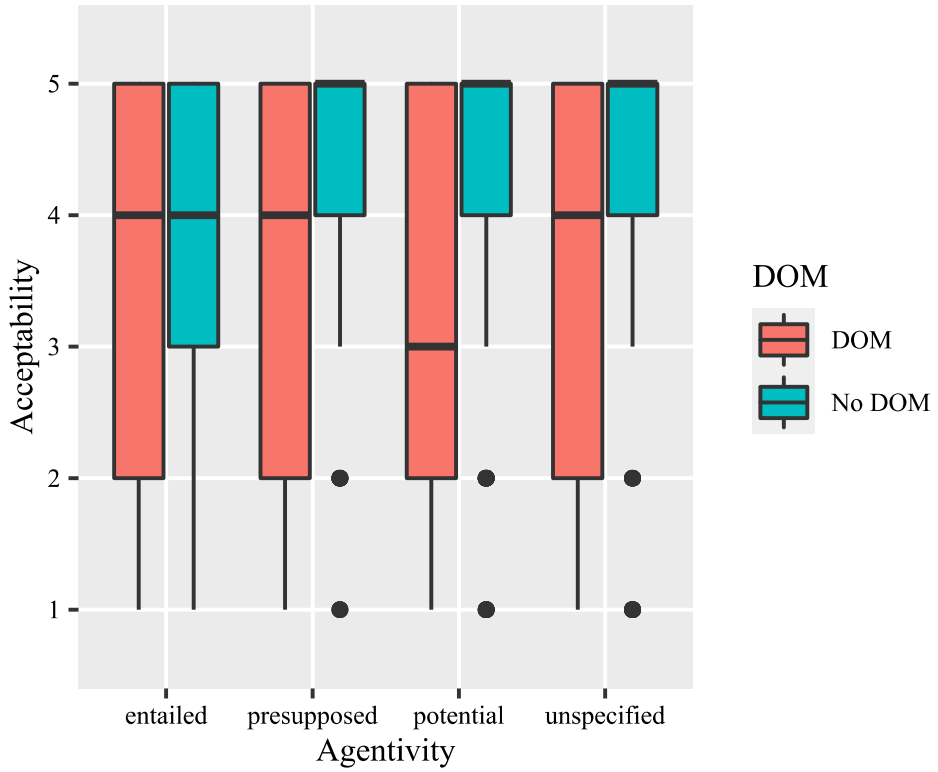


Figure 13. Boxplots of acceptability values in condition 2a (Agentivity & DOM) and 2b (Agentivity & no-DOM) with median values and outliers for Barcelona

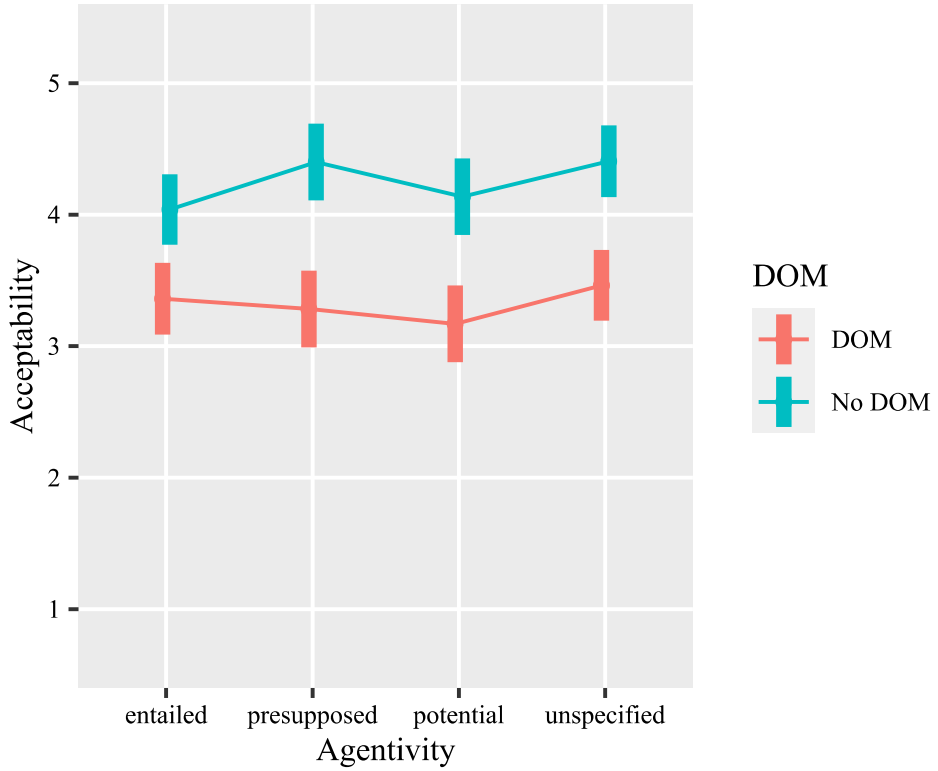


Figure 14. Estimated marginal means (EMM) in condition 2a (Agentivity & DOM) and 2b (Agentivity & no-DOM) for Barcelona, confidence level used: 0.95

Agentivity	EMM	SE	df	lower CL	upper CL
(i) entailed					
2a (DOM)	3.36	0.136	56.0	3.09	3.63
2b (no-DOM)	4.04	0.136	53.7	3.77	4.31
(ii) presupposed					
2a (DOM)	3.28	0.145	50.1	2.99	3.57
2b (no-DOM)	4.40	0.145	50.1	4.11	4.69
(iii) potential					
2a (DOM)	3.17	0.145	50.1	2.88	3.46
2b (no-DOM)	4.14	0.145	50.1	3.85	4.43
(iv) unspecified					
2a (DOM)	3.46	0.134	53.7	3.19	3.73
2b (no-DOM)	4.41	0.136	56.0	4.13	4.68

Table 23. Estimated marginal means (EMM) of acceptability values in condition 2a (Agentivity & DOM) and 2b (Agentivity & no-DOM) for Barcelona, degrees-of-freedom method: kenward-roger, confidence level used: 0.95

H2a (Barcelona)

As the mixed-effects models analysis reveals, the predictions for H2a (*Agentivity & DOM*) were not confirmed. The numeric differences between *entailed agentivity* vs. *presupposed agentivity* (diff1), *entailed agentivity* vs. *potential agentivity* (diff2) and *presupposed agentivity* vs. *potential agentivity* (diff4) were not statistically significant (cf. **Table 24**). Furthermore, also the unexpected differences between *entailed agentivity* vs. *unspecified for agentivity* (diff3), *presupposed agentivity* vs. *unspecified for agentivity* (diff5) and *potential agentivity* vs. *unspecified for agentivity* (diff6) did not attain statistical significance.

Fixed effects	Estimate	Standard error	*t*-value	Pr(>|t|)
Diff1: entailed vs. presupp. agentivity & DOM	-0.07807	0.16178	-0.483	0.633246
Diff2: entailed vs. potential agentivity & DOM	-0.19088	0.16178	-1.180	0.248273
Diff3: entailed vs. unspec. for agentivity & DOM	0.10174	0.15225	0.668	0.509702
Diff4: presupp. vs. potential agentivity & DOM	-0.11281	0.16941	-0.666	0.5112
Diff5: presupp. vs. unspec. for agentivity & DOM	0.17982	0.16001	1.124	0.2714
Diff6: potential vs. unspec. for agentivity & DOM	0.29263	0.16001	1.829	0.0789

Table 24. Parameter estimates for the fixed effects in condition 2a (Agentivity & DOM) for Barcelona

H2b (Barcelona)

With regard to H2b (*Agentivity & no-DOM*), the predictions were partially confirmed. The differences between *entailed agentivity* vs. *presupposed agentivity* (diff1) and *entailed agentivity* vs. *unspecified for agentivity* (diff3), with *entailed agentivity* being less accepted than *presupposed agentivity* and *unspecified for agentivity*, were statistically significant (cf. **Table 25**). The difference between *entailed agentivity* vs. *potential agentivity* (diff2), by contrast, was only minimal and hence not significant, neither was the (numerically greater) difference between *potential agentivity* vs. *unspecified for agentivity* (diff6). The two differences that did not match the predicted ranking, i.e. *presupposed agentivity* vs. *potential agentivity* (diff4) and *presupposed agentivity* vs. *unspecified for agentivity* (diff5), were also not significant.

Fixed effects	**Estimate**	**Standard error**	***t*-value**	**Pr(>|t|)**
Diff1: entailed vs. presupp. agentivity & no-DOM	0.36111	0.16001	2.257	**0.032646***
Diff2: entailed vs. potential agentivity & no-DOM	0.09742	0.16001	0.609	0.547897
Diff3: entailed vs. unspec. for agentivity & no-DOM	0.36687	0.15225	2.410	**0.023147***
Diff4: presupp. vs. potential agentivity & no-DOM	-0.263691	0.169406	-1.557	0.1314
Diff5: presupp. vs. unspec. for agentivity & no-DOM	0.005759	0.161779	0.036	0.9719
Diff6: potential vs. unspec. for agentivity & no-DOM	0.26945	0.16178	1.666	0.197

Table 25. Parameter estimates for the fixed effects in condition 2b (Agentivity & no-DOM) for Barcelona

(2) Girona

Figure 15 displays the distribution of acceptability values in condition 2a (*Agentivity & DOM*) and 2b (*Agentivity & no-DOM*). As the boxplots illustrate, the dispersion of the data is clearly higher for DOM (2a) than for no-DOM (2b) across agentivity levels. In the DOM condition, acceptability values for all four agentivity classes are distributed over the entire scale with 25% of the answers being lower than 3 or – in the case of *presupposed agentivity* – 2. The median lies at 4 for all four classes. In the no-DOM condition, by contrast, for *entailed agentivity*, *presupposed agentivity* and *potential agentivity*, acceptability values only range from 3 to 5 with 75% of the acceptability values lying between 4 and 5 (with a few outliers, indicated

by the dots). The median is at 5 for the three classes. The class *unspecified for agentivity* provides an even clearer picture: Except for a few outliers, there is no dispersion and nearly 100% of the acceptability values lie at 5. In **Figure 16** the estimated marginal means (EMM) of acceptability values in both conditions are given with 95% confidence intervals. The values are reported in detail in **Table 26** below. In condition 2a, as for Barcelona and contrary to prediction, *unspecified for agentivity* reaches the highest score of acceptability with an EMM of 3.98 (95% CI [3.74; 4.23]), followed by *entailed agentivity* (3.85, 95% CI [3.61; 4.10]), *potential agentivity* (3.79, 95% CI [3.55; 4.03]) and *presupposed agentivity* (3.65, 95% CI [3.39; 3.92]):

(40) Obtained scale of acceptability in condition 2a (*Agentivity & DOM*):

unspecified for agentivity & DOM > entailed agentivity & DOM > potential agentivity & DOM > presupposed agentivity & DOM

In condition 2b, as predicted, *unspecified for agentivity* (4.63, 95% CI [4.38; 4.87]) obtains the highest acceptability ratings. Moreover, *presupposed agentivity* (4.43, 95% CI [4.17; 4.70]) is better accepted than *entailed agentivity* (4.34, 95% CI [4.09; 4.58]). Contrary to expectations, *presupposed agentivity* and *entailed agentivity* are both rated better than *potential agentivity* (4.31, 95% CI [4.06; 4.56]):

(41) Obtained scale of acceptability in condition 2b (*Agentivity & no-DOM*):

unspecified for agentivity & no-DOM > presupposed agentivity & no-DOM > entailed agentivity & no-DOM > potential agentivity & no-DOM

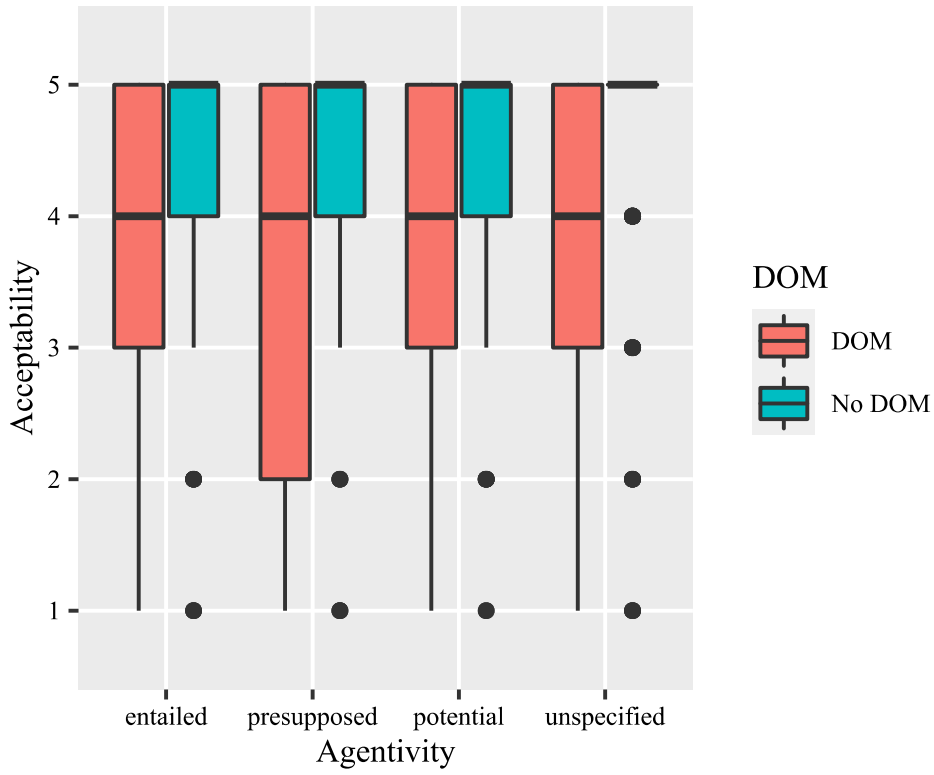


Figure 15. Boxplots of acceptability values in condition 2a (Agentivity & DOM) and 2b (Agentivity & no-DOM) with median values and outliers for Girona

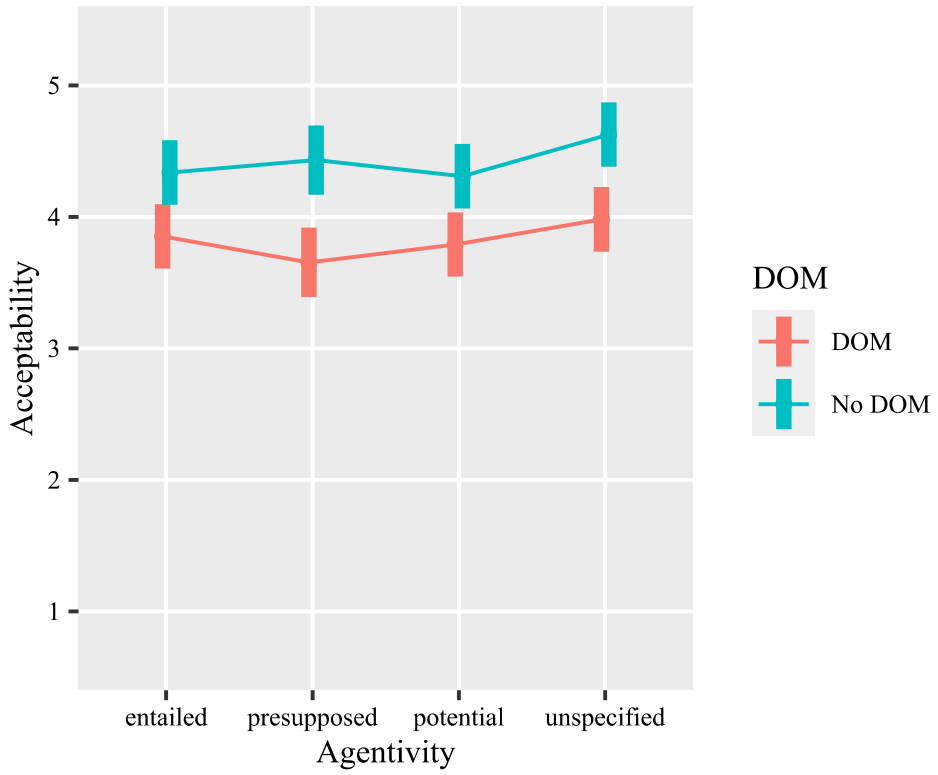


Figure 16. Estimated marginal means (EMM) in condition 2a (Agentivity & DOM) and 2b (Agentivity & no-DOM) for Girona, confidence level used: 0.95

Agentivity	EMM	SE	df	lower CL	upper CL
(i) entailed					
2a (DOM)	3.85	0.122	56.1	3.61	4.10
2b (no-DOM)	4.34	0.123	57.1	4.09	4.58
(ii) presupposed					
2a (DOM)	3.65	0.132	52.1	3.39	3.92
2b (no-DOM)	4.43	0.132	52.1	4.17	4.70
(iii) potential					
2a (DOM)	3.79	0.122	56.1	3.55	4.03
2b (no-DOM)	4.31	0.123	57.1	4.06	4.56
(iv) unspecified					
2a (DOM)	3.98	0.123	57.1	3.74	4.23
2b (no-DOM)	4.63	0.122	56.1	4.38	4.87

Table 26. Estimated marginal means (EMM) of acceptability values in condition 2a (Agentivity & DOM) and 2b (Agentivity & no-DOM) for Girona, degrees-of-freedom method: kenward-roger, confidence level used: 0.95

H2a (Girona)

Similar to the results for Barcelona, H2a (*Agentivity & DOM*) is also rejected for Girona. The numeric differences between *entailed agentivity* vs. *presupposed agentivity* (diff1) and *entailed agentivity* vs. *potential agentivity* (diff2) proved not statistically significant (cf. **Table 27**). Likewise, the differences between *entailed agentivity* vs. *unspecified for agentivity* (diff3), *presupposed agentivity* vs. *potential agentivity* (diff4) and *potential agentivity* vs. *unspecified for agentivity* (diff6), which all pointed in the reverse direction than predicted by the scale, were not significant. By contrast, the difference between *presupposed agentivity* and *unspecified for agentivity* (diff5) reaches statistical significance, which contradicts the prediction that *presupposed agentivity* would outrank *unspecified for agentivity*.

Fixed effects	**Estimate**	**Standard error**	***t*-value**	**Pr(>|t|)**
Diff1: entailed vs. presupp. agentivity & DOM	-0.19791	0.14795	-1.338	0.19221
Diff2: entailed vs. potential agentivity & DOM	-0.06275	0.13898	-0.452	0.65529
Diff3: entailed vs. unspec. for agentivity & DOM	0.12881	0.14027	0.918	0.36654
Diff4: presupp. vs. potential agentivity & DOM	0.1352	0.1480	0.914	0.3691
Diff5: presupp. vs. unspec. for agentivity & DOM	0.3267	0.1486	2.199	**0.0365***
Diff6: potential vs. unspec. for agentivity & DOM	0.19156	0.14027	1.366	0.183245

Table 27. Parameter estimates for the fixed effects in condition 2a (Agentivity & DOM) for Girona

H2b (Girona)

As for Barcelona, H2b (*Agentivity & no-DOM*) was partly confirmed. The differences between *entailed agentivity* vs. *unspecified for agentivity* (diff3) and between *potential agentivity* vs. *unspecified for agentivity* (diff6) were statistically significant in that in both cases, the absence of DOM was rated significantly better with *unspecified for agentivity* than with *entailed agentivity* and *potential agentivity*, respectively (cf. **Table 28**). Conversely, the numeric differences between *entailed agentivity* vs. *presupposed agentivity* (diff1) and *presupposed agentivity* vs. *unspecified for agentivity* (diff5) were not significant, neither were the unexpected differences between *entailed agentivity* vs. *potential agentivity* (diff2) and *presupposed agentivity* vs. *potential agentivity* (diff4).

Fixed effects	**Estimate**	**Standard error**	***t*-value**	**Pr(>|t|)**
Diff1: entailed vs. presupp. & no-DOM	0.09474	0.14861	0.638	0.52910

Diff2: entailed vs. potential agentivity & no-DOM	-0.02790	0.14036	-0.199	0.84390
Diff3: entailed vs. unspec. & no-DOM	0.28872	0.14027	2.058	**0.04926***
Diff4: presupp. vs. potential agentivity & no-DOM	-0.12264	0.14861	-0.825	0.416
Diff5: presupp. vs. unspec. & no-DOM	0.19398	0.14795	1.311	0.201
Diff6: potential vs. unspec. & no-DOM	0.31662	0.14027	2.257	**0.032231***

Table 28. Parameter estimates for the fixed effects in condition 2b (Agentivity & no-DOM) for Girona

In conclusion, the pattern of acceptability for DOM in Catalan can be described as follows: H1 (DOM vs. no-DOM) was confirmed. In both varieties, the ratings for direct objects exhibiting DOM differ significantly from those that do not exhibit DOM in that the former are significantly less accepted than the latter. H2 (Agentivity & DOM), by contrast, was only partly confirmed. In condition 2a (*Agentivity & DOM*), the statistic analysis did not confirm the predicted ranking of agentivity classes at all. There was only one significant effect that instead pointed in the opposite direction:

(42) Effect in condition 2a (*Agentivity & DOM*):

Girona:

unspecified for agentivity & DOM > presupposed agentivity & DOM

For condition 2b (*Agentivity & no-DOM*), the predicted ranking of agentivity classes was partly confirmed for both varieties: For both Barcelona and Girona, *unspecified for agentivity* was rated significantly better than *entailed agentivity*. Furthermore, also in line with the hypothesis, in the Barcelona study *presupposed agentivity* was rated significantly better than *entailed agentivity* and in the Girona study, *unspecified for agentivity* was significantly better accepted than *potential agentivity*:

(43) Effects in condition 2b (*Agentivity & no-DOM*):

a. Barcelona:

(1) presupposed agentivity & no-DOM > entailed agentivity & no-DOM

(2) unspecified for agentivity & no-DOM > entailed agentivity & no-DOM

b. Girona:

(1) unspecified for agentivity & no-DOM > entailed agentivity & no-DOM

(2) unspecified for agentivity & no-DOM > potential agentivity & no-DOM

The results are discussed in the following section.

5.2.7 Discussion

Regarding H1 (DOM vs. no-DOM), the prediction was confirmed: In both datasets, the ratings for DOM with definite human NPs significantly differed from the ratings for no-DOM. More precisely, it was the absence of DOM that was much better accepted than the occurrence of DOM for both Barcelona and Girona. The two varieties do not vary from one another considerably.[32] Taking into account the literature on DOM in spoken Central Catalan, which reports an extension of the *a*-marker to definite human NPs (cf. Escandell-Vidal 2009: 840f.; Hualde 1992: 86f., 237f.), the clear direction of the effect towards the absence of DOM seems surprising at first. As the target sentences were presented to the participants in spoken format and they had been previously instructed to imagine them occurring in informal settings of everyday life, it was expected that the effect would either point in the opposite direction or, at least, to be less clear. In my view, the reason for the results obtained might be the strong bias of Standard Catalan, which strictly prescribes, with some exceptions, that DOM is only grammatical with tonic pronouns (GIEC 2016). Apart from that, DOM is banned from Catalan grammar and commonly referred to as a 'castillanism'. The proposed sociolinguistic explanation is supported by findings of Adli (2015): In an acceptability judgement study conducted on French question types, participants were explicitly instructed to take into account the colloquial language only and to disregard the normative perspective. Yet the results, contrasted with production data for the same question types, suggest that they were still biased by the norm in their judgements. This leads Adli to the conclusion that, apparently, "speakers cannot *not* take the normative perspective into consideration when making acceptability judgments" (2015: 189).

As for H2 (Agentivity & DOM), the predictions are partly borne out for both Barcelona and Girona. In condition 2a (*Agentivity & DOM*), the predicted ranking of agentivity classes was not confirmed. To the contrary, for both varieties *unspecified for agentivity & DOM*, which was expected to be least accepted with the marker, obtained the highest score of acceptability. With the exception of this class, for Barcelona the cline of acceptability for the remaining three verb classes mirrored the agentivity scale (44a), though none of the effects had statistical significance. For Girona, *entailed agentivity & DOM* was the second-best accepted class, as well, but the rest of the scale did not correspond to the predictions with *potential agentivity & DOM* outranking *presupposed agentivity & DOM* (44b). Furthermore, the difference between *unspecified for agentivity & DOM* and *presupposed agentivity & DOM* was statistically significant.

[32] The mean acceptability values were slightly higher for Girona for both DOM and no-DOM but the difference between both values was almost equal for both varieties. The tendency reported by Pineda (2013: 99f.) that DOM is more extended in the metropolis of Barcelona, where Catalan constantly co-exists with Spanish and is thus more influenced by the latter than in other varieties of Central Catalan, is not mirrored in the acceptability data. Supplementary production data would be needed to prove this claim. Moreover, it must be noted that I took into account the whole province of Barcelona and not only the city.

(44) Obtained scale of acceptability in condition 2a (*Agentivity & DOM*):

a. Barcelona:

unspecified for agentivity & DOM > entailed agentivity & DOM > presupposed agentivity & DOM > potential agentivity & DOM

b. Girona:

unspecified for agentivity & DOM > entailed agentivity & DOM > potential agentivity & DOM > presupposed agentivity & DOM

At first glance, these results are puzzling and challenge the assumption of an influence of agentivity on DOM, as predicted by the scale. Especially the fact that for Girona the class *unspecified for agentivity*, where no agentivity information at all is conveyed to the direct object, was significantly better accepted than *presupposed agentivity*, for which agentivity is lexically encoded, is substantial counterevidence to my claim.

However, looking at the results for condition 2b (*Agentivity & no-DOM*), the predictions were partly confirmed. In both datasets, the absence of DOM is rated best for *unspecified for agentivity* (cf. (45)) and according to the predictions this class is rated significantly better than *entailed agentivity & no-DOM*. Furthermore, *presupposed agentivity & no-DOM* outranks *entailed agentivity & no-DOM*, though this effect is significant for Barcelona only. The effect between *unspecified for agentivity & no-DOM* and *potential agentivity & no-DOM*, by contrast, only reached statistical significance for Girona.

(45) Obtained scale of acceptability in condition 2b *(Agentivity & no-DOM)*:

a. Barcelona:

unspecified for agentivity & no-DOM > presupposed agentivity & no-DOM > potential agentivity & no-DOM > entailed agentivity & no-DOM

b. Girona:

unspecified for agentivity & no-DOM > presupposed agentivity & no-DOM > entailed agentivity & no-DOM > potential agentivity & no-DOM

The fact that *entailed agentivity & no-DOM* and *unspecified for agentivity & no-DOM*, respectively, behave differently in comparison to the other classes is supported by the dispersion of the data as displayed in the boxplots (**Figure 13** and **Figure 15**) in section 5.2.6.2. Whereas for Barcelona, *entailed agentivity & no-DOM* shows a larger range of acceptability values than all other three classes including lower values from 1 to 3, for Girona it is *unspecified for agentivity & no-DOM* that differs from the other three classes in that there is no dispersion of acceptability values: they are all concentrated at the highest score, namely 5.0.

A crucial question that must be addressed is why agentivity effects only showed up in the no-DOM but not in the DOM condition. The finding that effects only manifest

when DOM is absent was also made in the above-mentioned acceptability judgement study on nominal factors for DOM in Cuban Spanish in comparison to European Spanish (Caro Reina/García García/von Heusinger 2021). While no difference of acceptability was found in the DOM condition between definite and indefinite NPs both variety-internal and across varieties (all four values reached nearly the highest possible score), the no-DOM condition showed a different picture: First, for each variety, the absence of DOM was rated clearly better with indefinite than with definite NPs. Second, comparing the no-DOM condition for both varieties, it can be seen that both sorts of object NPs were rated clearly better in Cuban Spanish than in European Spanish (Caro Reina/García García/von Heusinger 2021: 358). A general interpretation of these findings would be, on the one hand, that DOM has grammaticalized more with definite than with indefinite NPs in both varieties and, on the other hand, that DOM seems to have grammaticalized to a high extent in European Spanish while it displays more optionality in Cuban Spanish (which – in the reasoning of the paper that additionally takes into account diachronic corpus data – indicates DOM retraction in Cuban Spanish). Accordingly, in the present study, we might interpret the significant interaction effects in the no-DOM condition as showing different degrees of grammaticalization of the verb classes in question. Given the absence of DOM was rated significantly worse with *entailed agentivity* than with *unspecified for agentivity*, I conclude that DOM has grammaticalized more with the former than with the latter class in both varieties. I propose the same explanation for the significantly better ratings for *presupposed agentivity & no-DOM* vs. *entailed agentivity & no-DOM* (Barcelona) and *unspecified for agentivity & no-DOM* vs. *potential agentivity & no-DOM* (Girona). All four effects are coherent with my hypothesis that the absence of DOM is better accepted the more a verb class is ranked to the right in the agentivity hierarchy, that is the less explicit agentivity is encoded in the verbal semantics. The lack of the same effects in the reverse DOM condition is interpreted as a matter of the bias by Standard Catalan, which might have led speakers to generally reject the *a*-marker with human definite NPs.

Still, the contradictory finding for condition 2a (*Agentivity & DOM*) for Girona that *unspecified for agentivity & DOM* is rated significantly better than *presupposed agentivity & DOM* requires further discussion. First, the unexpectedly high ratings for *unspecified for agentivity & DOM* that also show up for Barcelona, where it is the best-accepted verb class in condition 2a, demand explanation. Regarding Spanish, it was observed in the literature that a good number of verbs, partly overlapping with those of the class *unspecified for agentivity*, can undergo a semantic meaning shift (cf. Delbecque 2002; García García 2005). Thus, for example, the Spanish verbs *conocer* (46a) and *encontrar* (46b) show two alternate meanings that correlate with either the absence or the occurrence of the *a*-marker:

(46) a. *conocer* ø DO ‘to identify, to know’ vs.

conocer a DO ‘to get acquainted with, to enter in contact with’

b. *encontrar ø* DO ‘to find’ vs.

encontrar a DO ‘to encounter, to meet’ (Delbecque 2002: 93f.)

As argued by García García (2005: 26f.), the two meanings can be distinguished by their degree of informativeness: While the less informative meaning generally occurs without DOM, the more informative meaning correlates with the *a*-marker. Briefly, the account assumes that whenever a speaker makes use of the alternate without DOM, it can be inferred by a *quantity and quality based implicature* (qq-implicature) that the alternative, more informative meaning is not meant (García García 2005: 23). Importantly, for my purposes, the more informative meaning of the Spanish verbs *conocer* and *encontrar* listed in (46) requires more semantic truth-conditions in that it goes along with the entailment of the proto-agent property *sentience* for the direct object. The latter reading hence involves a shifted role-semantic configuration of the predication with the result that the verbs would enter into the highest-ranked class of the agentivity scale (cf. also section 4.3.4). Concerning the corresponding Catalan verbs *conéixer* and *trobar*, the two readings of (46) are also available, though I cannot say if there is indeed a tendency that they correlate with DOM and no-DOM in the same way as in Spanish. For the study design, it was important to control for the less informative meaning where the direct object is not assigned any proto-agent property, that is ‘to know’ and ‘to find’, respectively. However, it could still be that in the DOM condition 2a, participants coerced the intended non-agentive meaning into the agentive interpretation. This can be illustrated by the following example: The sentence *El/En Víctor coneix a l'escriptora de l'escola* with the intended meaning ‘Victor knows the writer from school’ could have been interpreted as ‘Victor gets in contact with/meets the writer of the school’ (also favoured by an ambiguity concerning the scope of the prepositional phrase *de l’escola* ‘from school/of the school’). Strikingly, for Girona, this item was the overall best rated item both with and without DOM. Likewise, the item *La Maria ha trobat ø/a la nena al supermercat* with the intended meaning ‘Maria has found the girl in the supermarket’ and the – possibly available – alternative agentive meaning in the DOM condition ‘Maria has encountered the girl in the supermarket’ was among the best rated items in both conditions. An explanation operating with meaning shifts could plausibly account for the especially good ratings of *unspecified for agentivity & DOM*, on the one hand, as well as for the equally high acceptability of *unspecified for agentivity & no-DOM* on the other hand. Certainly, the presented explanation is still speculative and requires a more profound analysis.

Second, the low ratings for *presupposed agentivity & DOM*, present not only in the Girona but also in the Barcelona dataset, were contradictory to the predicted agentivity cline. In my view, they can possibly be linked to an ongoing syntactic change in Central Catalan, namely the transitivization of originally “dative”-taking predicates. As analyzed by Pineda (2013, 2015) and only briefly sketched here, in most Central Catalan varieties a set of verbs including amongst others verbs of transfer of communication (e.g. *telefonar* ‘to phone’, *escribir* ‘to write’, *respondre* ‘to answer’), verbs of transfer of possession (e.g. *pagar* ‘to pay’, *robar* ‘to steal’) and

verbs of social interaction (e.g. *pregar* 'to pray', *obeir* 'to obey', *aplaudir* 'to applaud') increasingly select a direct instead of an indirect ("dative") object. Crucially, most verbs in question have the same role-semantic properties as the verbs chosen for the class *presupposed agentivity* as they select for an object argument that bears one or more presupposed proto-agent properties. Even though the five verbs selected for the present study had proved to be syntactically transitive and not partly alternating with an intransitive encoding in Central Catalan, they might well pattern with the aforementioned originally "dative"-taking verbs. Speaking with Blume's (2000: 200f.) terminology, it could be that for verbs that presuppose proto-agent properties for their object, the unmarkedness constraints have priority over the semantic licensing condition for marked valencies. This might also hold true for differentially marked direct objects.[33] This gets further support from the finding that across varieties the class *presupposed agentivity* was accepted surprisingly well without DOM in the reverse condition 2b.

In summary, H1 (DOM vs. no-DOM) was confirmed for both varieties in that human definite NPs were significantly better accepted without DOM than with the *a*-maker. This unambiguous result was ascribed to a bias towards the norm of Standard Catalan, which generally permits the use of DOM with personal pronouns only. In my view, the normative pressure can also account for the findings for H2 where the predicted agentivity scale was not confirmed in the DOM (2a) but instead (partly) in the no-DOM (2b) condition. Especially the fact that *unspecified for agentivity & no-DOM* was significantly better accepted than *entailed agentivity & no-DOM* for both Barcelona and Girona was interpreted as evidence for my hypothesis. DOM thus seems to have grammaticalized more with verbs that lexically imply the object's agentivity than with those that do not convey any agentivity information for the direct object.

The following section presents a contrastive discussion of the results for Sicilian and Catalan.

5.3 General discussion of the results

The acceptability studies conducted partially proved the validity of the proposed agentivity scale for DOM in Sicilian and Catalan. Comparing the results obtained for the two languages, it must be again stressed that the number of independent variables was different in the two studies: While for Sicilian, three independent variables (agentivity, referentiality and DOM) were tested, it was only two in the Catalan version (DOM and agentivity). The factor referentiality, by contrast, was kept constant in the latter study in that only definite human NPs were tested.

[33] Pineda (2013: 99f.) sees a similar link between the preference for "dative" marking and DOM in a variety. The opposite to Central Catalan can be seen in Valencian where both an intransitive case frame is preserved for the verbs in question and, at the same time, the phenomenon of DOM is more widespread. Hence, one could argue that in case of Valencian the semantic licensing condition for marked valencies outranks the unmarkedness constraints.

As for condition 2a (*definite NP & DOM*) in Sicilian, it was found that both DOM and no-DOM were approximately equally acceptable whereas for H1 (DOM vs. no-DOM) in both Central Catalan dialects, no-DOM was rated significantly better than DOM. On the basis of the literature on DOM such diverging outcomes are surprising at first glance given that the *a*-marker is reported to typically occur with human definite NPs in spoken Central Catalan (cf. Escandell-Vidal 2009: 840f.; Hualde 1992: 86f., 237f.). In Sicilian, DOM is said to optionally occur with definite human NPs (cf. e.g. Iemmolo 2010b; Guardiano 2000). In my view, these differing results can be linked to the status of the two languages in question as well as to the chosen methodological tool. While Sicilian lacks standardization and thus, concerning its *ausbau* (Kloss 1967, 1987), can be qualified more as a dialect than a proper language, in the case of Catalan a highly standardized variety exists, which generally forbids the use of the *a*-marker with NPs other than personal pronouns. Taking this difference into account, I interpret the judgements of Catalan speakers as being influenced by the norm, while the ratings of Sicilian speakers underpin the full optionality of DOM, which does not seem to have grammaticalized with definite human NPs yet.

For both languages, a partial impact of agentivity on DOM was found. Yet this effect showed up in different conditions. While for Sicilian, it was detected in condition 3a (*Agentivity & definite NP & DOM*), for both Catalan dialects it was present in condition 2b (*Agentivity & no-DOM*). More specifically, in the case of Sicilian, DOM with definite NPs is significantly better accepted with verbs that lexically encode (i.e. entail or presuppose) agentivity for their direct object than with verbs for which the object's agentivity is only pragmatically conveyed (*potential agentivity*). For Catalan, effects of agentivity solely occurred in the no-DOM condition: Especially verbs that do not encode any agentivity information (*unspecified for agentivity*) are significantly better accepted without the *a*-marker than those that lexically encode agentivity in form of a verbal entailment (*entailed agentivity*).

However, there are also contradictory evidence and tendencies to the predicted agentivity scale for both Catalan and Sicilian. A verb class that particularly challenges the working hypothesis is the lowest ranked class *unspecified for agentivity*: While for Sicilian, it was the best ranked verb class in condition 3c (*Agentivity & indefinite NP & DOM*) and the least accepted in condition 3d (*Agentivity & indefinite NP & no-DOM*), for Catalan, it was the numerically best rated class in condition 2a (*Agentivity & DOM*) for both Barcelona and Girona. Especially with respect to this piece of counterevidence, one could raise the question if the general hypothesis of an impact of the agentivity scale on DOM has been already falsified. As the following arguments show, this conclusion would be too hasty: First of all, we have seen that the predictions were partially confirmed. Second, I presented secondary explanations that could plausibly account for the counterevidence. It was argued, for example, that other (language-specific) competing verbal factors might be at play that demand further study. One shortcoming of the study design is that the type of adverbial phrase was not kept constant in both of the two studies. Since for

Sicilian, it was hypothesized as an intervening factor that interacts with the explanation profile of a verb and thus possibly affects the discourse status of a direct object, future studies should keep the type of adverbial constant across verb classes. Third, there is ample evidence from other Romance languages (cf. section 4.3) that at least supports a partial impact of the communicated agentivity for the object on DOM.

From a methodological perspective, the results of the studies conducted reveal one major advantage of acceptability judgement data for the investigation of DOM, namely the comparison of two alternate forms (DOM vs. no-DOM) for each item. On the one hand, the insertion of a no-DOM condition can control for the background acceptability of a sentence and help to locate possible outliers. On the other hand, it allows us to identify subtle gradual differences of grammaticalization. More precisely, if the absence of DOM is significantly better accepted with one agentivity class than with another, it is supposed that DOM has grammaticalized more with the latter than with the former class. Crucially, as can be seen by the Catalan study, such effects need not necessarily be mirrored in the reverse DOM condition. At the same time, an agentivity effect in the DOM condition does not need to be reflected in the no-DOM condition, as shown by the Sicilian data. Instead, the lack of such corresponding effect in the no-DOM condition points to the fact that (full) grammaticalization has not yet taken place. Nevertheless, the limits of acceptability judgements need to be recognized as well. Here it must be stressed that based on the obtained results, no conclusions can be drawn with respect to the actual *use* of the *a*-marker in Sicilian and Catalan since the (dis)preference of a certain linguistic form need not correlate with its frequency in production and vice versa (cf. i.a. Adli 2015; Bader/Häussler 2010). Thus, in order to gain a more comprehensive picture of the impact of agentivity on DOM in both languages, the judgement data should be complemented with production data in a further step.

5.4 Summary

The present chapter proved the working hypothesis elaborated in section 4.4 for Sicilian and Catalan. An acceptability judgement study was carried out for each language. As for Sicilian, a variety of Western Sicilian (Alcamo) was chosen and in the case of Catalan, I selected the two Central Catalan varieties of Barcelona and Girona. One main difference in the study design was the number of independent variables: While for Sicilian, agentivity, referentiality and DOM were manipulated, it was only agentivity and DOM in the Catalan study with the referentiality of the direct object kept constant (only definite human NPs). As for the factor agentivity, the critical items that were presented orally to native speakers included five verbs for each of the four agentivity levels. The animacy of both subject and direct object was kept constant in that only personal names were used as subjects and only human NPs as direct objects. Acceptability judgements were given on a five-point scale. Whereas the Sicilian study was conducted locally in the city of Alcamo, the Catalan

study was carried out online providing a separate version each for Barcelona and Girona. For the statistical analysis, a mixed-effects regression model was used.

For Sicilian, H1 (Agentivity & DOM) and H2 (Referentiality & DOM) were rejected, in other words no interaction effect of the single independent variables agentivity and referentiality with DOM was found. H3 (Agentivity & Referentiality & DOM), by contrast, was partly confirmed: For condition 3a (*Agentivity & definite NP & DOM*), the analysis revealed that both the classes *entailed agentivity* (47a) and *presupposed agentivity* (47b) significantly outrank the class *potential agentivity*:

(47) Effects in condition 3a (*Agentivity & definite NP & DOM*):

a. entailed agentivity & definite NP & DOM > potential agentivity & definite NP & DOM

b. presupposed agentivity & definite NP & DOM > potential agentivity & definite NP & DOM

Thus, as far as definite NPs are concerned, verbs that encode agentivity in their lexical semantics either as a verbal entailment or as a presupposition are more accepted with the marker than verbs that only entail proto-patient properties to their object and convey agentivity only pragmatically.

However, no similar effect was attested in the reverse condition 3b (*Agentivity & definite NP & no-DOM*), which was attributed to the fact that DOM has not fully grammaticalized for *entailed agentivity* and *presupposed agentivity*, respectively, in combination with human definite NPs. Its absence with the two verb classes therefore does not lead to perceived ungrammaticality of a sentence. Unexpectedly, for condition 3c (*Agentivity & indefinite NP & DOM*) as well as 3d (*Agentivity & indefinite NP & no-DOM*), the obtained effects and scalar rakings of verb classes instead point to the reverse direction, though not reaching statistical significance: In condition 3c, *entailed agentivity & indefinite NP & DOM* is rated worst while *unspecified for agentivity & indefinite NP & DOM* gets the best acceptability rate. Condition 3d provides exactly the opposite picture: *entailed agentivity & indefinite NP & no-DOM* is accepted best whereas *unspecified for agentivity & indefinite NP & no-DOM* gets the lowest acceptability score. It has been argued that a possible explanation for the contradictory results of condition 3a vs. 3c (and 3d) could be found in competing verbal factors that might have an impact on DOM with indefinite but not with definite NPs in Sicilian.

For Catalan, for both varieties H1 (DOM vs. no-DOM) was confirmed in that the absence of DOM was clearly better accepted than DOM. The highly significant effect was attributed to a bias by Standard Catalan, which generally bans DOM from object NPs other than personal pronouns. H2 (Agentivity & DOM) was partly confirmed for both varieties. The predictions for condition 2a (*Agentivity & DOM*) were entirely rejected and even given counterevidence for Girona where

unspecified for agentivity outranks *presupposed agentivity* (48), which calls into question the claimed agentivity cline.

(48) Effect in condition 2a (*Agentivity & DOM*):

Girona:

unspecified for agentivity & DOM > presupposed agentivity & DOM

The predictions for condition 2b (*Agentivity & no-DOM*), in contrast, were partly confirmed. For both varieties, *unspecified for agentivity & no-DOM* got significantly higher acceptability scores than *entailed agentivity & no-DOM*. Moreover, for Barcelona *presupposed agentivity & no-DOM* outranked *entailed agentivity & no-DOM*, while for Girona *unspecified for agentivity & no-DOM* was better accepted than *potential agentivity & no-DOM* (cf. (49)).

(49) Effects in condition 2b (*Agentivity & no-DOM*):

a. Barcelona:

(1) presupposed agentivity & no-DOM > entailed agentivity & no-DOM

(2) unspecified for agentivity & no-DOM > entailed agentivity & no-DOM

b. Girona:

(1) unspecified for agentivity & no-DOM > entailed agentivity & no-DOM

(2) unspecified for agentivity & no-DOM > potential agentivity & no-DOM

These results especially show a contrast between the outer edges of the predicted agentivity scale: The lack of DOM is hence more accepted with verbs that do not lexically assign or pragmatically convey any kind of proto-agent properties to their direct object, while it is less accepted for verbs that encode agentivity, more specifically *sentience*, as a verbal entailment.

In summary, while the type of obtained effects distinctly differed for Sicilian and Catalan, both studies provide evidence for an impact of agentivity on DOM. In particular verbs that encode agentivity in their lexical semantics as a verbal entailment systematically interacted with DOM. While for Sicilian, *entailed agentivity* correlated with the presence of DOM, for both Catalan varieties the absence of DOM with this verb class was significantly less accepted than with lower-ranked verb classes in the agentivity scale. Yet, note again that the agentivity effect for Sicilian was only measured in interaction with definite object NPs, which suggests that it is not the object's agentivity alone but its interplay with a highly prominent object in terms of referentiality that is decisive for DOM. Based on these findings, I will discuss the implementation of agentivity as a factor for DOM in Romance in detail.

6. Role of agentivity for DOM in Romance

This chapter aims to discuss the role of agentivity for DOM in Romance in a broader context. In chapter 4, I proposed a scalar account on the direct object's agentivity. The proposed scale, repeated in (1) for convenience, was confined to the type of communicated agentivity for the object.

(1) entailed agentivity for O > presupposed agentivity for O > potential agentivity for O > O unspecified for agentivity

The detailed subdivision of agentivity levels allowed me to systematically test the effect of (verb-dependent relational) agentivity on DOM in Sicilian and Catalan and to dissociate it from effects resulting from (object NP-inherent) animacy. As argued in section 4.3, the scale in (1) can be partially cross-classified with the agent-patient asymmetry and in this way include the thematic relation between subject and direct object. This latter type of agentivity was shown to have an impact on DOM with inanimate objects in Spanish by García García (2007, 2014).

Section 6.1 discusses agentivity as a factor for DOM in Romance. Starting from the results of the acceptability judgement studies on DOM with human direct objects in Sicilian and Catalan (6.1.1), I relate the findings to DOM with inanimate objects in Spanish and to its role-semantic explanation given by García García (2007, 2014) (6.1.2). Furthermore, both diachronic and synchronic evidence will be provided from other Romance languages and varieties further supporting an agentivity-based approach (6.1.3). Section 6.2 points to the fact, stressed throughout this work, that DOM in Romance is a phenomenon triggered by multiple factors: besides the role-semantic factor agentivity, these are the nominal factors animacy (6.2.1) and definiteness (6.2.2) as well as the information-structural factor topicality. Here it is of special interest to dissociate animacy from agentivity. A possible way that these factors can interact in grammaticalization is sketched in 6.2.3. Section 6.3 summarizes the main points of the discussion.

6.1 Agentivity as a factor for DOM in Romance

This subchapter discusses what the findings for Sicilian and Catalan can contribute to the discussion of the impact of agentivity on DOM in Romance. In 6.1.1, I will further elaborate on the acceptability judgement studies conducted for the two languages and also show the limits of the obtained empirical evidence. Section 6.1.2 relates the results on Sicilian and Catalan to the empirical data presented for *a*-marked inanimate objects in Spanish (García García 2007, 2014) and argues that both findings can be explained within the same role-semantic framework. Evidence provided from other Romance languages and varieties where DOM is (still) a more marginal phenomenon in section 6.1.3 supports my view: the verb classes showing a preference for DOM in these languages are ranked high in the agentivity scale, i.e. they entail or presuppose the object's agentivity.

6.1.1 Sicilian and Catalan

The acceptability judgement studies for Sicilian and Catalan were the first to systematically test the impact of communicated agentivity for the direct object on DOM. Importantly, agentivity levels were defined independently of animacy. The animacy of the direct object NP was kept constant across conditions in that only human NPs were tested. The results reveal at least a partial impact of agentivity on DOM. However, the measured effects were quite different for the two languages under investigation. For Sicilian, triple interaction effects were found in condition 3a (*Agentivity & definite NP & DOM*) but no interaction effects of agentivity and DOM (conditions 1a and 1b) or referentiality and DOM (conditions 2a and 2b), respectively. For Catalan, for which only definite NPs were tested, effects only occured in condition 2b (*Agentivity & no-DOM*).

It was shown that in Sicilian, predicates that entail agentivity – more precisely *sentience* – for their direct object (e.g. the OE-psych verbs *affènniri* 'to offend' or *nuciri* 'to bother') as well as verbs that presuppose agentivity for their object (e.g. *sentiri* 'to hear' or *sarvari* 'to save') were significantly better accepted with DOM than verbs selecting a potentially agentive direct object (e.g. *curari* 'to cure', *firiri* 'to injure'), as far as definite object NPs are concerned. No difference was found, by contrast, between the two verb classes *entailed agentivity* and *presupposed agentivity*. It can be thus concluded that for definite object NPs, verbs that lexically encode agentivity information for their direct object are better accepted with DOM in Sicilian than those for which agentivity (*sentience*) for the object arises only via conversational implicature.

In Catalan, the most robust finding – since it was the only one that showed up in both samples of Barcelona and Girona – was that the absence of DOM was significantly better rated for *unspecified for agentivity* than for *entailed agentivity*. Thus, predicates that do not convey any agentivity information for the object (e.g. *conéixer* 'to know', *veure* 'to see') are more unambiguously accepted without the *a*-marker, while there was more variation when agentivity was lexically encoded as a verbal entailment (e.g. *ofendre* 'to offend', *molestar* 'to bother'). Moreover, for each dataset there was another effect in condition 2b supporting the predicted agentivity cline:

(1) Effects in condition 2b (*Agentivity & no-DOM*):

  a. Barcelona:

    (1) presupposed agentivity & no-DOM > entailed agentivity & no-DOM

    (2) unspecified for agentivity & no-DOM > entailed agentivity & no-DOM

  b. Girona:

    (1) unspecified for agentivity & no-DOM > entailed agentivity & no-DOM

    (2) unspecified for agentivity & no-DOM > potential agentivity & no-DOM

Yet for both languages, counterevidence to the predicted gradual impact of agentivity on DOM was also found. For Sicilian, the obtained scales of acceptability in condition 3c (*Agentivity & indefinite NP & DOM*) and 3d (*Agentivity & indefinite NP & no-DOM*) correspond exactly to the opposite of what was predicted: The verb class *unspecified for agentivity* got particularly high acceptability values with indefinite NPs in the DOM condition, while the verb class *entailed agentivity* showed the lowest ratings in the same condition. As a mirror image, the class *unspecified for agentivity* was rated lowest in the reverse no-DOM condition 3d while *entailed agentivity* got the highest ratings. Note that none of the reported differences reaches statistical significance. Given the effect shown for *entailed* and *presupposed agentivity* vs. *potential agentivity* in condition 3a (*Agentivity & definite NP & DOM*), a similar tendency was expected in condition 3c (*Agentivity & indefinite NP & DOM*). A possible explanation for the contrary findings could be that the verbal factors determining DOM differ for definite and indefinite NPs. While agentivity demonstrably has an impact on DOM in interaction with referentially prominent (i.e. definite) NPs, it seems to be less relevant in interaction with indefinite NPs. Instead, I hypothesized that another verbal factor, namely implicit causality – or more generally the explanation type triggered by a verb – might be of relevance. This factor is located at the interface between lexical semantics and discourse structure and has to do with expectations about upcoming discourse that can be deduced by the lexical meaning of a predicate (cf. e.g. Bott/Solstad 2014, 2017). Therefore, although participants were exposed to isolated sentences, the predicate in question – sometimes in combination with the chosen adverbial phrase – evokes a certain explanation profile. For OE-psych verbs, which display an IC bias for the subject stimulus, this bias was pre-empted by the insertion of a manner or causal adverbial (with the exception of *nuciri* 'to bother'). So perhaps, in an out-of-the-blue context, such sentences are conceived of as more "complete" than they were without the simple cause given by the adverbial.[1] This might lessen the need for the arguments to be taken up in the subsequent discourse. One consequence could be a higher acceptability for the absence of DOM since from a discourse-contextual point of view the object is of low prominence.

In Catalan, contrary to the predictions for condition 2a (*Agentivity & DOM*), the presence of DOM was rated significantly higher with the verb class *unspecified for agentivity* than with *presupposed agentivity* in the Girona dataset. I have argued that the high acceptability for *unspecified for agentivity & DOM*, also attested for Barcelona, may be due to a meaning shift of the predicates in question (such as *conéixer* 'to know' or *trobar* 'to find') to a symmetrical reading in which the direct object is entailed *sentience* (i.e. *conéixer* 'to get acquainted with, enter in contact with', *trobar* 'to meet, encounter'). While such a shift would arise for *a*-marked objects, the intended (non-agentive) reading is unambiguously conveyed for non-marked objects. Note that this would match the finding that *unspecified for agentivity* is at the same

[1] As shown by Solstad/Bott (2017: 25), when no simple cause is provided for an OE-psych verb (e.g. *Mary fascinated John…*), 51% of continuations are explanations. When the IC bias for the subject is pre-empted by the insertion of a prepositional phrase (e.g. *Mary fascinated John with her frequent contradictions*), by contrast, explanations follow only in 19% of the cases.

time the best-rated class for both varieties in condition 2b (*Agentivity & no-DOM*). If such an attempt at an explanation, derived from Spanish (Delbecque 2002; García García 2005) proves true for Catalan, it would be further support for the agentivity hypothesis. Regarding the unexpectedly low ratings for *presupposed agentivity & DOM*, which also show up in the Barcelona dataset and are furthermore mirrored for both varieties by high ratings in the reverse condition 2b (*Agentivity & no-DOM*), there are some clues that they may be attributed to a more general syntactic change in Central Catalan. As was put forward by Pineda (2013, 2015), certain verb classes including verbs of transfer of communication or possession as well as verbs of social interaction do not select for an indirect ("dative") object anymore but instead undergo a process of transitivization. Crucially, these verbs typically show the same role-semantic properties as the verbs in the class *presupposed agentivity*. Therefore, I interpret the measured effects as a manifestation of the unmarkedness constraints (cf. Blume 2000: 200f.) for verbs which presuppose an agentive object participant, be it *a*-marking of an indirect or direct object.

Beyond verb semantics, my studies offer more general insights into the acceptability of DOM and no-DOM with human definite (and indefinite) NPs. Surprisingly at first, the Sicilian data seem to suggest full optionality of the *a*-marker for both definite NPs and indefinite NPs which obtained approximately equal ratings in the DOM (2a) and no-DOM (2b) condition. On the one hand, this broadly confirms what has been reported in the literature on the optionality of DOM with definite human NPs (cf. Iemmolo 2010b, Guardiano 2000). On the other hand, the fact that ratings for indefinite human NPs differ neither between DOM and no-DOM condition nor in comparison to definite NPs are a noteworthy finding that has never been shown for another Romance language. Rather, it is expected that at least in the no-DOM condition, the ratings for definite and indefinite human NPs mirror the referentiality scale, as has been shown for both European and Cuban Spanish (Caro Reina/García García/von Heusinger 2021: 358). In a nutshell, the general picture suggests that DOM has not grammaticalized for definite and indefinite human NPs in Sicilian but still functions as an optional means. This finding must be arguably linked to its status as a regional language or dialect that lacks standardization.

In Catalan, the results show for both samples of Barcelona and Girona that the absence of DOM is significantly better accepted than its presence. This is notable for the following reason: While DOM is only allowed with personal pronouns following prescriptive grammar of Standard Catalan (GIEC 2016), the spoken language provides a different picture. Escandell-Vidal (2009: 841) reports for the variety of Central Catalan that the marker is even obligatory with proper names as well as with definite human NPs. As for the latter, my data seem contradictory to this claim. Again, participants were exposed to spoken sentences intended to create a rather informal setting – the kind of setting in which DOM in Central Catalan is typically widespread. If DOM was compulsory with human definite NPs we would expect that the acceptability of DOM outranks the acceptability of no-DOM, which is not what was found. Instead, the data reflect that the no-DOM variant is preferred over the

DOM variant for nearly every item.[2] This finding was interpreted to be due to the prescriptive force of Standard Catalan that biases speakers' judgements, also in colloquial settings. Therefore, the acceptability ratings obtained should be clearly differentiated from generalizations made on DOM in spoken Central Catalan based on production data. Moreover, it can be argued that the measured effects for both samples in the no-DOM condition may be more reliable agentivity-related predictors than the results in the DOM condition: while the latter mirror a general rejection of the *a*-marker, the former display subtle differences between agentivity classes. Thus, for example, while there is large uniformity in accepting *unspecified for agentivity* in the no-DOM condition, there is more dissent for *entailed agentivity* in the same condition, which can be ascribed to the lack of DOM.

In summary, the acceptability judgement studies on Sicilian and Catalan provided evidence that the communicated agentivity for the direct object is a factor that has to be taken into account for DOM. It is thus suggested that proto-agent properties of the direct object may affect morphosyntactic marking even beyond an actual lack of thematic distinctness between subject and object. Note that for the sentences of the studies only role-semantic configurations were considered where argument selection is unambiguously performed. A deviation from the prototypical transitive scenario thus only consists in the type of agentive properties (mainly *sentience*) conveyed for the direct object. Certainly, further research must identify and investigate (language-specific) competing or interacting verbal factors. Since the study on Sicilian suggests that agentivity is only a relevant factor for DOM in interaction with definite but not with indefinite object NPs, the interaction between agentivity and referentiality deserves closer examination and should be considered for further studies on Catalan DOM as well.

6.1.2 Spanish

As the role-semantic approach on DOM has been developed based on *a*-marked inanimate objects in Spanish (García García 2007, 2014), my findings on Sicilian and Catalan have to be related to Spanish in a next step in order to see if they are explainable within the same account.

As a short recap, the role-semantic approach on DOM in Spanish has been developed based on systematic occurrences of DOM with inanimate objects as exemplified in (2).

(2) a. *El carnaval precede* ****ø/a*** *la Cuaresma.*
the carnival precede.PRS.3SG ø/DOM the Lent
'Carnival precedes Lent.' (García García 2014: 2)

[2] Exceptions are the verbs *consolar* 'to console', *insultar* 'to insult' and *matar* 'to kill' for the variety of Girona which are better accepted with the *a*-marker than without it.

b. *La euforia caracteriza* ***??ø/a*** *la situación.*
the euphoria characterize.PRS.3SG ø/DOM the situation
'Euphoria characterizes the situation.' (García García 2018a: 228)

In (2a), *preceder* 'to precede' is used as a reversible predicate. Both the subject *carnaval* 'carnival' and the direct object *Cuaresma* 'Lent' do not bear any involvement properties. Due to the lack of co-argument dependency, they can both be classified as weak proto-agents. This deviates from the prototypical transitive scenario in which the subject clearly outranks the object in terms of agentivity. In (2b), where *caracterizar* 'characterize' functions as a verb of attribution, we can detect an even greater deviation from the prototypical agent-patient asymmetry: While here, we also do not have a co-argument dependency relation between subject and direct object, with difference to (2a), the subject *la euphoria* 'euphoria' does not even have argument status. In both example cases, DOM functions as a means of disambiguation being required when subject and direct object are not sufficiently distinguishable in terms of their semantic roles, as captured in the generalization of thematic distinctness:

(3) Generalization of *thematic distinctness* (García García 2007: 71, 2014: 145, 2018a: 227):

> DOM in Spanish is required with inanimate objects when the subject does not outrank the direct object in terms of agentivity.

In order to make predictions for DOM with animate objects, I proposed making use of Blume's (2000) refined proto-role role model which has been elaborated to account for marked valencies (such as */nom/dat* case frames) in particular. As in García García's approach on DOM (2007, 2014), Blume (2000: 200f.) sees the non-asymmetric distribution of proto-roles as the decisive criterion for marked valencies to be licensed. For my purposes, there are two main advantages of Blume's proto-role model: First, it differentiates between potent and non-potent proto-agent properties. Whereas potent proto-agent properties, such as *volition* and *autonomous activity*, are linked to the nominative per default in accusative languages, non-potent proto-agent properties, such as *sentience*, can be linked to either case. In object position, the latter can overlap with entailed proto-patient properties (e.g. for the OE-psych verb 'frighten' in its causative reading). This last-mentioned role-semantic scenario is called *pseudo-symmetric* distribution of proto-agent properties (Blume 2000: 217) and might have an impact on formal marking: In a language in which the licensing condition for marked valencies is overgeneralized, verbs that entail proto-patient properties for the object, overlapping with non-potent proto-agent properties, can also get a marked case frame. Second, Blume not only considers *entailed* but also *presupposed* proto-agent properties relevant for the selection of marked valencies. Although the argument of the first implied subevent is linked to the nominative per default (Blume 1998: 269, 2000: 224), presupposed proto-agent properties for the object cross-linguistically matter for the selection of marked valencies, such as /*nom/dat* (e.g. for the interaction verb *helfen* 'to help' in German). Hence, the model

shows how (presupposed) proto-agent properties that are not decisive for nominative selection can influence the morphosyntactic marking of the object.

In section 3.3 and 4.1, I put forward two main points that, in my view, have to be defined more clearly in order to systematically investigate DOM with animate objects in Spanish – as well as in other languages – in an agentivity-based account. As I will argue later, these two points can be addressed by taking into account the above-sketched innovations of Blumes' (2000) proto-role model. The first concerns the more precise formulation of criteria for *potential agentivity*. I argued that the binary distinction between subcategorized and potential proto-agent properties made by Primus (2012a: 81) is not fine-grained enough. Instead, more sub-steps have to be included that should concern the thematic relation between the direct object and the predicate. An example can be given by the differentiation of verbs of auditory and visual perception. As von Heusinger/Kaiser (2011: 614) show, especially with indefinite human direct objects, the Spanish auditory perception verbs *escuchar* 'to listen' and *oír* 'to hear' already show a high preference for DOM in the 15th century. The visual perception verbs *mirar* 'to look at' and *ver* 'to see', by contrast, are considerably less attested with the marker. This different preference for DOM of the two types of perception verbs is supposedly due to the agentivity of the direct object, as formulated by García García (2018a: 227):

> While the verbs of auditory perception presuppose a noise-producing source as their object argument, i.e. a physically active and thus agent-like participant, the object argument of visual perception verbs need not be an agentive participant (Enghels 2007: 244–273).

For this reason, Blume (1998: 273) classifies 'listen to someone' in general (realized e.g. as */nom/dat zuhören* 'to listen to' or *lauschen* 'to listen to, eavesdrop' in German) as "an interaction verb rather than a perception verb, since it presupposes that the dative participant (protoagent property 'autonomous activity') produces an aesthetic or meaningful chain of sounds (si) and it implies a conscious and controlled act of perception by the nominative participant (proto-agent of s2)". Furthermore, the integration of presupposed proto-agent properties in linking theory could account for the fact that the Spanish auditory perception verb *oír* 'to hear' is found considerably more often with an indirect object clitic than the visual perception verb *ver* 'to see' (Roegiest 2003: 304).

The second is related to the more general prediction that the lack of the canonical agent-patient asymmetry between subject and direct object triggers DOM in Spanish (García García 2007: 71, 2014: 145). As I argue, besides the relative distribution of proto-properties between subject and object, the communicated agentivity for the direct object should also be considered. In this way, we could for example account for verbs of competition which optionally though systematically take DOM with inanimate objects in Spanish (García García 2014: 178-187), although they do imply an agent-patient asymmetry. Hence, in these cases, the subject clearly outranks the object in terms of agentivity and the generalization of thematic distinctness in the strict sense is not met.

(4) *La curiosidad vence* ***ø/a*** *la pereza.*

the curiosity overcome.PRS.3SG ø/DOM the laziness

'Curiosity overcomes laziness.' (García García 2014: 2)

In (4), the direct object *la pereza* 'laziness' is causally affected by the subject *la curiosidad* 'curiosity' and therefore stands in a causal dependency relation to the subject argument. Thus, in contrast to the examples in (2) above, the subject clearly outranks the direct object in terms of agentivity and linking can unambiguously take place. Yet, in addition to the proto-patient properties, the object also bears a presupposed proto-agent property: Competition verbs such as *vencer* 'win' presuppose that the object participant is autonomously active in a previous subevent. This presupposition can be proved by its constancy under negation: The sentence *La curiosidad no vence ø/a la pereza* 'Curiosity does not overcome laziness' implicitly states that curiosity and laziness were in competition with each other in a previous subevent. The general optionality of DOM with verbs of competition in contrast to the requirement of the marker with reversible predicates and verbs of attribution can be explained by the degree of deviation from the canonical agent-patient asymmetry: While *a*-marking is obligatory when the subject does not outrank the direct object in terms of agentivity, it is only optional when a canonical agent-patient asymmetry is indeed present but proto-patient properties of the object overlap with presupposed proto-agent properties.[3]

Briefly, by adding the communicated agentivity for direct object as a further role-semantic dimension to the thematic distinctness between subject and direct object, and thus combining García García's (2007, 2014) approach with Blume's (2000) modification of the proto-role model, a more fine-grained analysis of both *a*-marked animate and inanimate objects in Spanish can be achieved.

Beyond this, the established agentivity scale could help to systematize individual verbs that are recurrently mentioned in the literature for obligatorily also taking DOM with human indefinite objects in Spanish. As we know, unlike in Sicilian and Catalan, DOM has reached a high degree of grammaticalization in European Spanish so that differences in *a*-marking among verbs and verb classes are hard to detect as far as definite human NPs are concerned. Taking indefinite human NPs into consideration, subtle differences between individual verbs are expected to show up more clearly. Table 1 provides a collection of verbs for which DOM is reported to be compulsory with human indefinites in Spanish:

Verb	**References**
aborrecer 'to detest'	Leonetti (2004: 99)
acosar 'to harass'	Fábregas (2013: 28)
acusar 'to accuse'	Torrego (1998: 39), Fábregas (2013: 27)

[3] Note, however, that within the class of verbs of competition, there is considerable variation to take DOM with inanimate objects (cf. García García 2014: 178ff.).

admirar 'to admire'	Torrego (1998: 39), Fábregas (2013: 28), Romero Heredero/García García (2023: 275)[4]
afectar 'to affect'	Pensado (1995b: 33), Fábregas (2013: 28)
amar 'to love'	Leonetti (2004: 99), Romero Heredero/García García (2023: 275)
atacar 'to attack'	Torrego (1998: 30), Leonetti (2004: 84)
ayudar 'to help'	Fábregas (2013: 28)
castigar 'to punish'	Torrego (1998: 39), Leonetti (2004: 84), Fábregas (2013: 27)
defender 'to defend'	Torrego (1998: 30)
despreciar 'to despise'	Leonetti (2004: 99)
emborrachar 'to make drunk'	Torrego (1998: 30)
empapar 'to soak'	Torrego (1998: 30)
empujar 'to push'	Torrego (1998: 39)
encarcelar 'to jail'	Torrego (1998: 17)
entrevistar 'to interview'	Fábregas (2013: 28)
golpear 'to beat'	Torrego (1998: 39), Fábregas (2013: 27)
insultar 'to insult'	Torrego (1998: 30), Leonetti (2004: 84), Fábregas (2013: 27)
odiar 'to hate'	Torrego (1998: 39), Leonetti (2004: 84), Fábregas (2013: 27)
ofender 'to offend'	Torrego (1998: 39), Fábregas (2013: 27)
oír 'to hear'	Romero Heredero/García García (2023: 276)
operar 'to operate'	Torrego (1998: 23)
pervertir 'to pervert'	Torrego (1998: 30)
respetar 'to respect'	Torrego (1998: 39)
saludar 'to greet'	Pensado (1995b: 33), Leonetti (2004: 84), Fábregas (2013: 28)
sobornar 'to bribe'	Torrego (1998: 30), Leonetti (2004: 84), Fábregas (2013: 28)
soportar 'to put up with'	Leonetti (2004: 99)

Table 1. A collection of verbs that are recurrently reported to obligatorily take DOM with human indefinite direct objects in Spanish

Classifying the individual verbs along the agentivity hierarchy, as illustrated in Table 2, we see that they spread nearly along all agentivity classes. However, adding the

[4] Unlike the other references listed in Table 1, the indicated verbs of Romero Heredero/García García (2023: 275f.) stem from a forced-choice experiment. In the study, participants were given incomplete sentences lacking a direct object. For each test sentence, they had to choose between a variant with and without DOM while the objects were all human indefinites (e.g. *Valeria admiró (a un cantante / un cantante) durante toda su vida.* 'Valeria admired a singer all her life', cf. Romero Heredero/García García 2023: 285). As the results for 326 native speakers of European Spanish show, *admirar* 'to admire' was chosen with DOM in 88% of the cases, while it was 97% for *amar* 'to love' and 87% for *oír* 'to hear'.

verbs that are reported to take DOM only optionally[5] (or reject it)[6] with indefinite human NPs, a majority of verbs belongs to the lowest ranked class *unspecified for agentivity*.

AGENTIVITY INFORMATION OF O → ASYMMETRY ↓	entailed	> presupposed	> potential	> unspecified
S *less* agentive than O		-	-	-
S as agentive as O		*ayudar, entrevistar*, *oír*, *respetar*, *saludar*, *soportar*	-	-
S *more* agentive than O	*afectar*, *ofender*	*acosar*, *defender*, *castigar*, *encarcelar*, *pervertir*, *sobornar* [vs. ± DOM: *contratar*]	*acusar*, *atacar*, *emborrachar*, *empapar*, *empujar*, *insultar*, *golpear*, *operar* [vs. ± DOM: *besar*, *curar*, *examinar*, *matar*]	*aborrecer*, *admirar*, *amar*, *despreciar*, *odiar* [vs. ± DOM: *buscar*, *conocer*, *describir*, *encontrar*, *esconder*, *llevar*, *ver*] [vs. - DOM: *haber*, *tener*]

Table 2. Individual verbs that are recurrently reported to obligatorily take DOM with human indefinite direct objects in Spanish, categorized along the agentivity hierarchy (in square brackets verbs that optionally take the marker or reject it)

[5] DOM is said to be optional with human indefinite NPs for the following verbs:

Verb	**References**
besar 'to kiss'	Torrego (1998: 30)
buscar 'to look for'	Leonetti (2004: 84), Torrego (1998: 39)
conocer 'to know'	Leonetti (2004: 85), Torrego (1998: 39)
contratar 'to hire'	Leonetti (2004: 99)
curar 'to cure'	Leonetti (2004: 99)
describir 'to describe'	Leonetti (2004: 99)
encontrar 'to find'	Leonetti (2004: 99), Torrego (1998: 39)
esconder 'to hide'	Leonetti (2004: 84), Torrego (1998: 17)
examinar 'to examine'	Torrego (1998: 29f.)
llevar 'to take'	Leonetti (2004: 99)
matar 'to kill'	Torrego (1998: 30, fn. 17)
ver 'to see'	Leonetti (2004: 99), Torrego (1998: 39)

[6] DOM is excluded with human indefinite NPs for the following verbs:

Verb	**References**
haber 'there is/are'	Leonetti (2004: 81)
tener 'to have'	Leonetti (2004: 81)

For the verbs that take DOM obligatorily also with human non-specific indefinite NPs, Fábregas (2013: 27) gives the following examples for which I indicate the agentivity level in square brackets:

(5) a. *ofender* **(a)* *una* *persona* *cualquiera* [entailed agentivity]
offend DOM a person whatsoever

b. *golpear* **(a)* *un estudiante* *cualquiera* [potential agentivity]
beat DOM a person whatsoever

c. *afectar* **(a)* *una* *persona* *cualquiera* [entailed agentivity]
affect DOM a person whatsoever

d. *ayudar* **(a)* *una* *persona* *cualquiera* [presupposed agentivity]
help DOM a person whatsoever

As can be seen, the verbs in (5) are assigned different verb classes in the hierarchy in Table 2: while *ofender* 'to offend' (5a) and *afectar* 'to affect' (5c) entail sentience for its object, *ayudar* 'to help' (5d) presupposes an agentive object and *golpear* 'to beat, hit' (5b) conveys potential agentivity (sentience) for the object via generalized conversational implicature. Despite the different degrees of communicated agentivity for the object, the acceptability of DOM with human indefinite objects does not differ for these verbs – it is always obligatory. This may be due to the fact that DOM with human indefinite objects has already highly grammaticalized along the agentivity scale in Spanish. Such an assumption fits the observation that most verbs where DOM with indefinite human NPs is optional, as well as the ones where it is excluded, belong to the lowest ranked class *unspecified for agentivity*.

There are further indications that the communicated agentivity for the object should be considered as a factor for DOM in Spanish, too. Let us start with the counter hypothesis advanced by Torrego (1998: 27ff.; 1999: 1785f.) stating that it is the agentivity of the subject that affects *a*-marking in Spanish. On the one hand, she sees her hypothesis confirmed with verbs that can take both agentive and non-agentive subjects like *exigir* 'to demand, require' (6). With respect to *a*-marking, it is only the former (6a) but not the latter that allow it (6b). On the other hand, this may be illustrated by existential verbs like *haber* 'there is/are', which cannot take an agentive subject and reject *a*-marking (7).

(6) a. *El herido exigía* ***(a)*** *un médico.*
the injured require.PST.3SG DOM a doctor
'The injured demanded a doctor.'

b. *La situación exigía* ***(*a)*** *un médico.*
the situation require.PST.3SG DOM a doctor

‘The situation required a doctor.’ (Torrego 1998: 29; Fábregas 2013: 26f.)

(7) *Hay* ***(*a)*** *unas delegadas en la sala.*
there are.PRS DOM some delegates in the room
‘There are some delegates in the room.’ (Torrego 1999: 1785)

In (6a), the human subject *el herido* ‘the injured’ is interpreted as acting volitionally and being sentient relative to the event, while the direct object *un medico* ‘a doctor’ does not bear any proto-property. As indicated, DOM is optional in this case. In (6b), which shows the same predicate with the inanimate subject *la situación* ‘the situation’, an agentive reading cannot be obtained and DOM would be ungrammatical. Note that due to the lack of proto-properties for both co-arguments, *exigir* ‘to demand, require’ in its reading in (6b) bears an especially low degree of semantic transitivity. With respect to existential *haber* ‘there is/are’ in (7), Delbecque (2002: 107) criticizes that “by invoking the absence of an *agent* and the *non-specificity* of the DO entity (Torrego 1999: 1785, 1795), one tackles corollaries of the presentative and quantifying function of this predicate, but risks missing the point, viz. that the basic requirement of having two participants is not met”. It can even be questioned whether the direct object has argument status: while pronominalization by a direct object clitic is possible (*Las hay en la sala* ‘They are in the room’), passivization is barely acceptable (*?Unas delegadas son habido en la sala* ‘There are delegates in the room’). This is not surprising as we already deal with an impersonal verb that does not bear an agent that could be demoted in a passive sentence. As for its semantic transitivity, it must be noted that the object of *haber* does not bear any proto-properties. This distinguishes it from the direct object of the attribution verb *caracterizar* ‘to characterize’ (e.g. *La euforia caracteriza* ***??ø/a*** *la situación* ‘Euphoria characterizes the situation’). Just as with *haber*, the subject does not have argument status. However, unlike *haber*, the object of *caracterizar* is causally independent and functions as a weak proto-agent. The object hence outranks the subject in terms of agentivity, which is morphosyntactically reflected in the strong tendency to take DOM. This suggests that not the agentivity of the subject alone, but the role-semantic configuration of both co-arguments has to be taken into account in order to understand the presence or absence of DOM.

Counterevidence to Torrego’s (1998, 1999) claim also comes from non-causative OE-psych verbs such as *fascinar* ‘to fascinate’ or *interesar* ‘to interest’. Since these verbs do not assign potent proto-properties to either argument, they are of restricted semantic transitivity. However, unlike *haber* ‘there is/are’ and *exigir* ‘to demand, require’, a non-potent proto-agent property, namely *sentience*, is assigned to the direct object. As can be seen from the following examples, the non-causative psych verbs *fascinar* ‘to fascinate’ and *interesar* ‘to interest’ require DOM, regardless of the subject’s animacy:

(8) a. *María fascina/interesa* ****ø/a*** *muchos críticos culturales*

PN fascinate/interest.PRS.3SG ø/DOM many critics cultural

(con su obra).

(with her work)

'Maria fascinates/interests many cultural critics (with her work).'

b. *La obra fascina/interesa* ****ø/a*** *muchos críticos culturales.*

the work fascinate/interest.PRS.3SG ø/DOM many critics cultural

'The work fascinates/interests many cultural critics.'

The role-semantic configuration is the same in both cases: the direct object, exhibiting the proto-agent property of sentience outranks the subject, which does not bear any involvement property.[7] We thus deal with a transitive scenario that sharply deviates from the prototypical agent-patient asymmetry. As has been seen, the agentivity of the subject alone is not sufficient to predict *a*-marking of the direct object. Rather, the agentivity of the direct object has to be taken into account as well.

To sum up, a role-semantic account that considers the communicated agentivity for the direct object may provide an explanation for the special affinity for DOM of certain verbs in Spanish, also with human indefinite NPs. Moreover, as has been argued, the agentivity of the subject alone is no reliable predictor for the occurrence of DOM across predicates and predicate classes. Instead, the distribution of proto-properties for both co-arguments, which in particular also considers proto-agent properties of the object, can offer a more sophisticated understanding of *a*-marking patterns.

Yet there is one subclass of verbs, classified as *unspecified for agentivity*, which tends to always take DOM with indefinite human NPs and therefore contradicts my hypothesis: these are SE-psych verbs such as *odiar* 'to hate', *admirar* 'to admire', *aborrecer* 'to hate', *despreciar* 'to despise' and *amar* 'to love'. These verbs take a subject bearing the non-potent proto-agent property *sentience* and a direct object that does not exhibit any proto-property.[8] SE-psych verbs show yet another peculiarity in that they display an implicit causality (IC) bias for the object (cf. e.g. Kehler et al. 2008).[9] Put briefly, this is meant to say that their stimulus object is per default underspecified for a simple cause (Bott/Solstad 2014). To my knowledge, the IC bias of a verb has not been taken into account on studies of DOM so far. For Sicilian, data from my acceptability study have shown that the SE-psych verb *'un putiri viriri* 'to hate' displays notably high ratings with an indefinite object NP in the DOM condition, which

[7] Note that in the proto-role model of Dowty (1991), the subject as well as the direct object would bear the proto-agent property *independent existence* here.

[8] For a different view, see Dowty (1991: 579f.) who ascribes the proto-agent property causation to the stimulus object in this case. As argued in section 3.1, though, Dowty's role-semantic analysis of psych verbs is highly problematic.

[9] SE-psych verbs are special in yet another respect: they can be referred to as *non-incorporating* since they do not allow for bare plurals as direct objects, e.g. {**odiar/*admirar/*amar*} *personas* 'to hate/admire/love persons' (cf. Leonetti 2004: 99). For a profound analysis of the restrictions regarding the direct object of SE-psych verbs in Romance languages, cf. Seres/Espinal (2018).

has been hypothesized to be due to the expected continuation centring on the object NP. As for Spanish, there is an example given by Delbecque (2002) that supports this view. The Spanish verb *odiar* 'to hate', usually rejecting DOM with an inanimate object in a simple sentence (9a), seems to require the *a*-marker if a simple cause is spelled out for the object stimulus (9b). In contrast, DOM is ungrammatical if the IC bias for the object is overwritten by providing an internal cause that refers to the subject argument (9c).

(9) a. *Odia* ***ø/*a*** *las* *acelgas.*
hate.PRS.3SG ø/DOM the beets
'S/he hates beets.' (Torrego 1999: 1791; Delbecque 2002: 97)

b. *Odia* ****?ø/a*** *las acelgas* *por* *ser* *el* *plato* *típico*
hate.PRS.3SG ø/DOM the beets for be.INF the dish typical
del *seminario.*
of the seminary
'S/he hates beets since it is the typical dish of the seminary.'

c. *Odia* ***ø/*a*** *las acelgas* *porque* *no* *sabe*
hate.PRS.3SG ø/DOM the beets because NEG know.PRS.3SG
cómo *prepararlas.*
how prepare.INF=3PL
'S/he hates beets because s/he does not know how to prepare them.'
(Delbecque 2002: 97)

Delbecque (2002: 97) does not refer to implicit causality but gives a similar explanation for the difference with regard to the *a*-marker: In (9a), DOM is argued to be ungrammatical since *beets* as such generally do not provoke hatred. In (9b), by contrast, where DOM is nearly obligatory, they are construed as a cause for the hatred. In case the hatred lies outside the *beets*, namely in the thoughts of the subject participant, DOM is argued to be ungrammatical (9c).

What can be retained from this brief look at SE-psych verbs is the following: While they represent counterexamples to my agentivity-based hypothesis on DOM, they might support the view of another interacting verbal factor to agentivity. It is thus conceivable that the IC bias of a predicate, or more general, the default prediction of a certain explanation type for an argument, affects the preference for DOM (and therefore also the grammaticalization of the *a*-marker with certain verbs). Certainly, a study that systematically disentangles agentivity from implicit causality would be

needed to prove this claim.[10] Since the Sicilian data partly reveal contradictory results for one and the same verb class in the definite object NP vis-à-vis the indefinite object NP condition, a follow-up study should as well cross both referentiality types with both types of verbal factors.

In summary, my account is compatible with the generalization of thematic distinctness formulated for the occurrence of DOM with inanimate objects in Spanish (García García 2007, 2014). Operating with a more fine-grained distinction of the communicated agentivity for the object, as proposed by the agentivity scale, provides a subsumption of cases in which an asymmetric distribution of proto-properties is given but blurred because the direct object not only bears proto-patient but also (presupposed) proto-agent properties. Furthermore, it can possibly account for the different affinity of verb classes to take DOM with indefinite human NPs. It is left to further research to test the acceptability of DOM and no-DOM with the four classes of the agentivity scale in a controlled way for Spanish, too.

6.1.3 Other Romance languages

In the following, I will provide an overview of verbs (and verb classes) of a number of Romance languages and varieties that are reported to show an affinity for DOM in either diachrony or synchrony. Reviewing this cross-Romance evidence we shall see to what extent it may be subsumed under the present role-semantic account operating with a gradual notion of communicated agentivity for the object.

From a diachronic perspective, a majority of verbs that are reported to preferably select for a marked direct object in an early stage of the phenomenon were those that in some stage of Latin alternated between an accusative case frame and either a dative case frame or the equivalent AD + ACC (cf. e.g. Sornicola 1997, 2000, 2011 for Old Sicilian and Neapolitan; Pineda in press for Old Catalan[11]). Another related path of evolution is described by the term *karstification* (Sornicola 2011: 36), used to account for verbs that in archaic Latin selected for a dative that later disappeared in Classic Latin before then reappearing as *a*-marking in early stages of the Romance languages. Typical examples are the Sicilian lexemes *aiutari* 'to help', *ascultari* 'to listen to', *audiri* 'to hear', *clamari* 'to call', *cuntraddiri* 'to contradict' and *cunfurtari* 'to comfort' (Sornicola 1997: 72f.).

On that basis, Sornicola considers an etymological rather than a semantic-functional explanation for the affinity of the mentioned verb lexemes to take *a*-marked direct objects in diachrony. I would like to briefly state, however, why also from the role-

[10] The determination of the IC bias of an individual Spanish verb can be made with recourse to the study of Goikoetxea et al. (2008) who provide a list with 100 verbs with the calculated IC bias for each verb.

[11] Based on corpus data from Old Catalan (13th-15th century), Pineda (in press: 72-133) presents a systematic analysis of 33 alternating verbs in order to disentangle intransitive encodings from early instances of DOM. Among them are e.g. OE-psych verbs such as (*com)plaure* 'to please' or *satisfer* 'to satisfy' and interaction verbs like *ajudar* 'to help', *amenaçar* 'to threaten', *cridar* 'to call' or *perdonar* 'to forgive'.

semantic perspective adopted in this work, we can expect that these verbs are among the first to generalize the marker. As has been shown in section 4.1, Blume's (2000) licensing condition for marked valencies (e.g. */nom/dat*) and the generalization of thematic distinctness elaborated for DOM with inanimate objects in Spanish (García García 2007, 2014) formulate a similar role-semantic configuration: a non-asymmetric distribution of proto-properties among the two co-arguments. Based on this, the introduced cross-classification of the agent-patient asymmetry with the communicated agentivity for the object is hypothesized to apply to both marked valencies and DOM. In the beginning stage of DOM it is expected that verbs, especially those that are highly ranked in the agentivity scale and deviate from the canonical agent-patient asymmetry, are preferably marked. These are exactly the predicates that fulfil the licensing condition for marked case frames. That marked valencies and DOM are closely linked is borne out by the typological tendency that, besides locatives, it is dative or indirect object markers that serve as the most frequent source for newly developed differential object markers (Bossong 1985: 109f., 1991: 157f.; Primus 2012a: 73, 79; Iemmolo in preparation: 266).[12] Choosing the same morphosyntactic marker for both indirect objects and differentially marked direct objects thus mirrors semantic similarities that not only concern the prototypical animacy and individuation of the participant in question but also its relative agentivity – mostly *sentience* – in the given event.

The following table presents a collection of verbs that are (increasingly) attested with DOM in Old Ligurian, Old Neapolitan, Old Portuguese, Old Romanian and Old Sicilian:

Language	**Verbs**	**References**
Old Ligurian (14th century)	*adorar* 'to adore', *demandar* 'to ask for', *ihamar* 'to call', *creer* 'to believe', *fare* 'to make'[13], *ferir* 'to injure', *guardar* 'to look at', *lassar* 'to leave', *menare* 'to threaten', *meritar* 'to deserve', *offender* 'to offend', *parlar* 'to speak', *perdonar* 'to forgive', *pregar* 'to pray', *satisfar* 'to satisfy', *servir* 'to serve'	Parry (2003: 116-120)
Old Neapolitan (16th century)	*aiutare* 'to help', *(res)guardare* 'to protect, look at', *offendere* 'to offend', *perdonare* 'to forgive', *pregare* 'to pray', *recusare* 'to reject', *servire* 'to	Ledgeway (2009: 831f.),

[12] Iemmolo (in preparation: 266) postulates the following grammaticalization path of *a(d) + object NP* in Romance:
(1) locative, allative > (topic) > dative > (differential) direct object marker
The topic stage is put in brackets because he does not classify *ad* as having a proper topic marking function in Latin. As he notes, the dative stage is not considered necessary. This can be illustrated by Romanian where the form of the differential object marker *pe* (> Lat. *per* 'through') is different from the dative. Correspondingly, Antonov/Mardale (2014: 47) propose the following adapted grammaticalization path for Romanian:
(2) perlocative > locative > (dislocated) topic marker > DOM

[13] The verbs *fare* 'to make and *lassar* 'to leave' are listed here due to their appearance in causative constructions (Parry 2003: 119f.).

	serve', *scontrare* 'to collide', *soccorrere* 'to save', *supplicare* 'to beg'	Sornicola (1997: 74f.)
Old Portuguese (16th century)	*abrandar* 'to calm', *acompanhar* 'to accompany', *ajudar* 'to help', *comover* 'to move', *confortar* 'to comfort', *desamparar* 'to leave without support', *espantar* 'to scare', *estimar* 'to estimate', *ferir* 'to injure', *louvar* 'to praise', *matar* 'to kill', *receber* 'to receive', *tratar* 'to treat', *vencer* 'to win'	Delille (1970: 63)
Old Romanian (16th-17th century)	*a dărui* 'to present someone with something', *a împuternici* 'to empower', *a însura* 'to marry', *a întreba* 'to ask', *a înzestra* 'to endow', *a logodi* 'to engage', *a milui* 'to mercifully endow', *a peţi* 'to woo', *a pierde* 'to kill', *a pofti* 'to invite', *a porunci* 'to order', *a robi* 'to enslave', *a ruga* 'to kindly ask/pray', *a scuti* 'to spare', *a turci* 'to forcefully convert to Islam', *a ucide* 'to kill', *a vătăma* 'to wound', *a adăuga* 'to add', *a chema* 'to call', *a lovi* 'to hit', *a pomeni* 'to mention', *a prinde* 'to catch', *a slobozi* 'to free', *a urî* 'to hate', *a vinde* 'to sell'	Mardale (2015: 236)
Old Sicilian (13th-15th century)	*adurari* 'to adore', *aiutari* 'to help', *amari* 'to love', *ascultari* 'to listen to', *auchidiri* 'to kill', *audiri* 'to hear', *aviri in odiu* 'to hate', *basari* 'to kiss', *battiri* 'to beat', *charmari* 'to fascinate', *clamari* 'to call', *consolari* 'to console', *cunfurtari* 'to confort', *cunsiglari* 'to deliver', *cuntraddiri* 'to contradict', *dampnari* 'to condemn', *dampnificari* 'to harm', *disfamari* 'to defame', *firiri* 'to injure', *ingannari* 'to deceive', *offendiri* 'to offend', *honorari* 'to honour', *iastimari* 'to curse', *inculpari* 'to accuse', *inflamari* 'to inflame', *iniuriari* 'to injure', *laudari* 'to praise', *perdunari* 'to forgive', *prigari* 'to pray', *reprindiri* 'to take over', *riquediri* 'to request', *serviri* 'to serve', *toccari* 'to touch'	La Fauci (1991: 391-393), Sornicola (1997: 72f.)

Table 3. Verbs attested for different Romance languages and varieties showing a diachronic preference for DOM

As a first very general observation, one can state that the majority of the listed predicates denotes either events that prototypically take place between humans, be it either in social or physical interaction, or emotional states. Among the attested verbs, there are a couple of OE-psych verbs and one symmetrical predicate, all entailing the object's agentivity:

(10) OE-psych verbs and symmetrical predicates (*entailed agentivity*):

a. Old Ligurian: *offender* 'to offend', *satisfar* 'to satisfy'

b. Old Neapolitan: *offendere* 'to offend'

c. Old Portuguese: *abrandar* 'to calm', *comover* 'to move', *confortar* 'to comfort', *espantar* 'to scare'

d. Old Romanian: *a însura* ‘to marry’

e. Old Sicilian: *charmari* ‘to fascinate’, *consolari* ‘to console’, *cunfurtari* ‘to comfort’, *ingannari* ‘to deceive’, *offendiri* ‘to offend’

The class of interaction verbs, presupposing the object’s agentivity, is by far the largest subgroup: it comprises inter alia verbs of communication and religious practices, helping and auditory perception verbs as well as verbs of threatening and punishment. Note that here only predicates are listed that allow for a straightforward classification as interaction verbs:

(11) Interaction verbs (*presupposed agentivity*):

a. Old Ligurian: *demandar* ‘to ask for’, *ihamar* ‘to call’, *menare* ‘to threaten’, *parlar* ‘to speak’, *perdonar* ‘to forgive’, *pregar* ‘to pray’, *satisfar* ‘to satisfy’, *servir* ‘to serve’

b. Old Neapolitan: *aiutare* ‘to help’, *perdonare* ‘to forgive’, *pregare* ‘to pray’, *recusare* ‘to reject’, *servire* ‘to serve’, *soccorrere* ‘to save’, *supplicare* ‘to beg’

c. Old Portuguese: *acompanhar* ‘to accompany’, *ajudar* ‘to help’, *louvar* ‘to praise’, *receber* ‘to receive’, *vencer* ‘to win’

d. Old Romanian: *a împuternici* ‘to empower’, *a întreba* ‘to ask’, *a peți* ‘to woo’, *a pofti* ‘to invite’, *a robi* ‘to enslave’, *a ruga* ‘to kindly ask/pray’, *a turci* ‘to forcefully convert to Islam’, *a chema* ‘to call’, *a prinde* ‘to catch’, *a slobozi* ‘to free’

e. Old Sicilian: *aiutari* ‘to help’, *ascultari* ‘to listen to’, *audiri* ‘to hear’, *clamari* ‘to call’, *cuntraddiri* ‘to contradict’, *dampnari* ‘to condemn’, *disfamari* ‘to defame’, *honorari* ‘to honour’, *laudari* ‘to praise’, *perdunari* ‘to forgive’, *prigari* ‘to pray’, *serviri* ‘to serve’

Moreover, verbs of injuring and killing involving the potential agentivity of the object participant occur across languages:

(12) Verbs with an affected animate object (*potential agentivity*):

a. Old Ligurian: *ferir* ‘to injure’

b. Old Portuguese: *ferir* ‘to injure’, *matar* ’to kill’

c. Old Romanian: *a pierde* ‘to kill’, *a ucide* ‘to kill’, *a vătăma* ‘to wound’

d. Old Sicilian: *auchidiri* ‘to kill’, *inculpari* ‘to accuse’, *iniuriari* ‘to injure’

Interestingly, contradictory to the predicted agentivity scale but in line with the previously discussed synchronic data from Spanish, the given diachronic data also contain some SE-psych verbs that display an affinity for the *a*-marker:

(13) SE-psych verbs (*unspecified for agentivity*):

a. Old Ligurian: *adorar* ‘to adore’

b. Old Portuguese: *estimar* ‘to estimate’

c. Old Romanian: *a urî* ‘to hate’

d. Old Sicilian: *adurari* ‘to adore’, *amari* ‘to love’, *aviri in odiu* ‘to hate’

The fact that in Old Romanian, where the differential object marker *pre* (> Lat. *per* ‘through’) and the dative show no syncretism, verbs similar to those in the other languages are attested with DOM and, likewise, alternate between dative and transitive encodings (Mardale 2015: 216ff.), is a further argument in favour of a semantic rather than a purely etymological explanation of emerging DOM. In example (14), it is illustrated by the verb *a ruga* ‘to kindly ask/pray’ how the different case frames oscillate in Old Romanian texts: While the verb occurs with a morphological dative in (14a), it is used transitively with DOM in (14b) and without DOM in (14c).

(14) a. *Rogu-mă* *măriei* *tale* *[…]* [Dative]
pray.PRS.1SG=REFL highness.DAT 2SG
‘I ask your highness […]’ (1598)

b. *Eu,* *împăratul,* *rog* ***pre*** *domneta […]* [DOM]
1SG king.DEF pray.PRS.1SG DOM highness.2SG
‘I, the emperor, ask your highness […]’ (1600)

c. *Rugîm* ***ø*** *milostivul* *dumnedzău* [ø DOM]
pray.PRS.1PL ø merciful.DEF God
‘We pray [to] God [who is] merciful.’ (1610)

(Mardale 2015: 216f.; Antonov/Mardale 2014: 39; emphasis: S.M.)

The presented diachronic evidence broadly matches the distinction of agentivity classes drawn and provides first clues that this approach may be indeed suitable to capture a cross-Romance verb-semantic factor for DOM that already mattered in early stages of the phenomenon. Yet, quantitative corpus data including an in-depth analysis of relevant examples are needed to properly prove this hypothesis.

From a synchronic perspective, a look at Romance languages in which DOM is in an incipient stage, such as colloquial Standard Italian or French, provides further evidence for my account. Let us take the example of colloquial Italian first. As has been reported by several authors (Berretta 1989, 1991; Benincà 1986; Iemmolo 2010a, in preparation; Belletti 2018), *a*-marking of the direct object systematically occurs under certain information-structural and semantic conditions. Typical examples are given in (15):

(15) a. ***A*** *me* *preoccupa* *Torino:* *è* *una* *città* *difficile*
DOM 1SG worry.PRS.3SG Turin be.PRS.3SG a city difficult
‘Turin worries me: it is a difficult city.’

b. ***A*** *loro* *le* *aspettava* *Adone* *in* *doppio* *petto* *blu*

DOM 3PL CL.3PL wait for.PST.3SG PN in double-breasted blue

'Adone waited for them in a blue double-breasted (suit).'

(Berretta 1991: 228; 225)

In both (15a) and (15b), the *a*-marked direct object is a strong personal pronoun that occurs in non-canonical position: while it is – most probably – clause-internal and thus pre-posed in (15a), it occurs in left-dislocated position in (15b) being resumed by a clause-internal clitic. The two examples are part of a corpus of spontaneous speech of colloquial Italian, compiled by Berretta (1991). Concerning the verbs that are attested with DOM in Italian, she identifies the following three verb classes (Berretta 1991: 137-138):[14]

1. OBJECT-EXPERIENCER PSYCH-VERBS:
affascinare 'to fascinate', *attrarre* 'to attract', *colpire* 'to impress, strike', *confortare* 'to comfort', *consolare* 'to cheer up', *convincere* 'to convince', *deludere* 'to disappoint', *disturbare* 'to disturb', *divertire* 'to entertain', *eccitare* 'to excite', *entusiasmare* 'to excite', *incantare* 'to enchant', *ingannare* 'to deceive', *innervosire* 'to irritate', *invitare* 'to tempt', *mettere (di buon umore)* 'to cheer', *persuadere* 'to persuade', *preoccupare* 'to worry', *rattristare* 'to sadden', *spaventare* 'to frighten',

2. CAUSATIVE CONSTRUCTIONS with *fare* 'to make'/ *lasciare* 'to let' + infinitive:
fare dormire 'to make sleep', *fare ingrassare* 'to fatten', *fare morir dal ridere* 'to make die laughing', *fare ridere* 'to make laugh', *fare piangere* 'to make cry', *fare sentir male* 'to make feel bad', *lasciare salire* 'to pick up', *lasciare stare* 'to let be/go', and

3. OTHER VERBS:
accompagnare 'to accompany', *aspettare* 'to expect', *chiamare* 'to call', *coccolare* 'to cuddle', *conoscere* 'to (get to) know', *dannegiare* 'to damage, harm', *fermar*e 'to stop', *fregare* 'to wipe', *fucilare* 'to shoot', *graffiare* 'to scratch', *incastrare* 'to trap', *informare* 'to inform', *incolpare* 'to blame', *lasciare (in pace)* 'to leave sb alone', *mandare* 'to send', *mettere (in galera)* 'to jail', *pagare* 'to pay', *picchiare* 'to beat', *portare* 'to carry', *prendere* 'to take', *proteggere* 'to protect', *pungere* 'to offend', *ringraziare* 'to thank', *rovinare* 'to ruin', *salutare* 'to greet', *seppellire* 'to bury', *sposare* 'to marry', *stancare* 'to tire sb out', *stendere* 'to hang', *svegliare* 'to wake sb up', *temere* 'to fear', *toccare* 'to touch', *trattare (male)* 'to misuse', *vedere* 'to see'

While the first two groups of verbs are precisely classified as (1) OE-PSYCH VERBS and (2) CAUSATIVE CONSTRUCTIONS, the third group is not further specified, being simply labelled (3) OTHER VERBS. Yet, as has been argued by Cassarà/Mürmann (2021: 68), a good number of verbs of class (3) overlap with Blume's (1998) category of *interaction verbs* presupposing an agentive object in a temporally prior subevent such as *chiamare* 'to call', *ringraziare* 'to thank' or *salutare* 'to greet'.

However, it is important to stress that agentive properties of the verb alone can never trigger DOM in Italian: as already suggested by the above examples in (15), (i) the object NP must be highly ranked on the referentiality scale and (ii) the direct object

[14] All relevant examples of the corpus are listed in the appendix of her article (Berretta 1991: 225-229).

must occupy a non-canonical sentence position (either dislocated or pre-posed). This is borne out by (16) where *a*-marking in canonical SVO word order is ungrammatical regardless of the presence of an OE-psych verb and a tonic pronoun as object NP:

(16) *La soluzione ha soddisfatto solo* ***ø/*a*** *noi*[15]
the solution have.PRS.3SG satisfy.PTCP only ø/DOM 1PL
'The solution satisfied only us.' (cf. Benincà 1986: 231)

In an online acceptability judgement study Cassarà/Mürmann (2021) tested the reported preference of certain verb classes to take DOM in Italian. Participants from Northern Italy (mainly Lombardy) were exposed to spoken sentences containing *a*-marking in clitic left-dislocation structures which they had to rate on a five-point scale. Verb type and NP type of the object were manipulated. As for the former, OE-psych verbs (e.g. *affascinare* 'to fascinate'), interaction verbs (e.g. *salutare* 'to greet') and highly transitive verbs (e.g. *ferire* 'to injure') were tested. In a nutshell, the data reveal the following cline of acceptability of verb classes (Cassarà/Mürmann 2021: 91ff.):

(17) Acceptability of DOM for verb type:
OE-psych verbs > interaction verbs > highly transitive verbs

These results confirm the hypothesis presented here: while OE-psych verbs entail the object's sentience, interaction verbs presuppose an agentive participant and highly transitive verbs solely assign proto-patient properties to their object (with sentience being only pragmatically conveyed). Note that for OE-psych verbs, it was controlled for a non-causative reading in that only inanimate subjects have been used. The role-semantic configuration thus corresponds to the most unprototypical pattern in the agentivity hierarchy where the object outranks the subject in terms of agentivity.

Another language where DOM is an equally marginal phenomenon is regional and colloquial French. As reported by Iemmolo (2010a: 253f., in preparation: 256-259), in Languedoc French for instance, DOM only occurs in dislocated constructions. Two examples are given below: In (18a), we find a right-dislocated pronominal object, while (18b) involves a left-dislocated human NP accompanied by the possessive article.

(18) a. *il faut l' aider,* ***à*** *elle*
3SG must.PRS.3SG CL.F.3SG.ACC help.INF DOM 3SG
'As for her, we should help her.'

[15] This example is given by Benincà (1986: 231f.) to prove that the *a*-marked (left-)dislocated complement occurring with OE-psych verbs in Italian (as in *A noi, la soluzione non ci ha soddisfatti* 'The solution did not satisfy us') is indeed a direct and not an indirect object. Otherwise, the *a*-marker would appear in canonical word order, too.

b. ***à*** *ton père, je l' ai vu*

DOM your father 1SG CL.M.3SG.ACC have.PRS.1SG see.PTCP

'As for your father, I saw him.' (Iemmolo in preparation: 256f.; Rohlfs 1971: 68)

Likewise, in Brussels French and spoken French only dislocated direct objects can obtain the *a*-marker. Here, DOM only appears with personal pronouns, as exemplified in (19). As for Languedoc French, the dislocated object is resumed by a clause-internal clitic.

(19) a. *Je la suivais,* ***à*** *elle*

1SG CL.F.3SG.ACC follow.IPRF.1SG DOM 3SG

'I followed her.'

b. ***à*** *moi, ça me touche beaucoup*

DOM 1SG this CL.1SG touch.PRS.3SG a lot

'As for me, I'm very touched by that.' (Iemmolo in preparation: 257f.)

Fagard/Mardale (2014) provide a first more extensive investigation of DOM in French. They base their study on corpus data from spoken corpora of different varieties of French as well as written, internet and diachronic corpora. The authors show that, although extremely rare, the phenomenon of DOM can be attested across varieties in colloquial registers and, moreover, already existed in prior language stages. In addition to *inherent factors* of the object NP they qualify so-called *global factors* as triggering *a*-marking; these not only point to topicality but also to verb type. As for the latter, the authors list the following verb lexemes as favouring DOM (Fagard/Mardale 2014: 158):

(20) *aider* 'to help', *aimer* 'to love', *applaudir* 'to applaud', *blesser* 'to hurt, injure', *écouter* 'to listen to', *empêcher* 'to prevent', *escroquer* 'to swindle', *faire* 'to make', *faire chier* 'to get up sb's nose', *frapper* 'to hit', *insulter* 'to insult', *intéresser* 'to interest', *interroger* 'to query', *kiffer* 'to like', *préoccuper* 'to worry', *prier* 'to pray', *regarder* 'to look at', *soigner* 'to take care of', *toucher* 'to touch', *tuer* 'to kill', *voir* 'to see'

Note that, here again, more than half of the predicates are either OE-psych verbs (*faire chier* 'to get up sb's nose', *intéresser* 'to interest', *préoccuper* 'to worry') or interaction verbs (*aider* 'to help', *applaudir* 'to applaud', *écouter* 'to listen to', *empêcher* 'to hinder', *escroquer* 'to swindle', *interroger* 'to query', *prier* 'to pray', *soigner* 'to take care of'). The authors, too, identify semantic similarities among the verbs, more specifically the semantic role of beneficiary (and maleficiary). The preference of these verbs to take DOM in Modern French is analyzed as a reaction to a mismatch between the typically "dative" semantic role and its morphosyntactic encoding as a direct object ("accusative"):

> Dans notre corpus, un certain nombre d'exemples sont susceptibles d'être analysés de deux manières, comme des constructions accusatives (MDO) ou datives. C'est le cas en

particulier de verbes comme *aider* qui, en français standard, sont construits avec un objet direct *(aider quelqu'un)* dont le rôle sémantique est BÉNÉFICIAIRE/ DESTINATAIRE. Etant donné cette discordance entre sémantique et morphosyntaxe, il n'est pas surprenant de trouver des occurrences comme (15):

(15) j'aidais **à** mon, **à** papa pendant que mon frère était parti (PFC, Normandie)

(Fagard/Mardale 2014: 157)

In summary, the presented survey of diachronic and synchronic data suggests that the consideration of the communicated agentivity for the object might serve as one connecting link for understanding the preference of certain predicates to generalize DOM across Romance languages. While this section was confined to agentivity, it is left to section 6.2.3 to take a closer look at interacting constraints concerning the referential status of the object and its degree of topicality.

6.2 Interaction of agentivity with nominal factors (and topicality)

While the previous section focused on agentivity as a factor for DOM in individual Romance languages, in the following, I will take a closer look at the relation between agentivity and the nominal factors animacy (6.2.1) and definiteness (6.2.2). This is indispensable since the phenomenon of DOM, despite my focus on agentivity, is the outcome of an interplay of various factors. As a further step, I will outline how agentivity could be integrated into the overall picture of DOM in Romance (6.2.3). Importantly, this does not only include its relation to semantic-pragmatic properties of the object NP but also to the information-structural factor topicality, syntactically mirrored by constraints in object position.

6.2.1 Agentivity and animacy

This section aims at revisiting the relation between agentivity and animacy based on the view taken in this study. Since the objective was to explore the influence of agentivity on DOM while keeping the object's animacy constant, it was necessary to clearly dissociate both concepts in a first step. This was particularly challenging for the notion of *potential agentivity* as referring to animate objects as inherently agentive participants, which are interpreted as potential proto-agents in a given event (cf. García García 2014: 129). In this way, it was possible to subsume the obligatory marking of animate, or more precisely, human definite direct objects in Spanish under a proto-role approach to DOM.

While such argumentation is theoretically anchored in the view of human entities being conceived of as typical agents in a cognitive ontology (cf. Dahl/Fraurud 1996, Fraurud 1996, Dahl 2008) and further supported by neurolinguistic studies (cf. García García 2014: 134-137 for an overview), it serves rather as a general explanation for the impact of animacy on grammatical coding alternations. For empirical purposes, though, the notion of inherent potential agentivity has been argued to be too broad since it does not provide criteria to properly differentiate between subclasses of verbs that typically select an animate (or human) direct object. Primus (2012a)

further elaborates on a role-semantic account on *a*-marked animate objects in Spanish. In her extended dative default, she makes a distinction of two cases in which the object can obtain proto-agent properties:

(21) The dative is used for an argument with a low number of proto-agent properties that are **i) subcategorized by the verb** or **ii) assigned according to the intrinsic meaning of the direct object referent.** (Primus 2012a: 81; emphasis: S.M.)

The first case (21i) concerns proto-agent properties entailed by the verb, while the second case (21ii) points to inherent properties of the object referent derived from its animacy, qualifying it to act as a proto-agent in the event denoted by the predicate (cf. Primus 2012a: 75). The paper remains unclear about the extent to which direct objects can be included under subcategorized proto-agent properties (21i). Apparently, these are understood as solely applying to indirect objects while intrinsic proto-agent properties (21ii) refer to differentially marked direct objects.[16] Yet, in a later work of García García/Primus/Himmelmann (2018) it becomes clear that case (21i) also applies to transitive predicates, such as the entailment of *sentience* for the object of transitive OE-psych verbs:

> Replacing animacy with sentience in the lexical entry of these verbs eliminates some of the substantial disadvantages of older conceptualizations of selectional restrictions. First, the feature [+ Animate] is too broad and cannot separate animate object verbs into distinct, grammatically relevant subclasses. Besides psychological predicates such as *frighten* and *scare*, there are communication verbs such as *nominate* and *acclaim*, interactional verbs such as *help* and *assist*, and diverse others such as *kill* and *envy*.
> (García García/Primus/Himmelmann 2018: 35f.)

As stated above, besides OE-psych verbs, there are other verb classes that typically select for an animate direct object. I proposed integrating two more sub-steps between the two poles of subcategorized and inherent agentivity in order to differentiate also between these classes: *presupposed agentivity* and *potential agentivity*. Both are defined in a strictly relational sense. Presupposed agentivity is derived based on the work of Blume (1998, 2000). It is the defining characteristic of objects of so-called (social) interaction verbs. Relational potential agentivity has been introduced as resulting from the generalized conversational implicature AFFECTED + ANIMATE +> SENTIENT which is given for the objects of verbs such as 'to kill', 'to injure' or 'to insult'. Table 4 sums up the differences of the four classes:

[16] In García García (2014: 147-170) this is specified in more detail: He shows that entailed proto-agent properties also occur with transitive predicates and defines several predicates for this purpose, e.g. ACT-LIKE. However, the concrete way in which this ACT-LIKE is carried out, i.e. which proto-agent properties are present, remains tied to the respective context.

Communicated agentivity for O → Further characterization ↓	**entailed >**	**presupposed >**	**potential >**	**unspecified**
type of relational agentivity	verbal entailment	(semantic) presupposition	generalized conversational implicature	potential subject status of the object
meaning dimension	semantics	semantics	semantics/ pragmatics	pragmatics
example verbs	*to frighten, to molest, to surprise*	*to help, to greet, to threaten*	*to injure, to insult, to kill*	*to find, to love, to see*
animacy preference of the example verbs	+ animate	+ animate	+ animate	± animate

Table 4. Characterization of the four types of communicated agentivity for the object

In the following, more general remarks on the relation between animacy and agentivity are made that consider first and foremost the implicational relations that are discussed to exist between both concepts and, closely linked to the previous point, the difference between inherent and relational agentivity.

García García (2014: 129ff.), after having given an agentivity-based explanation for DOM with inanimate objects in Spanish, also states the relation between animacy and agentivity in more precise terms. In this way, the obligatory marking of nearly every human definite direct object in Modern Spanish should be accounted for by agentivity as well. He establishes the following unilateral implications from the individual proto-agent properties to the animacy status of the participant:

(22) Unilateral Implications from proto-agentivity to animacy:

a. CTRL(x, …) → HUMAN(x, …) → ANIMATE(x, …)

b. POSSESS(x, …) → HUMAN(x, …) → ANIMATE(x, …)

c. EXPER(x, …) → ANIMATE(x, …)

d. MOVE(x, …) → HUMAN(x, …) → ANIMATE(x, …)

(García García 2014: 131)

Primus (1999a: 50) formulates similar implications, though a bit weaker, restricting them to prototypical control (P-CONTROL) and prototypical experience (P-EXPER). The implications from (alienable) possession to animacy, strictly speaking to human beings (22b), and from (autonomous) movement to animacy are made by Primus as well (2012a: 85). Hence, independently of verb- and context-specific requirements, one could say that proto-agent argument properties, and sub-types of them,

most prototypically refer to animate or human entities. Reversing the perspective by distributing the proto-agent properties along the animacy scale, it can be seen that due to their inherent properties, human beings can accumulate more proto-agent properties than non-human animate beings, which, in turn, can accumulate more proto-agent properties than inanimates:

human >	**animate** >	**non-animate**
+ control		
+ alienable possession		
+ sentient	+ sentient	
+ movement	+ movement	
+ causation	+ causation	+ causation
+ independent existence	+ independent existence	+ independent existence

Table 5. Animacy and potential agentivity (García García 2018b: 60)

It is argued that due to their inherent agentivity potential, "arguments that refer to animate or human referents, respectively, are interpreted as participants that are able to potentially control, perceive or dispose of another participant" (García García 2014: 134; translation S.M.).[17] This can be exemplified by the following sentence in which the *a*-marked human object participant can function as potential subject participant of the same event, as the added predication in brackets illustrates:

(23) *El jardinero vio **a** la asesina (y ella también vio **al** jardinero).*

'The gardener saw the murderess (and she also saw the gardener).'

EXPER(x, y) & HUMAN(x) & HUMAN(y) (& EXPER(y, x))

(García García 2014: 134)

The direct object participant of the first predication (*la asesina* 'the murderess') is qualified as inherently sentient due to its property of being human. For this reason, it can be interpreted as relationally sentient in a seeing event, too, and potentially function as subject in the very same event. As further illustrated by García García, the same would not be possible for an inanimate object argument:

(24) *El jardinero vio la pala (*y ella también vio **al** jardinero).*

'The gardener saw the shovel (*and it also saw the gardener).'

EXPER(x, y) & HUMAN(x) & INANIMATE(y) (*& EXPER(y, x))

The ungrammaticality of the added predication in (24) can be explained by the lack of inherent agentive potential of the inanimate object *la pala* 'the shovel' which, in its literal meaning, cannot obtain the relative proto-agent property *sentience* in a seeing event and thus fails to function as a subject in the same event. In this argumentation, only animate (or human) entities qualify as potential agents due to their inherent

[17] "Argumente, die sich auf belebte bzw. menschliche Referenten beziehen, werden als Partizipanten verstanden, die ein Ereignis potenziell kontrollieren, wahrnehmen oder über einen anderen Partizipanten verfügen können." (García García 2014: 134).

agent properties. From a syntagmatic perspective, they compete with the subject, which bears relational proto-agent properties. Primus (2012a: 85) explains the kind of pragmatic inference from the animacy of an argument to its potential agentivity as *abductive reasoning*. Levinson (2000: 43) defines abduction as "the hypothesizing of a fact that would, in conjunction with a general law, explain an observed fact". Taking example (23) above, abductive reasoning can be described as follows:

(25) a. EXPER(x,…) → ANIMATE(x,…) (known law)

b. ANIMATE(x,…) (observed fact)

----------Abductive reasoning

c. Potential EXPER(x,…) (hypothesized explanation)

Hence, given the premises that (25a) (prototypical) experience implies animacy and that (25b) an object argument is observed as animate, the animate object argument is interpreted as bearing potential sentience (25c). Just as for a conversational implicature, the informativity of an argument is increased in this way (Primus 2012a: 85). Likewise, it can be cancelled in context:

(26) *La asesina vio **al** jardinero (pero el no podía ver **a la** asesina. Estaba inconsciente.)*

'The killer saw the gardener (but he could not see the killer. He was unconscious).'

The cancelation of the potential agentivity of the object participant in the same event does not seem to affect the insertion of DOM. It must be noted, however, that abduction is less easily verifiable than a "normal" generalized conversational implicature of proto-agent properties. While the former points to a further additive predication, which is a reversal of the actual event and hence highly hypothetical, the latter is true between a participant and the predicate of an event, as exemplified for the subject in (27):

(27) *Pepe mató a la actriz.*

'Pepe killed the actress.'

CAUSE(x, BECOME(KILLED(y)) & MOVE(x, y) & **HUMAN**(x)

+> **CTRL**(x, BECOME(KILLED(y))

+> **EXPER**(x, BECOME(KILLED(y))

Without further information, the human subject of *matar* 'to kill' is interpreted to act volitionally and hence with sentience. The implicatures of proto-agent properties as illustrated in (27) thus concern the thematic relation between subject and verb in the *actual* event.

It is questionable then, how the hypothetical explanation for inherent potential agentivity can be applied to individual verbs. As I shall further argue, the crucial point is to draw a distinction between inherent (proto-)agentivity, on the one hand, and

relational (proto-)agentivity on the other hand. Let us have a look at the following examples, which seem to be contradictory to the premise in (25a) stating that certain proto-agent properties (or sub-properties) imply animacy. The exceptions concern intelligent machines and natural forces which, as already mentioned by Cruse (1973: 16, 21) and Dowty (1991: 574) among others, can bear a certain degree of agentivity in certain events. In (28), we see that for the English verb *to calculate*, which entails *volition* and *sentience* for its subject, not only human entities (like *the secretary*) appear very naturally as subjects but also intelligent machines like *the computer*. Inanimate objects such as *the book*, in contrast, are odd in subject position.

(28) *The secretary/The computer/#The book calculated the price.*

Another example is given in (29a). The verb *to batter* entails *volition*, *sentience* and *autonomous movement* to its subject. However, it cannot only select human beings as subjects but also natural forces such as *the storm*. Inanimate objects such as *the book*, again, cannot function as subjects unless their movement is externally caused. The capability of autonomous movement is also given for intelligent machines for certain predicates as in (29b). Since the machine is programmed to switch itself off at a specific time, it can be autonomously active in the limited range of the programmed commands.

(29) a. *The mob/The storm/#The book battered the windows.* (Schumacher 2018: 88)

b. *The machine automatically switches (itself) off at 6 p.m.* (Cruse 1973: 21)

Interestingly, for my purposes, also in object position machines can appear with verbs that prototypically select a human object, such as the interaction verb *llamar* 'to call' in Spanish. This verb presupposes an agentive direct object that has the capacity to perceive the call and react to it. Such a role is typically taken by humans but can also be occupied by certain machines like an elevator, which is programmed to react to the call of another participant. In these cases, we find *a*-marking in Spanish as shown in (30). Note that the verb *llamar* 'to call' and the direct object *el ascensor* 'the elevator' are both used in their literal meaning here (cf. García García/Primus/Himmelmann 2018: 30).

(30) *¡Hans, puñeta, llama* ***al*** *ascensor!*

'Hans, damn, call the elevator!'

(García García 2014: 189; García García/Primus/Himmelmann 2018: 30)

This example is an argument in favour of the relevance of relational instead of inherent agentivity for DOM in Spanish: Machines such as an elevator can only be autonomously active and sentient with respect to a programmed reaction but not verb-independently. That is, they cannot fill any slot that presupposes an agentive object argument (e.g. *#Hans greeted the elevator.*), neither they can function as subject in the same event *(*e.g. *#The elevator called Hans.*). The given examples illustrate the verb-dependent compatibility of proto-agent properties like *sentience*, *volition* and

autonomous motion with certain inanimate entities, more precisely with intelligent machines and natural forces, as well as the gradience of animacy. Importantly, none of these uses involves a conceptual shift from an inanimate to an animate entity. A similar explanation can be proposed for other interaction verbs that systematically take DOM with inanimate objects in Spanish (García García 2014: 187f.), such as *temer* 'to fear' and *amenazar* 'to threaten'. In contrast to example (30) above, the *a*-marked direct objects in (31a) and (31b) denote abstract referents.

(31) a. *Jano, por norma, no temía **a** la soledad.* [ADESSE: CAR:070.32][18]

'Jano, generally, did not fear loneliness.'

b. *el sistema totalitario que ha usurpado el nombre y la tradición del socialismo-- amenaza sobre todo y ante todo **a** la democracia.* [ADESSE: TIE:123.18]

'The totalitarian system that has usurped the name and tradition of socialism-- threatens first and foremost democracy.'

The verb *temer* 'to fear', typically described as an SE-psych verb, can be at the same time described as an interaction verb since it presupposes that the object participant displays a certain agentivity potential that is apt to evoke fear in the subject participant. The example in (31a) does not involve a conceptual shift of the inanimate object referent. Rather, the concept of *la soledad* 'loneliness' is used in its literal meaning, namely as an emotional state caused by the lack of social contacts or deeper relationships. Since this state increases the risk for mental diseases or disorders, such as depression or alcoholism, it can be ascribed an agentivity potential for which it is not necessary to shift 'loneliness' into an animate entity. The verb *amenazar* 'to threaten' in (31b), too, presupposes an agentive object participant. According to Blume (1998: 273; 279, fn. 17), it can only be felicitously uttered if the object participant fears something before the threat of the subject participant is announced. Democracy as a system of government is an abstract entity and as such cannot experience fear in the common sense. However, it is characterized by a range of activities, such as elections of representatives, which can be suppressed by competing systems of government like autocracy or oligarchy. When elections are rigged for example, the system of democracy is in danger. The threatening of democracy hence presupposes certain suppressable activities. In this line of argumentation, also for *amenazar* 'to threaten' as used in the above example, no animacy shift is needed. Instead, the occurrence of DOM can be accounted for by the presupposed agentivity of the direct object relative to the predicate in question.

A different case is given with animacy shifts.[19] In these cases, an NP that is considered inanimate from its lexical meaning is conceptualized as animate in a given context (Nieuwland/Van Berkum 2006; Vogels et al. 2013; Rosenbach 2008; Yamamoto

[18] Both examples (31a-b) are taken from the ADESSE corpus (http://adesse.uvigo.es/) which was the basis of García García's (2014) study of DOM with inanimate objects in Spanish.
[19] The understanding of *animacy shift* is grounded within type-theoretical semantics referring to a shift from an inanimate type to a human being/a person (cf. e.g. de Swart/de Hoop 2018).

1999). Nieuwland/Van Berkum (2006) provide experimental evidence that discourse contextual cues can overrule the semantic-lexical animacy of an argument in language processing. As they show for Dutch, predicates that require agentive properties for their arguments, such as the object of 'comfort' or the subject of 'be in love', and which are therefore typically inappropriate with inanimate arguments filling these slots, can be processed without difficulties also with inanimate arguments given the appropriate contextual cues. The study is based on two experiments in which the test persons listened to short narratives while an EEG was recorded. By means of event-related brain potentials (ERP), it could be shown that sentences like *the girl comforted the clock* or *the peanut was in love* elicited an N400 effect without a corresponding context. This means that 400 milliseconds after playing such a sentence, a negative amplitude was recorded. The effect, whose description goes back to Kutas/Hillyard (1980), is explained by difficulties in processing the semantic content of linguistic utterances. However, such an effect did not occur if a discourse context was given in which the clock was conceptualized as human being, such as in a cartoon-like story in which the clock is described as feeling depressed, or in a context in which a peanut is introduced as dancing with a big smile on its face and singing about a girl it has met (Nieuwland/Van Berkum 2006: 1106). Thus, once inanimate entities like a clock or a peanut are attributed behaviours and feelings that are characteristic of humans, they can be equally well processed with predicates that are typically reserved for animate entities.[20] Furthermore, in such contexts, for instance in a story about an amorous peanut, a predicate like *the peanut was in love* is shown to be easier to process than canonical predicates such as *the peanut was salted*. This manifested itself in an N400 effect for the latter. Although role-semantic properties are not mentioned by the authors, they could as well serve as an explanation for the measured effects. Since both the predicates 'comfort' and 'be in love' entail *sentience* – 'comfort' for its object and 'be in love' for its subject – these slots cannot be filled with entities that cannot exhibit sentience relative to the event. Yet, once conceptualized as entities that are capable of having mental states, a clock or a peanut can be sentient relative to the events mentioned and hence satisfy the proto-agent property entailed by the verb.

Coming back to DOM in Spanish, there are indeed cases where DOM appears with abstract entities being conceptualized as animate in the given context. An example is shown in (32):

(32) *Llamó* ***a*** *la muerte.*

[20] Another kind of animacy shifts is given in cases like *The ham sandwich is sitting at table 7* (Nunberg 1979: 149) or *The nurse told the doctor that the hepatitis had called a few minutes ago* (Schumacher 2011: 200). In these cases, a contextually salient property of a human individual is used to refer to this person, i.e. 'the person who ordered the ham sandwich' in a restaurant setting or 'the person suffering from a hepatitis' in the context of a doctor's office. As shown by Schumacher (2011) in an ERP-study, such cases require increased processing costs, indicated by a late positivity. This effect is accounted for by "processes associated with the modification and updating of discourse representation structures that are crucially induced by pragmatic requirements" (Schumacher 2011: 200).

'S/he called out to Death.' (García García/Primus/Himmelmann 2018: 29)

The verb *llamar* 'to call', in the given reading, presupposes a sentient direct object participant, one that is capable of perceiving the call of the subject participant and to react to it. With the literal meaning of *la muerte* 'death' this is hardly possible since the abstract entity 'death' cannot be ascribed an inherent proto-agent property per se. Therefore, in order for *la muerte* to fill the object slot we might have to assume a conceptual shift from an inanimate to an animate entity. Note, however, that it is not animacy as such that is relevant here but rather the inherent proto-agent property of sentience, which is required in order to satisfy the predicate's restriction to select an object that could perceive the subject's call and react to it (cf. García García/Primus/Himmelmann 2018: 31). This would be the same kind of explanation that has been given for DOM with *llamar* 'to call' in its literal meaning in (30) above (*llamar al ascensor* 'call the elevator').

Hence, an approach operating with (relational) proto-agent properties instead of animacy can reconcile two different cases of DOM with inanimate objects in Spanish: on the one hand, it captures cases in which the object is used in its literal meaning, as with machines like elevators that can obtain a certain degree of agentivity in certain events, and on the other hand, it can account for cases in which the object undergoes a conceptual shift from an inanimate to an animate entity.

Although I argued in favour of *relational* agentivity, it must be stressed that the mentioned implications from *inherent* proto-agent properties to animacy, as stated in (22) above, are neither considered invalid nor of less importance for an in-depth understanding of animacy-based DOM systems. Instead, we are dealing with two different, methodically contrary approaches to semantic roles as further specified by Jacob (2004): one that defines semantic roles as (verb-independent) apriori-deductive categories and one that treats semantic roles as (verb-dependent) empirical-inductive derivable properties. Models of generalized semantic roles such as Dowty (1991) or Primus (1999a, b) cannot be assigned to one of the two poles but can rather be characterized as a combination of both approaches. This is also argued to be the main theoretical problem of these accounts:

> Vom theoretisch-methodischen Standpunkt her dürfte auch in den rollensemantischen Ansätzen jüngeren Datums (Van Valin/LaPolla 1997, Van Valin *passim*, Dowty 1991, Primus 1999, etc.; vgl. den Überblick von Kailuweit 2004) das Hauptproblem in dem ungeklärten Verhältnis von apriorisch-deduktiven Komponenten der "Logical (!) structure" (man denke etwa an die Diskussion um die Rolle CAUS) einerseits und der Reminiszenz an empirisch-induktive Gegebenheiten der einzelsprachlich vorfindbaren Verben andererseits liegen. (Jacob 2004: 108)

While this problem is of less relevance when it comes to practice where often no consequent differentiation between both perspectives is taken (cf. Jacob 2004: 107f.), I think that it is crucial for the discussed relation between animacy and agentivity. The implication from proto-agent properties to animacy, such as EXPER(x) → ANIMATE(x), takes an apriori-deductive perspective on semantic roles: Proto-agent properties are regarded "as those cognitive-semantic properties that characterize how

a human individual, in particular oneself, is involved in situations" (Primus 2012a: 87). By abductive reasoning, an animate (or human) object participant is interpreted as potentially agentive in the same event. This serves as a plausible general explanation for the relation between animacy and (potential) agentivity and is supported by neurolinguistic evidence from language processing. However, it is not tangible enough to be applied to individual verbs in an empirical approach since it does not rely on lexical properties of verbs but on the (cognitive-) ontological status of an entity. In this way, it is hard to establish criteria according to which verbs, that for example preferably select a human direct object, can be classified. From an empirical-inductive perspective, by contrast, the (deductively derived) proto-agent properties, such as *sentience*, are regarded as lexical properties of verbs and are hence determined in relation to the event in question. This verb- (and language-)dependent view allows for a classification of predicates depending on if they for instance entail or presuppose an agentive direct object. It is also partly verb-dependent if a conversational implicature of sentience may be conveyed for the object: whenever the predicate entails proto-patient properties for its direct object and the object participant is at the same time animate as in (33a) below, such agentive interpretation of the object typically arises. In case an inanimate object is selected, such interpretation is ruled out (33b).[21]

(33) a. *Peter hit his colleague in front of the town hall.*

+> his colleague experienced pain (+> sentience)

∀x[CAUSE(x, y...) & ANIMATE(y) +> EXPER(y, x...)]

b. *Peter hit the bench in front of the town hall.*

#+> the bench experienced pain (#+> sentience)

To conclude, my account supports the view that operating with agentivity instead of animacy leads to a more comprehensive understanding of DOM patterns with both animate and inanimate objects, as shown by the example of Spanish. Crucially, in order to implement a role-semantic approach on DOM that takes verb meanings as the starting point, it is essential to disentangle relational, lexically encoded agentivity from inherent agentivity derived from the (cognitive-) ontological status of an animate being.

6.2.2 Agentivity and definiteness

In the following, I will deal with the relation between agentivity and definiteness. In general, the semantic-pragmatic concept of definiteness can be described as a factor that cross-linguistically determines DOM (cf. e.g. Bossong 1991, Aissen 2003, Witzlack-Makarevich/Seržant 2018). A crucial distinction has to be made between "individuality, which presupposes referentiality, on the one hand, and discourse-related

[21] Besides, note that in both (33a) and (33b) the conversational implicature of volition and sentience is conveyed for the subject.

definiteness on the other hand" (Bossong 1991: 159, cf. (34a)). As further noted by Bossong (1991: 159), each of these notions can be further subdivided. So, *definiteness* for example can involve different constellations of identifiability of the referent by speaker and hearer: (i) the referent is identifiable for both speaker and hearer, (ii) the referent is identifiable for the speaker only, (iii) the referent is identifiable for neither speaker nor hearer.[22] The ranking of the corresponding NP types can be exemplified by a version of the Referentiality Scale by von Heusinger (2008: 5) in (34b):

(34) a. The domain of reference (Bossong 1991: 160):

[+ individuality] > [± referentiality] / [± definiteness]

b. The referentiality scale (von Heusinger 2008: 5):

personal pronoun > proper name > definite NP > indefinite specific NP > non-specific NP > non-argumental

In languages in which DOM depends – fully or partly – on the definiteness of the direct object NP, the extension of marking, including the cut-off-point from obligatory to optional marking, is typically described by means of the referentiality scale in (34b).

With difference to the interaction between animacy and definiteness, whose impact on DOM is often studied (e.g. for Spanish), the interplay between agentivity and definiteness lacks a systematic investigation. In the acceptability judgement studies conducted for this work, definiteness was manipulated only in the study on Sicilian but not in the Catalan version. Instead, for the latter language definiteness was kept stable by testing only definite human NPs. The main reason that indefinite NPs were not tested for Catalan is that DOM is generally reported to be ungrammatical with indefinite NPs in colloquial Central Catalan (Escandell-Vidal 2009: 840). Interestingly, although DOM with definite human NPs is stated to be obligatory in the spoken variety (Escandell-Vidal 2009: 841), the acceptability judgement test revealed a main effect for *no-DOM* being significantly better accepted than *DOM* for both Barcelona and Girona. Such an effect has been attributed to the influence of Standard Catalan, which prescribes the use of DOM solely with personal pronouns (GIEC 2016).

The main finding of the Sicilian study was a triple interaction effect of *Agentivity*, *Referentiality* and *DOM* in the conditions *entailed agentivity & definite NP & DOM* and *presupposed agentivity & definite NP & DOM*. This suggests that *a*-marking is particularly favourable to highly prominent direct objects in terms of both the dimensions of agentivity and referentiality. Interaction effects of the single variables agentivity and referentiality with DOM, by contrast, were not attested. This was fairly surprising with regard to definiteness where acceptability ratings in the conditions

[22] As remarked by Bossong (1991: 159), "the fourth logically possible case of identification of the referee by the hearer alone does not seem to play any role in the grammatical organization of natural language; it occurs only in marginal discourse types like police interrogation".

definite NP & DOM, *definite NP & no-DOM*, *indefinite NP & DOM* and *indefinite NP & no-DOM* were almost numerically identical. Such a finding was unexpected since in similar acceptability studies for Spanish varieties a clear difference in acceptability between *definite NP & no-DOM* and *indefinite NP & no-DOM* has been attested (Caro Reina/García García/von Heusinger 2021: 20f.; Montrul/Bhatt/Girju 2015: 582). In these studies, the lack of DOM was rated considerably worse for definite NPs than for indefinite NPs following the referentiality scale. The full optionality for DOM and no-DOM in Sicilian, in contrast, deserves further research. At the same time, it is interesting that the interaction of referentiality and agentivity only holds true for definite NPs but not for indefinite NPs. As the results for the individual verbs revealed (cf. Appendix 1.4), there is even the reverse tendency that verbs that get higher ratings with definite NPs and DOM are less accepted with indefinite NPs and DOM, and vice versa. Therefore, it should be also considered whether different, possibly competing verbal factors might influence DOM, depending upon the referentiality of the direct object (definite vs. indefinite object NP).

So far, my focus has been on the role-semantic factor of agentivity and its relation to nominal factors characterizing the direct object participant, such as animacy and definiteness. From a broader perspective, that takes into account general motivations for the marker to appear, I considered both deviations from canonical thematic distinctness of subject and object in a transitive sentence (García García 2014) and deviations from natural transitivity (Comrie 1989). The common motive for DOM in these approaches is a distinguishing one, namely the disambiguation of subject and direct object of a transitive sentence. Yet, one crucial factor for DOM in Romance languages has been disregarded up to now, namely the topicality of the direct object. As shown in more detail in the next section, DOM especially in incipient stages often involves a deviation from the prototypical information-structural repartition of a sentence in that the object functions as primary topic and appears in (left-)dislocated position.

6.2.3 Interaction of factors in grammaticalization

After having discussed the relation of the role-semantic factor agentivity to the nominal factors animacy and definiteness, I shall address how agentivity could be integrated in a model of DOM in Romance languages. In addition to the synchronic perspective, I will argue that the consideration of agentivity may also lead to a better understanding of the diachronic development of DOM in Romance.

As a general finding, the acceptability judgements studies conducted on Sicilian and Catalan confirm the view that animacy is only to a limited extent a reliable criterion to predict DOM, previously put forward with regard to systematic instances of DOM with inanimate objects in Spanish (García García 2007, 2014). If we broaden our view on DOM in further Romance languages, there is more evidence for the assumption that animacy, also beyond role semantics, is neither a necessary nor a sufficient criterion to account for DOM to appear. Let us take the example of Corsican. As illustrated in (35a), personal names and personal pronouns receive the *a*-marker. The

same holds for kinship names like *mamma* 'mum' (Neuburger/Stark 2014: 377) as well as kinship terms lacking a definite article such as *me frateddu* 'my brother' (*aghju vistu à me frateddu* 'I have seen my brother', Ledgeway 2016a: 226). However, as can be seen by the following examples, animacy is neither a necessary nor a sufficient criterion for grasping the DOM pattern in Corsican: also place names (35b), which have an inanimate referent, are overwhelmingly *a*-marked (cf. Neuburger/Stark 2014: 376). Moreover, definite human NPs, marked with a definite article (35c) or with a demonstrative (35d), consistently reject the marker.

(35) a. personal name + personal pronoun [+ animate]:

Vegu *chè* *tù* *preferisci* *più* ***à*** *Peneloppe*

see.PRS.1SG that you prefer.PRS.2SG more DOM PN

chè ***à*** *mè.*

than DOM 1SG

'I see, you prefer Peneloppe rather than me.' (Neuburger/Stark 2014: 366)

b. place name [- animate]:

Vinz, […] decide *d'occupà* ***à*** *Calinzana.*

PN decide.PRS.3SG to occupy.INF DOM Calinzana

'Vinz, […] decides to occupy Calinzana.'

c. definite human NP [+ animate]:

Vigu ***ø/*à*** *l'omu.*

see.PRS.1SG ø/DOM the man

'I see the man.' (Marcellesi 1986: 137; Neuburger/Stark 2014: 376)

d. demonstrative NP [+ animate]:

Un tempu dopi, *i* *pastori* *trovanu* ***ø/*à*** *issu* *zitellu, […]*

a time later the shepherds find.PRS.3PL ø/DOM this child

'Some time later the shepherds find that child, […].' (Neuburger/Stark 2014: 379)

As this pattern seems to be stable in a corpus of written Corsican compiled by the authors, they assume that from a semantic-functional perspective, "DOM in Corsican is a marking strategy for highly contextually localized and therefore individuated referents, i.e., DOM in Corsican is perhaps a marker of contextual [individuation]" (Neuburger/Stark 2014: 385). It can be seen as a further indication that subconcepts of animacy, such as individuation, can account for distribution patterns of DOM rather than animacy in general, as represented by the animacy scale. Yet, the authors

indicate that this certainly has to be proved on a larger database, also taking into account spoken material and acceptability judgement data.

Besides nominal and verbal factors, the information-structural parameter topicality is discussed as a further predictor for DOM to appear in Romance languages, both from a synchronic and diachronic perspective. Here, we have to differentiate approaches that rely on the impact of topicality for the emergence and incipient stages of DOM in Romance languages (cf. Iemmolo 2010a; Melis 1995; Pensado 1995c) and those that still see topicality as a triggering factor in languages that have highly grammaticalized the marker (cf. Laca 1995; Leonetti 2004; Dalrymple/Nikolaeva 2011). The topicality of a direct object NP typically correlates with a peripheral syntactic position of the argument in question, being either left- or right-dislocated.

In order to see how topicality interacts with role semantics and referentiality in different Romance languages, let us consider some examples. As has been shown for Spanish (6.1.2), as far as verbs of competition are concerned, DOM with inanimate objects is optional. This optionality is illustrated in (36):

(36) *El entusiasmo vence* ***ø/a*** *la dificultad.*
the enthusiasm win.PRS.3SG ø/DOM la difficulty
'Enthusiasm overcomes difficulty.' (García García 2014: 185)

The optional occurrence of *a*-marking with *vencer* 'to win' can be accounted for by the role-semantic configuration: On the one hand, the direct object is causally affected by the subject participant and thus possesses a typical proto-patient property. This is why García García's (2007: 71, 2014: 145) generalization of thematic distinctness in the strict sense does not apply. On the other hand, the object is presupposed to be autonomously active in a previous subevent. This can be tested by its constancy under negation: *El entusiasmo no vence ø/a la dificultad* 'Enthusiasm does not overcome difficulty' states that there must have been a prior "fight" or conflict between both moods. The prototypical agent-patient asymmetry is hence present for *vencer* 'to win' but could be blurred as for the object proto-patient and presupposed proto-agent properties overlap. Interestingly, once the direct object is topicalized and appears in left-dislocated position, as in (37), DOM becomes obligatory.

(37) ***ø/A*** *la dificultad la* *vence* *el entusiasmo.*
ø/DOM the difficulty CL.F.3SG.ACC win.PRS.3SG the enthusiasm
'Difficulty is overcome by enthusiasm.' (García García 2014: 185)

As has been pointed out by García García (2014: 185), this obligatory marking is not due to the topicality of the object alone but seems to be the result of the interaction of the object's presupposed agentivity and its topicalization, the latter being syntactically mirrored by its movement to the left periphery and the presence of a clause-

internal resumptive clitic. That topicality alone is no sufficient criterion for DOM to appear is borne out by the following example:

(38) a. *El virus ha provocado* ***ø/*a*** *la enfermedad.*
the virus have.PRS.3SG provoke.PTCP ø/DOM the disease
'The virus has caused the disease.'

b. ***ø/*A*** *la enfermedad la ha provocado el virus.*
ø/DOM the disease CL.F.3SG.ACC have.PRS.3SG provoke.PTCP the virus
'The disease has been caused by the virus.' (García García 2014: 185)

The verb *provocar* 'to cause' selects for an effected object that can be classified as a maximal proto-patient since it is assigned the whole range of proto-patient properties. The clear-cut agent-patient asymmetry for this verb serves as an explanation of why *a*-marking of the direct object is ungrammatical both in canonical (38a) and in left-dislocated position (38b).

A close interplay between topicality and agentivity, along with referentiality, can be furthermore attested for DOM in Italian. While the marker is often said to be absent in Italian, we can find systematic occurrences in the colloquial spoken variety. As described in 6.1.3, DOM in colloquial Italian preferably occurs with left-dislocated first- and second-person pronouns of OE-psych verbs and in causative constructions (Berretta 1991: 221ff.). The impact of role semantics has been systematically tested in a recent acceptability judgement study on *a*-marked clitic-left-dislocated direct objects in colloquial Italian (Cassarà/Mürmann 2021). The study confirms that OE-psych verbs are indeed significantly better accepted than interaction verbs or highly transitive verbs. For the sake of clarity, a typical example is given again in (39):

(39) a. ***A*** *me, non (mi) convince questo*
DOM 1SG NEG CL.1SG convince.PRS.3SG this
'I am not convinced by this.'

b. **Me non mi convince questo*
1SG NEG CL.1SG convince.PRS.3SG this
'This does not convince me.' (Iemmolo in preparation: 249)

c. *Non convince* ***ø/*a*** *me*
NEG convince.PRS.3SG DOM 1SG
'It does not convince me.' (cf. Berretta 1991: 223)

The sentence in (39a), uttered in an informal context, contains the OE-psych verb *convincere* 'to convince'. The direct object, the first-person tonic pronoun *me* 'me', appears in left-detached position and is *a*-marked. The clause-internal resumptive clitic is optional in this case. As can be proven by (39b), DOM is obligatory under the given conditions. In postverbal clause-internal position, in contrast, the marker is ungrammatical (39c). Iemmolo (in preparation: 249f.) supposes a twofold information-structural function for DOM in Italian: on a sentence level, DOM marks the primary topic. This is a deviation from the prototypical information-structural repartition that predicts the subject is topical while the direct object is non-topical. In addition, the analysis of larger discourse units in spoken corpus data reveal that DOM in Italian typically carries a topic-shift function (cf. also Berretta 1991: 216). Importantly, for the purposes of this study, the example of Italian nicely illustrates that role-semantic factors can be also relevant in a language that exhibits DOM in an incipient stage. Yet, entailed agentivity for the object alone is not sufficient. Rather, it is the co-occurrence of high agentivity, referentiality and topicality that results in the obligatoriness of DOM.

Given these conditions for DOM in Italian, which in a similar way seem to apply to instances of the marker in colloquial and regional French,[23] Iemmolo hypothesizes DOM to be "a strategy that emerges in pragmatically and semantically marked contexts, namely personal pronouns in (mainly left) dislocation contexts" (2010a: 247). While he does acknowledge the affinity of OE-psych verbs and causative constructions to take DOM (2010a: 250, in preparation: 244), he does not derive verb semantic properties that could be entered into his approach. In Cassarà/Mürmann (2021), it has been claimed that the consideration of role-semantic factors on a par with referentiality-based and information-structural constraints could lead to a more comprehensive picture of the phenomenon in Italian. Accordingly, the emergence of DOM with OE-psych verbs can be accounted for by their non-prototypical role-semantic configuration: on the one hand, the direct object is always assigned the proto-agent property sentience. On the other hand, the agent-patient asymmetry between subject and object may be absent, as far as the non-causative reading of these verbs is concerned (e.g. *Peter's reaction disturbed Maria*, where an agent-patient asymmetry is lacking, is more unprototypical than *Peter (intentionally) disturbed Maria*).

As further indicated by Iemmolo (2010a: 248) on the basis of Dalrymple/Nikolaeva (2011: 194ff.), two main paths of grammaticalization are conceivable for incipient DOM constructions in left- (or right-) dislocation contexts:

(40) a. DOM may be extended to non-topical objects which share features of topicworthiness, as in Modern Sicilian. In this way, the link with information structure is lost or at least weakened.

b. DOM may be restricted to topical objects only, as in [Balearic] Catalan.

[23] Note that unlike in Italian, *a*-marking in French seems to mainly occur with right- rather than left-dislocated direct objects that are at the same time accompanied by a clause-internal resumptive clitic (e.g. *[il] nous escroquait pas à nous* '[he] wasn't scamming us', Fagard/Mardale 2014: 153).

The first path not only corresponds to the situation in Modern Sicilian but also to Modern Spanish as well as to (spoken) Peninsular Catalan. In all these languages or varieties, the connection between DOM and topicality is weakened in that direct objects can also get *a*-marked in canonical, that is postverbal (non-topical) position. The second path is exemplified by Balearic Catalan where only pronouns are marked in situ, while all other types of object NPs can only get DOM if occurring in left- or right-dislocated structures as shifting or continuing topics (cf. Escandell-Vidal 2009).

In order to systematize the different dimensions of factors that are reported to correlate in incipient stages of DOM, I propose to make use of Himmelmann's (1997: 28f., 2004: 32f.) model of grammaticalization. This model defines grammaticalization as "context expansion" of a construction that take places on the following three levels: host-class formation ("host-class expansion"), change of syntactic context ("syntactic context expansion") and change of semantic-pragmatic context ("semantic-pragmatic context expansion"). As shown below, the multiple constraints on incipient DOM, as described by the example of Italian, can be classified along these levels. Host-class constraints concern object the NP type (1a) and verb type (1b). Syntactic constraints refer to the position of the direct object (2). Semantic-pragmatic constraints are the referentiality of the direct object participant (3a) and its agentivity relative to the event (3b). Note that information-structural constraints, here the topicality of the direct object (4), have been added as a fourth level.

(1) Host class constraints (Object NP type and Verb type):

a. Object NPs:

deictic/1,2 pronoun > 3 pronoun > proper name[24] > def. NP > indef. NP
(cf. Bossong 1991)

b. Verbs which encode the following O properties:

entailed agentivity > presupposed agentivity > potential agentivity > unspecified for agentivity

(2) Syntactic constraint (Object position):[25]

[24] Caro Reina (2020) proposes a modification of the extended animacy hierarchy of Silverstein (1976) in which he distinguishes different subtypes of proper names. This finer distinction is a good indicator for the stage of development of DOM in Romance languages: As has been shown above for Corsican, place names, which are inanimate, also regularly get *a*-marked (Neuburger/Stark 2014: 376). At the same time, DOM is ungrammatical with human definite NPs being accompanied by the definite article, a possessive or a demonstrative. This can be compared to Old Spanish where in contrast to Modern Spanish, place names were *a*-marked, whereas human definite NPs were only optionally marked (Caro Reina 2020: 249ff.). Once animacy grammaticalized further as a proper semantic factor for DOM in Spanish, place names at some point stopped taking the marker (cf. also Kabatek 2016: 230). The retraction of DOM with inanimate proper name classes can thus be seen as an indicator that there is a competing factor, namely animacy, which gains in importance.

[25] Besides the dislocation of the object, there are other syntactic constraints for the occurrence of DOM in early stages as for example coordination, secondary predication or a pre-posed object (cf. e.g. Delille 1970: 36ff. for diachronic Portuguese; Pineda in press: 57ff. for diachronic Catalan). Since dislocation is the prevalent syntactic constraint, I listed only this one.

Non-canonical (peripheral) position (LD or RD) > canonical (clause-internal) position (SVO)

(3) Semantic-pragmatic constraints (Referentiality and Agentivity status of the O):

a. Referentiality:

O inherently definite and referential (= individuation) > not inherently definite and referential but discourse-related definiteness (= definiteness) (cf. Bossong 1991)

b. Agentivity:

O agentive (= agentivity) > non-agentive but potentially agentive (= humanness/animacy)

(4) Information-structural constraint (Topic status of the O):

O Topic > non-topic but topicworthy (animate and identifiable)

Certainly, the factors displayed for the individual dimensions are only preliminary and need to be further modified. It goes beyond the scope of this work to enlarge upon grammaticalization theory and a possible grammaticalization process of DOM in different Romance languages. In any case, extensive (diachronic and synchronic) corpus studies are needed to precisely capture the emergence of DOM and to clarify the claimed role of agentivity as an initial triggering semantic factor.

In summary, the integration of agentivity into a model on DOM in Romance could yield new insights on the grammaticalization of the phenomenon. From a synchronic perspective, a look at languages where DOM is in an incipient stage, such as colloquial Italian or French, could enhance the understanding of the interplay of role-semantic, referentiality-based and information-structural constraints on *a*-marking. At least for Italian, there is evidence to claim that DOM emerges in contexts where a deviation from the prototypical direct object in all three dimensions coincides, that is to say a pronominal object occurring in (left-)detached position for which agentivity is entailed by the predicate. Taking into account the diachronic development of DOM in languages that at present show a medium level (Sicilian, Central Catalan) or high (Spanish) degree of grammaticalization could reveal if agentivity was already among the initial triggering factors. If this turns out to be the case, it would further underpin the view defended in García García (2014: 224) for Spanish DOM, stating that the impact of animacy is only a symptom of, and hence epiphenomenal to, agentivity. The relevance of animacy as a factor for DOM would then only be the grammaticalized result of the impact of its linguistically manifested sub-concepts such as individuation, person, topicality and (relational) agentivity.

6.3 Summary

Based on the findings of the acceptability studies on Sicilian and Catalan, this chapter further discussed the role of agentivity for DOM in Romance. As suggested by the results obtained for the two languages, the gradual communicated agentivity for the

direct object should be considered a relevant verbal factor for *a*-marking. The presented account can be regarded as complementary to the generalization of thematic distinctness developed for DOM with inanimate objects in Spanish (García García 2007, 2014) in that it adds a further dimension that makes a fine-grained distinction of agentivity levels for the direct object. In this way, more DOM-sensitive predicates in Spanish could be explained within a role-semantic model including a good number of those that are reported to obligatorily take DOM with indefinite human NPs. Beyond this, diachronic and synchronic evidence from DOM-sensitive verb classes in other Romance languages further supports an impact of agentivity on *a*-marking. Another focus was on the question of how agentivity interacts with animacy and definiteness, respectively, in determining DOM. With regard to animacy, it was argued that a notion of relational agentivity is preferable over one of animacy or inherent agentivity since only the former can capture *a*-marking patterns both with animate and inanimate objects triggered by role-semantic properties of the predicate in question. Concerning definiteness, which had been shown to interact with agentivity for Sicilian, further systematic studies are needed that prove the claim that the interplay of high referentiality with high agentivity especially favours *a*-marking. Finally, it has been proposed how agentivity could be integrated into a model that describes the interaction of the mentioned factors in grammaticalization. Applying this model to individual Romance languages – preferably from both a synchronic and diachronic perspective – could advance our understanding of the role of agentivity for the emergence and development of the phenomenon and its relation to referentiality-based and information-structural factors.

7. Conclusion

In the following, I will provide a summary of the main findings of this study and put them into the overall context of differential morphosyntactic argument realization in general and DOM in Romance in particular. I will furthermore review the two main functions discussed for DOM in light of the results and delineate further research questions that can be derived from the empirical findings and theoretical implications of this study.

In the introduction, I formulated three main research questions that are taken up again to serve as a guide for the structure of the summary. They are repeated here for convenience and will be answered in sequence.

1. DOM from a linking perspective (RQ 1)

   How can we include DOM in a broader theory on linking? In particular, how can we capture proto-agent properties of the direct object in more precise terms in order to clearly disentangle agentivity from animacy?

2. DOM and acceptability in Sicilian and Catalan (RQ 2)

   Why are acceptability judgement tasks of particular relevance for a synchronic view of DOM? Is there a difference in acceptability of DOM between different degrees of communicated agentivity in the two languages?

3. Implications for a model on DOM in Romance (RQ 3)

   What do the results for Sicilian and Catalan tell us about the impact on the object's agentivity on DOM and its interaction with nominal factors? Can they be related to findings for DOM in Spanish? As for DOM in Romance in general, how can we model the impact of role-semantic factors and the interaction with other factors?

As for **RQ 1**, I have argued that DOM can be captured as a linking phenomenon. Just as non-differential morphosyntactic case selection and argument selection, DOM can be described within a framework of generalized semantic roles (Dowty 1991, Primus 1999a, b, 2006). One main achievement of these approaches is that they provide general linking principles that allow us to predict (morpho-)syntactic argument realization by the relative number of proto-agent and proto-patient properties of the co-arguments of a transitive sentence. The approach by García García (2007, 2014) on DOM with inanimate objects in Spanish shows that such models can be successfully applied to differential morphosyntactic argument realization: He establishes the *generalization of thematic distinctness* that predicts DOM will occur if the subject does not outrank the direct object in terms of agentivity, i.e. when the direct object either bears more proto-agent properties than the subject or when both co-arguments exhibit the same number of proto-agent properties. In order to apply such a linking principle to DOM with animate objects, further theoretical work was needed in a first step since the role-semantic property agentivity had to be clearly dissociated from the object NP-inherent property animacy. For this purpose, the proto-role model of Blume (2000) was considered particularly suitable. The model is a modified version

of Dowty's (1991) and Primus' (1999a, b) proto-role model and should especially account for the selection of marked valencies, such as */nom/dat*, from a cross-linguistic perspective. In comparison to the former two models, Blume (2000) makes the following two refinements, which are essential for this study: First, she not only takes entailed but also presupposed and conversationally implicated proto-agent properties for the object participant into account. Importantly, they can all affect the selection of marked valencies; presupposed proto-agent properties, for instance, are central for explaining the linking of interaction verbs such as 'help' or 'obey' cross-linguistically (Blume 1998). Second, she distinguishes between potent and non-potent proto-agent properties. Whereas potent proto-agent properties, such as control, are linked to the nominative by default, non-potent proto-agent properties, such as *sentience*, can be linked to any argument or morphosyntactic case. Sentience is thus the proto-agent property that is most likely to occur in object position. On the basis of Blume's model, it is possible to draw a fine distinction between three different types of communicated agentivity for the object: *entailed agentivity*, *presupposed agentivity* and *potential agentivity*. While the first two can be deduced from the lexical semantics of the predicate, the third is conveyed by generalized conversational implicature, which speakers are assumed to make based on the verb meaning and the animacy of the object participant. Together with the category *unspecified for agentivity*, indicating that the direct object is not assigned proto-agent properties of any kind, these levels were arranged on a four-point scale. For each of the four levels, semantic diagnostics were presented that allow us to clearly differentiate them and, in a further step, attribute verb classes to them. The type of communicated agentivity by the object can be partially cross-classified with the agent-patient asymmetry, i.e. the relative distribution of proto-properties between subject and object, illustrated again in Table 1. As can be seen, the first two classes do not form a homogeneous class with regard to the thematic asymmetry between subject and object: while for *entailed agentivity*, there are three different configurations, it is two for *presupposed agentivity*. For the classes *potential agentivity* and *unspecified for agentivity*, only a configuration where the subject outranks the object in terms of agentivity is conceivable.

COMMUNICATED AGENTIVITY FOR O → ASYMMETRY ↓	entailed (e.g. *to fascinate, to resemble, to frighten*)	> presupposed (e.g. *to call, to greet, to help, to defeat*)	> potential (e.g. *to injure, to insult, to kill*)	> unspecified (e.g. *to admire, to see*)
S *less* agentive than O	+	-	-	-
S as agentive as O	+	+	-	-
S *more* agentive than O	+	+	+	+

Table 1. Cross-classification of agentivity scale (horizontal) and agent-patient asymmetry (vertical)

The red-marked fields represent the role-semantic configurations that, according to Blume (2000: 200), license the selection of marked valencies. The main criterion put forward in her licensing condition is the non-asymmetric assignment of proto-roles in the argument relation. It is important to note that the majority of DOM-sensitive verb classes identified by García García (2007, 2014) for Spanish, such as reversible predicates (*preceder* 'to precede', *seguir* 'to follow', *sustituir* 'to substitute') or verbs of attribution (*caraterizar* 'to characterize', *definir* 'to define') arguably fulfil the same role-semantic criteria that license marked case frames. In the present work, it is the blue-marked fields that were of special interest. Here, with the exception of the last-ranked class *unspecified for agentivity*, an agent-patient asymmetry between subject and direct object is always given. As a consequence, marked valencies are not licensed, the relevant verb classes are predicted to select for a transitive case frame. Yet a particularity is that proto-patient properties for the direct objects of the three first classes overlap with at least one proto-agent property. This role-semantic configuration has not gone unnoticed in previous work but has never been systematically described for two-place predicates. Primus (1999a: 54ff., 1999b: 145) introduces a third generalized semantic role, namely a proto-recipient that is characterized by an accumulation of both proto-patient and proto-agent properties (*sentience* or *possession*). From a linking perspective, proto-recipients are often realized by morphological datives or indirect objects. In Primus (2012a: 73) she applies the hybrid character of proto-recipients to *a*-marked direct objects in Spanish stating, "[a]nimacy-driven DOM involves a similar role overlap. It occurs whenever a noun phrase subcategorized for patient properties acquires potential agent properties due to its intrinsic meaning". Blume (2000: 214-217), in her cross-linguistic study of marked valencies, reports the finding that so-called overgeneralizations of the licensing condition for marked valencies can be attested when an asymmetric distribution of proto-roles is indeed present but the object bears proto-agent properties at the same time. Furthermore, García García (2014: 178-191) finds that verbs of competition (e.g. *vencer* 'to win', *derrotar* 'to defeat') show a preference for DOM with inanimate objects in Spanish. For these verbs, the subject clearly outranks the object in terms of agentivity. However, the direct object is presupposed to be agentive in a previous subevent, which seems to be morphosyntactically reflected.

Put briefly, the agentivity scale introduces a further dimension in which the role-semantic configuration could deviate from a prototypical transitive scenario, i.e. when the subject has potent proto-agent properties and the direct object exhibits proto-patient properties but in addition to that, the latter also bears proto-agent properties. While this scenario does not affect linking in a narrow sense – the selected case frame should typically be */nom/acc* – it is argued to influence the surface expression of the direct object as either morphologically marked or unmarked. Since a verbal entailment is by definition stronger than a presupposition or a generalized

conversational implicature, I hypothesized DOM would be more likely to occur the more a verb class can be graded at the left of the scale.

As for **RQ 2**, I presented an acceptability judgement study that investigated the impact of role-semantic factors on DOM in Sicilian (variety of Alcamo, Western Sicilian) and Central Catalan (varieties of Barcelona and Girona). On the basis of the agentivity scale (cf. (1)) four verb classes were defined, i.e. one class for each type of communicated agentivity for the object (*entailed agentivity*, *presupposed agentivity*, *potential agentivity*, *unspecified for agentivity*). In order to prevent an effect of the relative distribution of proto-properties among co-arguments, the primary focus was on role-semantic configurations where the subject was more agentive than the direct object (the blue-marked fields in Table 1).[1] The animacy of subject and direct object in the test items was kept constant in that only human subjects and direct objects were chosen. For Sicilian, referentiality was added as a further independent variable, i.e. each verb was tested with a definite as well as with an indefinite human NP. The choice of gradient acceptability judgements as empirical method was due to the fact that it offers a direct comparison of two competing forms, one with DOM and one without DOM, for each predicate. In addition, testing five lexicalizations for each of the four verb classes is a challenging task in a corpus since we do not necessarily attain a balanced number of tokens for each verb in question. Yet, most importantly, larger corpora of spontaneous speech are lacking for Sicilian.

The results of the acceptability judgement tests revealed clear differences of acceptability between the four agentivity classes in both Sicilian and Catalan. However, the measured effects were quite different for the two languages under investigation. For Sicilian, Hypotheses 1 (Agentivity & DOM) and 2 (Referentiality & DOM) have not been confirmed, that is no effects of the single independent variables agentivity and referentiality in interaction with DOM have been found. Hypothesis 3 (Agentivity & Referentiality & DOM), conversely, was partly confirmed: In condition 3a (*Agentivity & definite NP & DOM*) a triple interaction effect was found. More precisely, for definite object NPs, the presence of DOM was significantly better accepted with verbs that either entail (e.g. *affènniri* ‘to offend’) or presuppose (e.g. *aiutari* ‘to help’) agentivity for their direct object than with verbs for which the object’s agentivity is only conveyed via conversational implicature (e.g. *firiri* ‘to injure’):

(1) Sicilian: Effects in condition 3a (*Agentivity & definite NP & DOM*):

a. entailed agentivity & definite NP & DOM > potential agentivity & definite NP & DOM

[1] This is true with one exception: for the class of presupposed agentivity, two verbs have also been included that show a symmetric argument relation between subject and direct object: ‘help’ and ‘hear’. I think, though, that this inconsistency in the study design would, if at all, only have a minor effect on the results since relative linking can unambiguously take place for interaction verbs like the two mentioned: the argument bearing *entailed* proto-agent properties, i.e. the subject of ‘help’ and ‘hear’, is cross-linguistically linked to the nominative by default, while the argument with *presupposed* proto-agent properties is always syntactically realized as the object (Blume 1998: 269, 2000: 224).

b. presupposed agentivity & definite NP & DOM > potential agentivity & definite NP & DOM

By contrast, in the same condition no effect was found with respect to the lowest ranked class of the agentivity scale (*unspecified for agentivity & definite NP & DOM*). The shown effects were not mirrored in the reverse condition 3b (*Agentivity & definite NP & no-DOM*). Moreover, the predicted ranking of agentivity classes was rejected for indefinite object NPs in the conditions 3c (*Agentivity & indefinite NP & DOM*) and 3d (*Agentivity & indefinite NP & no-DOM*). The results for Sicilian thus point to an impact of agentivity on DOM only in interaction with definite, i.e. referentially prominent object NPs, while no such impact could be identified in interaction with indefinite, i.e. less prominent object NPs.

For Catalan, for which only definite NPs were tested, Hypothesis 1 (DOM vs. no-DOM) was confirmed: For both Barcelona and Girona, the absence of DOM was significantly better accepted than the presence of DOM. Hypothesis 2 (Agentivity & DOM) was partly confirmed. While for the presence of DOM no predicted interaction effect with agentivity in neither of the two datasets was found (condition 2a, *Agentivity & DOM*), for the absence of DOM my predictions have been partly confirmed (condition 2b, *Agentivity & no-DOM*). For both varieties, *unspecified for agentivity & no-DOM* was significantly better accepted than *entailed agentivity & no-DOM*. In addition, for Barcelona *presupposed agentivity & no-DOM* outranked *entailed agentivity & no-DOM*, while for Girona *unspecified for agentivity & no-DOM* got higher acceptability scores than *potential agentivity & no-DOM*.

(2) Catalan: Effects in condition 2b (*Agentivity & no-DOM*):

a. Barcelona:

(1) presupposed agentivity & no-DOM > entailed agentivity & no-DOM

(2) unspecified for agentivity & no-DOM > entailed agentivity & no-DOM

b. Girona:

(1) unspecified for agentivity & no-DOM > entailed agentivity & no-DOM

(2) unspecified for agentivity & no-DOM > potential agentivity & no-DOM

Thus, across varieties, verbs that do not encode any agentivity information for the object (*unspecified for agentivity*) receive higher acceptability values in the no-DOM condition than those which lexically encode agentivity as a verbal entailment of the predicate (*entailed agentivity*). While the results for both languages do indicate an impact of agentivity on the acceptability of DOM – for Sicilian proved only for the presence and for Catalan solely for the absence of the *a*-marker – there was also counterevidence to the agentivity scale for both Sicilian and Catalan. As explained in more detail in section 5.3, these contrary findings can be potentially ascribed to competing verbal factors deserving further investigation.

From a methodological point of view, the obtained acceptability data offer a detailed picture of the impact of agentivity on the perceived grammaticality of DOM and no-DOM in both languages. Note, however, that given these findings no conclusions can be drawn with respect to the actual *use* of the *a*-marker. This seems to hold true especially for Central Catalan for which the results have to be seen in the light of the sociolinguistic context: DOM is a phenomenon that occurs in the colloquial spoken language only. Prescriptive grammar bans the use of the marker referring to DOM as a Castillanism and allows only a few exceptions. One is the use of DOM with personal pronouns, the only NP type with which the marker is declared to be obligatory in Standard Catalan. Although participants were given spoken target sentences and instructed to imagine an informal situation from everyday life, they might still have been biased by the norm in their acceptability judgements (cf. Adli 2015 for a more differentiated view on a similar finding for French).

As for RQ 3, the results of the conducted studies suggest at least a partial influence of the role-semantic factor agentivity on DOM. The fact that the four verb classes were tested both in the DOM and in the no-DOM condition allowed me to capture very subtle differences in acceptability. For Sicilian, as far as definite object NPs are concerned, verbs that encode the object's agentivity in their lexical semantics received significantly higher acceptability values with DOM than those for which agentivity is solely pragmatically conveyed. More specifically, predicates either entailing or presupposing the object's agentivity (e.g. *affènniri* 'to offend' or *sentiri* 'to hear') clearly outranked verbs conveying agentivity only via conversational implicature (e.g. *feriri* 'to injure'). For Catalan, it was the absence of DOM that revealed agentivity effects. For both varieties, verbs that are unspecified for agentivity showed significantly higher acceptability scores than verbs that entail agentivity for their object. That is, for verbs that do not assign any kind of proto-agent properties to their object (e.g. *conéixer* 'to know') an omission of the *a*-marker is better accepted than with verbs that encode the object's agentivity their verbal semantics (e.g. *ofender* 'to offend'). Moreover, for Barcelona, also a difference between predicates for which the object's agentivity is entailed as opposed to such for which it is presupposed (e.g. *salvar* 'to save') was detectable with the latter outranking the former. For Girona, conversely, a further effect of predicates unspecified for agentivity being better accepted without the marker than those with potential agentivity, which pragmatically convey sentience (e.g. *curar* 'to cure), supports the proposed ranking. Comparing the findings for both languages, it is particularly verbs that entail agentivity for their object, more specifically the proto-agent property sentience, that seem to have generalized the *a*-marker, as predicted by the agentivity scale. While for Sicilian, this is indicated by high acceptability values for the *presence* of DOM with this class, for Catalan it is deduced as an explanation for the low ratings in case of the *absence* of DOM with these verbs.

As for the relation between agentivity and the nominal property animacy, the verb-based account offered a more detailed perspective than the general criterion of (inherent) potential agentivity. Since the animacy of both subject and direct object was

kept stable for all test items, the results can be interpreted as being clearly independent of the object's animacy. Yet, it must be stressed that they are not independent of the object's definiteness: The interaction effect measured for Sicilian revealed that the object's agentivity *per se* is not decisive but its interplay with an object NP highly ranked in the referentiality scale, namely a definite NP. Whether the same interaction of factors also applies to Catalan, where only definite NPs were tested, must be left to future research. I argue that my findings can be integrated into the role-semantic explanation on DOM given by García García (2014) for DOM with inanimate objects in Spanish. It has been proposed to add a further dimension to the thematic asymmetry between subject and object, namely the communicated agentivity for the direct object. The relation of agentivity to other factors, such as definiteness and topicality, showing up for DOM in Romance languages is delineated in section 6.2.3. As suggested by languages that exhibit DOM in an incipient stage like colloquial Standard Italian, role-semantic factors should be treated on a par with referentiality-based and information-structural constraints. Put briefly, I have argued that the non-prototypicality of the object on all three levels contributes to its preference to obtain the marker.

In the following, I will summarize what my findings contribute to the question of the general motivation of DOM. As mentioned earlier, there are two main functions put forward for case marking in general, which have also been proposed as explanations for systems of differential case marking: the *distinguishing* (or *discriminatory*) *function* and the *indexing function*. According to the distinguishing function, direct objects are morphologically marked in order to distinguish them from the subject of the sentence and thus avoid ambiguity of the co-arguments (cf. e.g. Bossong 1991; Aissen 2003). The indexing function, in contrast, assumes that languages formally mark direct objects that exhibit a high degree of transitivity (cf. e.g. Hopper/Thompson 1980; Næss 2004). While both distinguishing and indexing function fundamentally differ in the kind of explanation they provide for case marking to appear, they make roughly the same predictions for the occurrence of DOM, namely that direct objects bearing a high degree of animacy and definiteness are preferably marked. DOM is thus argued to be consistent with both functions (Malchukov 2008: 209-211; de Swart 2007: 127f.; Primus 2012a: 84). Given neither of the two functions offer a satisfactory explanation for systematic occurrences of DOM with inanimate objects in Spanish, García García (2014) proposes a revised distinguishing approach that accounts for DOM in terms of role semantics. It takes the lack of thematic distinctness of a prototypical transitive sentence as the main criterion; the insertion of the *a*-marker thus functions as a formal means of *actual* role-semantic disambiguation between subject and direct object. While the generalization of thematic distinctness in the strict sense does not apply to DOM with animate objects in Spanish, it has been proposed that although role-semantic distinctness is given, the direct object can be analyzed as *potentially* agentive in the event denoted by the predicate due to its intrinsic meaning properties (Primus 2012a: 75). Hence, in these cases DOM would function as a means of *potential* role-semantic disambiguation.

The present study captures animacy-based DOM in more precise lexical role-semantic terms. A detailed analysis of transitive predicates that exhibit an agent-patient asymmetry and preferably select for an animate object has revealed that proto-patient properties for the direct object may overlap with a proto-agent property. This communicated agentivity for the direct object may be conveyed by verbal entailment, presupposition or generalized conversational implicature. The relevant proto-agent property is *sentience*, referred to as a non-potent proto-agent property by Blume (2000: 137) due to its subordinated role for morphosyntactic linking. Although *sentience* is classified as a non-potent proto-agent property, its presence in object position overlapping with proto-patient properties can be considered a deviation from the prototypical transitive scenario in which subject and direct object are expected to be maximally distinct, i.e. the subject bears a maximal number of proto-agent and the direct object a maximal number of proto-patient properties. Any deviation from this configuration can be argued to blur thematic distinctness between the two arguments.

At the same time, my proposal is not generally incompatible with an indexing function for DOM. Verbs such as causative OE-psych verbs (e.g. 'frighten') are close to canonical transitive verbs and can unambiguously be linked to a transitive case frame. As a particularity, the overlap of affectedness and sentience for the direct object might open a further dimension of affectedness, namely a mental one. Note that in this sense, the relevance of sentience for DOM alleged here can be reasonably connected to the explanation of indexing approaches of why the object's animacy matters for DOM:

> […] affectedness correlates with the degree of P's prominence/individuation, and hence indirectly with its animacy features. All other things being equal animate Ps are conceived as more affected, as compared to inanimates (cf. Wierzbicka, 1981; Næss, 2004). The reason for this connection, admittedly more subtle, seems to be that inanimates are affected inasmuch as they undergo a change of a physical state in the course of an action, while animates can also **undergo a change of a mental/emotional state** to qualify as affected. (Malchukov 2008: 211; emphasis: S.M.)

The crucial difference is that the present study does not consider mental affectedness to increase the patienthood of an object but, quite the contrary, to decrease it since it is derived from the proto-agent property sentience that overlaps with the proto-patient property affectedness. To sum up, the consideration of proto-agent properties that co-occur with proto-patient properties for the direct object is suitable to draw a link between cases in which DOM can be ascribed an actual distinguishing function and those cases in which DOM is obligatory despite an agent-patient asymmetry. Crucially, both cases share one central property, namely the relative agentivity of the direct object in the given event.

It would be interesting to explore how a role-semantic approach on DOM might relate both distinguishing and indexing function to each other. One plausible assumption is that an indexing function developed out of a mere strategy of distinguishing. A similar connection of both functions has been proposed by de Swart (2007: 130-133) for DOM in Spanish, though for the nominal property of animacy. He shows

that while DOM in initial stages can be proved to serve as a means of disambiguation ("recoverability") between subject and object, the usual obligatory marking of animate objects in Modern Spanish can be subsumed under an indexing strategy ("prominence marking"). Such a change in function might be taken into account from a role-semantic perspective as well. As has been shown in section 6.1.3, in languages with incipient DOM such as Italian, non-causative OE-psych verbs (e.g. *preoccupare* 'to worry') for which the object outranks the subject in terms of agentivity are considerably more attested and accepted with the marker than highly transitive verbs, even if the latter convey sentience for the object via implicature (e.g. *ferire* 'to injure').

Though the studies on Sicilian and Catalan suggest a systematic influence of agentivity on DOM, much more evidence is needed to prove the robustness of the presented hypothesis. This should comprise mainly three types of language data. First, as has been previously mentioned, synchronic follow-up studies on the investigated Sicilian and Catalan varieties should be carried out. Ideally, as a next step, the judgement data should be complemented with production data for both Central Catalan and Sicilian by, for example, eliciting the very same sentences in an informal setting. A further step would be a corpus analysis of a corpus of spontaneous speech, which would reveal insights on the frequency of the relevant forms. This is necessary as previous studies show that the perceived grammaticality of a linguistic form does not have to correlate with its frequency in a corpus and vice versa (cf. i.a. Adli 2015; Bader/Häussler 2010). Moreover, it would be interesting to take micro-variation into account as well and apply the acceptability judgement tests to other varieties of Catalan (e.g. Valencian or Balearic Catalan) and Sicilian (e.g. Central Metafonetica or Southeast Metafonetica). As it is well known, DOM is a phenomenon that occurs across varieties in both languages. It is unknown, however, whether the impact of role-semantic factors is relatively stable or varies a great deal as far as different varieties are taken into account. Importantly, these tests have to respect variety-specific characteristics of DOM such as the cut-off point on the animacy and referentiality scale as well as peculiarities in sentence position of marked objects. Accordingly, in Valencian Catalan where the distribution of DOM is as widespread as in European Spanish with human definite NPs being obligatorily marked (Sancho Cremades 1995: 199f.), indefinite human NPs where DOM is reported to be optional should be tested. For the same reason, in Balearic Catalan indefinite human NPs should also be tested and, since DOM occurs only in left- or right-detached position (Escandell-Vidal 2009), the sentence structure of the test items has to be adapted. Second, future research has to prove the hypothesis for a larger set of Romance languages. A special focus should be put on Romance languages such as Italian or French that display DOM in an incipient stage (cf. e.g. Berretta 1991; Fagard/Mardale 2014; Iemmolo 2010a). The fact that the marker has not been grammaticalized yet in these languages allows for a detailed analysis of role-semantic, referential and syntactic factors that have to interact systematically for the marker to occur. Taking the example of colloquial Standard Italian, a direct object must be non-prototypial in all three dimensions to be compulsatorily *a*-marked, that is a pronominal left-dislocated direct object of an OE-psych verb (cf. Cassarà/Mürmann 2021). These initial multi-dimensional

factors are difficult to disentangle in Modern Spanish since almost every human definite direct object is *a*-marked. However, in order to gain a more complete picture, it would be desirable to test the verb classes of the agentivity scale for Spanish, as well. To minimalize the risk of ceiling effects, an object NP type should be chosen for which DOM is optional, for example human indefinite NPs. Third, and closely linked to the previous point, it is indispensable to investigate the impact of agentivity on DOM from a diachronic perspective, too. As delineated in section 6.2.3, the assumption that DOM spreads along the agentivity scale implies that role-semantic factors might have been already present at the emergence of the phenomenon. Here, diachronic corpus studies on the Romance languages in question could yield new insights. Old Sicilian is a case in point: There are some indications that DOM was favoured with verbs selecting a non-prototypical object (cf. La Fauci 1991; Sornicola 1997; Volo 2010). La Fauci (1991) analyzed four Old Sicilian texts of the *Confessionali*[2] dating from the end of the 13th century until approximately 1480. In his argumentation, DOM correlates with verbs that are thematically "dative". Interestingly, a good number of these verbs take an agentive object (e.g. *ingannari* 'to deceive', *offendiri* 'to offend', *charmari* 'to charm') or an object with presupposed agentivity (e.g. *honorari* 'to honour', *perdunari* 'to forgive', *laudari* 'to praise'). Yet these indications have to be systematically proved in an empirical study.

Adopting a broader perspective, the agentivity scale introduced in this work, cross-classified with the agent-patient asymmetry between subject and object, can capture subtle differences in semantic transitivity. Taking the communicated agentivity of the object as the decisive criterion, it allows for a detailed classification of two-place predicates that, roughly speaking, range from lower to higher semantic transitivity.[3] Crucially, the definition of subclasses is solely based on role-semantic criteria and is thus motivated independently of the preference of certain predicate classes to select for DOM or marked valencies.

Gradience is a core characteristic of approaches to prototypical semantic transitivity (e.g. Hopper/Thompson 1980; Næss 2007). Hopper/Thompson (1980: 252f.) define transitivity as a multifactorial phenomenon that comprises verbal features as well as inherent properties of the object, i.e. its degree of individuation. The more features of high transitivity are given in a sentence, the more transitive it is, while no relative ranking of the individual features is made. Næss (2007: 44) defines the transitive

[2] Edited by F. Branciforti in the third volume *Rules, constitutions, confessionals and rituals* of the Collection published by the Centre for Philosophical and Linguistic Studies of Sicily.

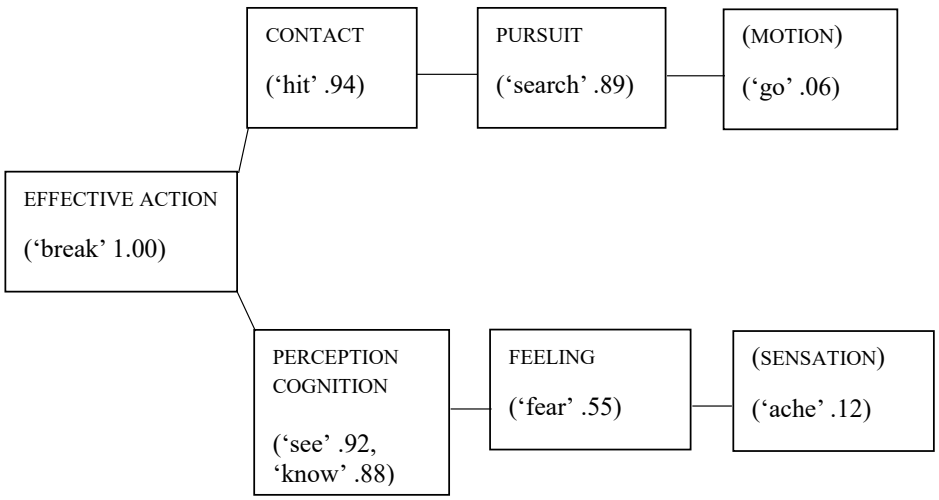
[3] Note, however, that the agentivity scale, cross-classified with the agent-patient asymmetry, does not rank semantic transitivity *as such* (from low to high). What matters is the presence of proto-agent properties for the object. As a result, the lowest-ranked class (*unspecified for agentivity*) also may include verbs that have a low degree of semantic transitivity, such as SE-psych verbs (e.g. Spanish *querer* 'to like') or possession verbs (e.g. *tener* 'to have') for which the subject only exhibits the proto-agent property of sentience. Aptly enough, these predicates typically have a lexical converse with the inverse linking pattern, i.e. where sentience is assigned to the object (in this case, of the OE-psych verb *gustar* 'to like' or the possession verb *pertenecer* 'to belong to'). This is predicted by Blume's (2000: 224) linking principle, stating that for verbs that only assign non-potent proto-agent properties to their arguments linking cannot be determined.

prototype by the maximal distinctness of participants in an event. While the agent is characterized by the features [+ volitionality] and [+ instigation], the patient most typically is [+ affected]. Besides these verb-based properties, she further includes the animacy of the object as a factor that increases the distinctness.[4] The point of departure of these accounts is the prototypical transitive scenario. The more a transitive sentence deviates from this prototype by the lack of core properties, such as volitionality for an agent or affectedness for a patient, the less prototypical it is.

Based on the concept of semantic transitivity, typological research has formulated generalizations on verb-based splits in case selection. Tsunoda (1981, 1985) argues in favour of a ranking of Hopper/Thompsons' (1980) transitivity parameters depending on the morphosyntactic phenomenon in question. As for the selection of a transitive case frame, he considers the affectedness of the patient decisive rather than the volitionality or agency of the agent. Based on a sample of ergative and accusative languages, he establishes the following scale of verb classes along which a decrease of unmarked case frames can be attested from left to right:

(3) (1) ACTION ('break', 'hit') > (2) PERCEPTION ('see', 'look at') > (3) PURSUIT ('search') > (4) KNOWLEDGE ('know') > (5) FEELING ('love') > (6) RELATIONSHIP ('possess') > (7) ABILITY ('capable')

Malchukov (2005: 80f.), though, identifies two different kinds of deviation from the transitive prototype in Tsunoda's scale: (i) a deviation from the prototypical affected patient and (ii) a deviation from both the prototypical patient *and* the agentive agent. As a consequence, he decomposes Tsunoda's scale into a two-dimensional hierarchy. Haspelmath (2015: 144) broadly confirms both sub-hierarchies in a quantitative study that ranks 80 verbs in terms of their transitivity prominence in a sample of 36 languages as illustrated in Figure 1.



[4] For a critique on the integration of animacy into Næss' model cf. García García (2014: 124, fn. 13) and Iemmolo (in preparation: 38f.).

Figure 1. Haspelmath's (2015) data in Malchukov's (2005) two-dimensional hierarchy

Proto-role models, which represent the theoretical framework of my account, offer yet another perspective to capture deviations from the prototypical transitive scenario in that they allow us to formulate a *positive* definition of non-prototypical thematic configurations. Note that one advantage of proto-role models, a merit of Dowty's (1991) approach, is that one argument can accumulate both proto-agent and proto-patient properties at the same time and that both co-arguments can bear identical proto-properties. Corollary 1 of his argument selection principle states that when both arguments bear approximately the same number of proto-properties, either argument may be realized as the subject or direct object (Dowty 1991: 576). García García (2007: 71, 2014: 145) further specifies this corollary in order to account for the selection of DOM for inanimate objects in Spanish. He pins down two scenarios in which a lack of thematic distinctness can be detected between both arguments, that is either (i) when both subject and object bear the same number of proto-agent properties or (ii) when the object has more proto-agent properties than the subject. Blume's (2000: 200) licensing condition for marked valencies also positively defines deviations from the prototypical assignment of proto-roles. The present work incorporates these different configurations of deviation, cross-classifying a horizontal dimension, i.e. the decrease of communicated agentivity for the object, and a vertical dimension, i.e. the increase towards a canonical agent-patient asymmetry. Further research must show both to what extent the scale is suitable for capturing systematic cross-linguistic tendencies in linking beyond DOM and in what way adaptions have to be made on a more theoretical level. As for the former point, there is evidence to suggest that the agentivity hierarchy not only allows to capture empirical generalizations on DOM in Romance languages but that it may be applied to non-differential case marking from a typological perspective, too. Blume's (2000: 214-217) cross-linguistic study on marked valencies points out that languages that rank the licensing condition for marked case frames higher than the unmarkedness constraint are prone to overgeneralize it. More specifically, in languages like Finnish or Czech marked case frames are not only selected for verbs that display a non-asymmetric argument relation but also for verbs where an agent-patient asymmetry is present though overlapping with proto-agent properties for the object in one subevent.[5] Further instances are Icelandic and Faroese where dative case marking extends to verb classes with prototypical patient objects (Jónsson 2009; Maling 2001). Interestingly, in Modern Faroese a valency change from dative to accusative can be attested for a good number of verbs (Jónsson 2009: 217-219). The fact that dative-loss is detected first with verbs that solely assign proto-patient but no proto-agent properties to their object, such as *kasta* 'to throw' or *læsa* 'to lock', suggests that the retraction of marked valencies might develop along the agentivity scale.

[5] Cf. for instance Finnish *kiusata* 'to harrass', *suudella* 'to kiss', *hyväillä* 'to spoil', *siunata* 'to bless' and *rangaista* 'to punish', *haukkua* 'to tell off', *loukata* 'to offend' and Czech *zlořečit* 'to punish', *nabít, namlátit*, *naplácat, nasekat* 'to punish by beating, spanking' and *nafackovat* 'to punish with a slap in the face', *nesedět* 'to harrass', *ubližovat* 'to offend' (Blume 2000: 216f., appendix).

In conclusion, this study can be also seen as a contribution to argument realization theory in that it proposes a fine-grained approach to the gradience of communicated agentivity – mainly *sentience* – for the direct object, which is shown to affect its morphosyntactic expression in Romance languages. It thus provides a detailed picture of one more facet of possible deviations from the prototypical agent-patient asymmetry, which may be reflected on the morphosyntactic level. In this regard, it represents an example case for the general claim made by Levin/Rappaport Hovav: "Despite the pervasiveness of the agent–subject and patient–object associations, the bottom line is that there are also many other rich empirical generalizations underlying the association of semantic roles with their morphosyntactic realizations" (2005: 32).

References

Ackerman, Farrell & John Moore. 1999. Syntagmatic and paradigmatic dimensions of causee encodings. *Linguistics and Philosophy* 22(1). 1–44.

Ackerman, Farrell & John Moore. 2001. *Proto-properties and grammatical encoding. A correspondence theory of argument selection*. Stanford: Center for the Study of Language and Information.

Adli, Aria. 2015. What you like is not what you do: Acceptability and frequency in syntactic variation. In Aria Adli, Marco García García & Göz Kaufmann (eds.), *Variation in language: System- and usage-based approaches*, 173–199. Berlin: de Gruyter.

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21(3). 435–483.

Alexiadou, Artemis & Elena Anagnostopoulou. 1998. Parametrizing Agr: word order, verb-movement and EPP-checking. *Natural Language and Linguistic Theory* 16(3). 491–539.

Alfonzetti, Giovanna. 1998. The conversational dimension in code-switching between Italian and dialect in Sicily. In Peter Auer (ed.), *Code-switching in conversation*, 180–214. London: Routledge.

Allen, Cynthia L. 1995. *Case marking and reanalysis: Grammatical relations from Old to Early Modern English*. Oxford: Clarendon Press.

Alsina, Alex. 2016. Catalan. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 363–381. Oxford: Oxford University Press.

Andrews, Avery D. 1985. The major function of the noun phrase. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 1, Clause structure, 62–154. Cambridge: Cambridge University Press.

Antonov, Anton & Alexandru Mardale. 2014. From perlative to differential object marking. The curious case of Romanian PE. Paper presented at the Workshop *The diachronic typology of Differential Argument Marking*. Universität Konstanz, April 5–6 2014.

Bader, Markus & Jana Häussler. 2010. Towards a model of grammaticality judgments. *Journal of Linguistics* 46(2). 273–330.

Barðdal, Johanna. 2008. *Productivity: Evidence From Case and Argument Structure in Icelandic*. Amsterdam: John Benjamins.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1). 1–48.

Beavers, John. 2010. The structure of lexical meaning: Why semantics really matters. *Language* 86(4). 821–864.

Bel, Aurora. 2002. Les funcions sintactiques. In Joan Solà, M. Rosa Lloret, Joan Mascaró & Manuel Pérez Saldanya (eds.), *Gramàtica del català contemporani.* Amb la col·laboració de Gemma Rigau, vol. 2, Sintaxi, 1075–1147. Barcelona: Empúries.

Belletti, Adriana. 2018. On a-Marking of object topics in the Italian left periphery. In Roberto Petrosino, Pietro Cerrone & Harry van der Hulst (eds.), *Beyond the veil of Maya. From sounds to structures*, 445–466. Berlin: de Gruyter.

Benincà, Paola. 1986. Il lato sinistro della frase italiana. *Balkan-Archiv* 11. 213–243.

Bentley, Delia. 2008. The interplay of focus structure and syntax: evidence from two sister languages. In Robert D. Van Valin Jr. (ed.), *Investigations of the syntax-semantics-pragmatics interface*, 263–284. Amsterdam: John Benjamins.

Berretta, Monica. 1989. Sulla presenza dell'accusativo preposizionale in italiano settentrionale: note tipologiche. *Vox Romanica* 48. 13–37.

Berretta, Monica. 1991. Note sulla sintassi dell'accusativo preposizionale in italiano. *Linguistica* 31. 211–232.

Blume, Kerstin. 1998. A Contrastive Analysis of Interaction Verbs with Dative Complements. *Linguistics* 36. 253–280.

Blume, Kerstin. 2000. *Markierte Valenzen im Sprachvergleich: Lizensierungs- und Linkingbedingungen*. Tübingen: Niemeyer.

Bossong, Georg. 1982. Historische Sprachwissenschaft und empirische Universalienforschung. *Romanistisches Jahrbuch* 33. 17–51.

Bossong, Georg. 1985. *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.

Bossong, Georg. 1991. Differential object marking in Romance and beyond. In Dieter Wanner & Douglas A. Kibbee (eds.), *New analyses in Romance linguistics*. Selected papers from the XVIII Linguistic Symposium on Romance Languages, Urbana-Champaign, April 7-9, 1988, 143–170. Amsterdam: John Benjamins.

Bossong, Georg. 1998. Le marquage différentiel de l'objet dans les langues d'Europe. In Jack Feuillet (ed.), *Actance et valence*, 193–258. Berlin: de Gruyter.

Bossong, Georg. 2008. *Die romanischen Sprachen. Eine vergleichende Einführung*. Hamburg: Buske.

Bott, Oliver & Torgrim Solstad. 2014. From verbs to discourse: A novel account of implicit causality. In Barbara Hemforth, Barbara Mertins & Cathrine Fabricius-Hansen

(eds.), *Psycholinguistic approaches to meaning and understanding across languages*, 213–251. New York: Springer.

Carlson, Greg N. 1977. *Reference to kinds in English*. Amherst: University of Massachusetts dissertation.

Caro Reina, Javier. 2014. The grammaticalization of the terms of address *en* and *na* as onymic markers in Catalan. In Friedhelm Debus, Rita Heuser & Damaris Nübling (eds.), *Linguistik der Familiennamen*, 175–204. Hildesheim: Olms.

Caro Reina, Javier. 2019. *Central Catalan and Swabian: A study in the framework of the typology of syllable and word languages*. Berlin: de Gruyter.

Caro Reina, Javier. 2020. Differential Object Marking with proper names in Romance languages. In Luise Kempf, Damaris Nübling & Mirjam Schmuck (eds.), *Linguistik der Eigennamen*, 225–258. Berlin: de Gruyter.

Caro Reina, Javier, Marco García García & Klaus von Heusinger. 2021. Differential Object Marking in Cuban Spanish. In Johannes Kabatek, Philipp Obrist & Albert Wall (eds.), *Differential Object Marking in Romance – The third wave*. Berlin: de Gruyter.

Cassarà, Alessia & Sophie Mürmann. 2021. Role-semantic parameters for DOM in Italian. In Johannes Kabatek, Philipp Obrist & Albert Wall (eds.), *Differential Object Marking in Romance – The third wave*. Berlin: de Gruyter.

Chiriacescu, Sofiana & Klaus von Heusinger. 2010. Discourse prominence and *pe*-marking in Romanian. *International Review of Pragmatics* 2(2). 298–332.

Chiriacescu, Sofiana & Klaus von Heusinger. 2011. Discourse Structuring Potential of Definite Noun Phrases in Natural Discourse. In Iris Hendrickx, Sobha Lalitha Devi, António Branco & Ruslan Mitkov (eds.), *Anaphora and Reference Resolution. Proceedings of the 8th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC 2011)*, 151–162. Lisbon: Edições Colibri.

Comrie, Bernard. 1989. *Language universals and linguistic typology. Syntax and morphology*. 2nd ed. Oxford: Blackwell.

Cortés, Susana, Conxita Lleó & Ariadna Benet. 2009. Gradient merging of vowels in Barcelona Catalan under the influence of Spanish. In Kurt Braunmüller & Juliane House (eds.), *Convergence and divergence in language contact situations*, 185–204. Amsterdam: John Benjamins.

Croft, William. 1986. Surface Subject Choice of Mental Verbs. Paper presented at the *Annual Meeting of the Linguistic Society of America*. New York.

Croft, William. 1993. Case Marking and the Semantics of Mental Verbs. In James Pustejovsky (ed.), *Semantics and the Lexicon*, 55–72. Dordrecht: Kluwer.

Cruschina, Silvio. 2012. *Discourse-Related Features and Functional Projections*. Oxford: Oxford University Press.

Cruse, D. Alan. 1973. Some thoughts on agentivity. *Journal of Linguistics* 9. 11–23.

Dahl, Östen. 1990. Standard Average European as an exotic language. In Johannes Bechert, Giuliano Bernini & Claude Buridant (eds.), *Toward a Typology of European Languages*, 3–8. Berlin: de Gruyter.

Dahl, Östen. 2008. Animacy and egophoricity: Grammar, ontology and phylogeny. *Lingua* 118(2). 141–150.

Dahl, Östen & Kari Fraurud. 1996. Animacy in grammar and discourse. In Thorstein Fretheim & Jeanette K. Gundel (eds.), *Reference and referent accessibility*, 47–64. Amsterdam: John Benjamins.

D'Alessandro, Roberta. 2017. When you have too many features: Auxiliaries, agreement and clitics in Italian varieties. *Glossa* 2(1). 1–36.

Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press.

D'Aprile, Vincenzo. 2005. Analfabetismo: Italia - Censimento 2001: Cinque Grandi Ripartizioni Geografiche e Venti Regioni. *Educazione&Scuola* (http://www.edscuola.it/archivio/statistiche/analfabetismo_01.pdf, last accessed 30 June 2023).

Delbecque, Nicole. 1998. Why Spanish has two transitive construction frames. *Leuvense Bijdragen* 87. 387–415.

Delbecque, Nicole. 2002. A construction grammar approach to transitivity in Spanish. In Kristin Davidse & Béatrice Lamiroy (eds.), *The nominative & accusative and their counterparts*, 81–130. Amsterdam: John Benjamins.

Delille, Karl Heinz. 1970. *Die geschichtliche Entwicklung des präpositionalen Akkusativs im Portugiesischen*. Bonn: Universität Bonn.

Döhla, Hans-Jörg. 2014. Diachronic convergence and divergence in differential object marking between Spanish and Portuguese. In Kurt Braunmüller, Steffen Höder & Karoline Kühl (eds.), *Stability and Divergence in Language Contact. Factors and Mechanisms*, 265–289. Amsterdam: John Benjamins.

Dowty, David. 1991. Thematic Proto-Roles and Argument Selection. *Language* 67. 547–619.

Dubert, Francisco & Charlotte Galves. 2016. Galician and Portuguese. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 411–446. Oxford: Oxford University Press.

Dufter, Andreas & Elisabeth Stark. 2008. Double indirect object marking in Spanish and Italian. In Elena Seoane & María José López-Couso (eds.), *Theoretical and Empirical Issues in Grammaticalization*, 111–129. Amsterdam: John Benjamins.

Egetenmeyer, Jakob. 2019. *Der Verbalanschluss im Spanischen: Kognitiv-syntaktische Analyse nominaler und satzwertiger Akkusativobjekte*. Berlin: de Gruyter.

Engelberg, Stefan. 1995. Event Structure and the Meaning of Verbs. In Per Bærentzen (ed.), *Aspekte der Sprachbeschreibung. Akten des 29. Linguistischen Kolloquiums, Aarhus 1994*, 37–41. Tübingen: Niemeyer.

Engelberg, Stefan. 2000a. *Verben, Ereignisse und das Lexikon*. Tübingen: Niemeyer.

Engelberg, Stefan. 2000b. Verb Meaning as Event Structure. In Alan K. Melby & Arle R. Lommel (eds.), *LACUS Forum XXVI. The Lexicon*, 257–268. Fullerton, CA: LACUS.

Engelberg, Stefan. 2006. A Theory of Lexical Event Structures and its Cognitive Motivation. In Dieter Wunderlich (ed.), *Advances in the Theory of the Lexicon*, 235–285. Berlin: de Gruyter.

Enghels, Renata. 2007. *Les modalités de perception visuelle et auditive. Différences conceptuelles et répercussions sémantico-syntaxiques en espagnol et en français*. Tübingen: Niemeyer.

Enghels, Renata. 2013. Transitivity of Spanish Perception Verbs: a Gradual Category? *Borealis* 2(1). 35–56.

Enghels, Renata & Eugeen Roegiest. 2012a. L'expression syntaxique de la causation négative avec *laisser*, *dejar*, *deixar*: trois systèmes romans différents. In Marleen Van Peteghem, Peter Lauwers, Els Tobback, Annemie Demol, & Laurence De Wilde (eds.), *Le verbe en verve: réflexions sur la syntaxe et la sémantique verbales*, 167–187. Ghent: Academia Press.

Enghels, Renata & Eugeen Roegiest. 2012b. Los Verbos De Causación Negativa *Dejar* y *Laisser*: Sintaxis y Polisemia. In Valeriano Bellosta von Colbe & Marco García García (eds.), *Aspectualidad, Transitividad y Referencialidad: Las Lenguas Románicas En Contraste*, 87–119. Frankfurt am Main: Peter Lang.

Enghels, Renata & Eugeen Roegiest. 2013. *Dejar*: entre verbo causativo y verbo de control. *Bulletin of Hispanic Studies* 90(5). 505-522.

Escandell-Vidal, María Victoria. 2009. Differential object marking and topicality. The case of Balearic Catalan. *Studies in Language* 33(4). 832–885.

Fábregas, Antonio. 2013. Differential object marking in Spanish: State of the art. *Borealis* 2. 1–80.

Fagard, Benjamin & Alexandru Mardale. Non, mais tu l'as vu à lui? Analyse(s) du marquage différentiel de l'objet en français. *Verbum* 36(1). 143–168.

Fernández Ramírez, Salvador. 1986. *Gramática española*, vol. 4, El verbo y la oración. Madrid: Arco Libros.

Fillmore, Charles J. 1968. The case for case. In Emmon Bach & Robert T. Harms (eds.), *Universals in linguistic theory*, 1–88. New York: Holt, Rinehart & Winston.

Fleischhauer, Jens. 2018. Animacy and Affectedness in Germanic Languages. *Open Linguistics* 4. 566–588.

Foley, William A. & Robert D. Van Valin Jr. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

Fraurud, Kari. 1996. Cognitive Ontology and NP Form. In Thorstein Fretheim & Jeanette K. Gundel, *Reference and Referent Accessibility*, 65–87. Amsterdam: John Benjamins.

Frege, Gottlob. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100. 25–50. Reprinted in Gottlob Frege, 2008, *Funktion, Begriff, Bedeutung. Fünf logisch Studien*, 23–46. Göttingen: Vandenhoeck & Ruprecht.

García García, Marco. 2005. Differential Object Marking and informativeness. In Klaus von Heusinger, Georg A. Kaiser & Elisabeth Stark (eds.), *Proceedings of the workshop "Specificity and the evolution/emergence of nominal determination systems in Romance"*, 17–31, Arbeitspapier 119, Fachbereich Sprachwissenschaft. Konstanz: Universität Konstanz.

García García, Marco. 2007. Differential object marking with inanimate objects. In Georg A. Kaiser & Manuel Leonetti (eds.), *Proceedings of the workshop "Definiteness, specificity and animacy in Ibero-Romance languages"*, 63–84, Arbeitspapier 122, Fachbereich Sprachwissenschaft. Konstanz: Universität Konstanz.

García García, Marco. 2010. *Differentielle Objektmarkierung bei unbelebten Objekten im Spanischen*. Freiburg: Albert-Ludwigs-Universität Freiburg dissertation.

García García, Marco. 2014. *Differentielle Objektmarkierung bei unbelebten Objekten im Spanischen*. Berlin: De Gruyter.

García García, Marco. 2018a. Nominal and verbal parameters in the diachrony of DOM in Spanish. In Ilja A. Seržant, Alena Witzlack-Makarevich & Kelsey Mann (eds.), *The diachronic typology of differential argument marking*, 209–242. Berlin: Language Science Press.

García García, Marco. 2018b. Competing motivations for DOM in Spanish? Paper presented at the Workshop *DOM in Spanish - diachronic change and synchronic variation*. Universität Zürich, June 4–5 2018.

García García, Marco, Beatrice Primus & Nikolaus P. Himmelmann. 2018. Shifting from animacy to agentivity. *Theoretical Linguistics* 44(1–2). 25–39.

GIEC (= Institut d'Estudis Catalans). 2016. *Gramàtica de la llengua catalana*. Barcelona: Institut d'Estudis Catalans.

Goikoetxea, Edurne, Gema Pascual & Joana Acha. 2008. Normative study of the implicit causality of 100 interpersonal verbs in Spanish. *Behavior Research Methods* 40(3). 760–772.

Grice, Paul. 1975. Logic and conversation. In Peter Cole & James L. Morgan (eds.), *Syntax and semantics: Speech acts*, 41–58. New York: Academic.

Grice, Paul. 1989. *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Gries, Stefan Th. 2015. The most under-used statistical method in corpus linguistics: multi-level (and mixed-effects) models. *Corpora* 10(1). 95–125.

Guardiano, Cristina. 2000. Note sull'oggetto diretto preposizionale nel siciliano. *L'Italia Dialettale* 51. 1–35.

Haspelmath, Martin. 1998. How young is Standard Average European? *Language Sciences* 20. 271–287.

Haspelmath, Martin. 2015. Transitivity prominence. In Andrej L. Malchukov & Bernard Comrie (eds.), *Valency classes in the world's languages: A comparative handbook*, vol. 1, 131–147. Berlin: de Gruyter.

Heusinger, Klaus von. 2008. Verbal semantics and the diachronic development of DOM in Spanish. *Probus* 20(1). 1–31.

Heusinger, Klaus von. 2011. Specificity. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics. An International Handbook of Natural Language Meaning*, vol. 2, 1025–1058. Berlin: de Gruyter.

Heusinger, Klaus von & Georg A. Kaiser. 2005. The evolution of differential object marking in Spanish. In Klaus von Heusinger, Georg A. Kaiser & Elisabeth Stark (eds.), *Proceedings of the workshop "Specificity and the evolution/emergence of nominal determination systems in Romance"*, 33–69, Arbeitspapier 119, Fachbereich Sprachwissenschaft. Konstanz: Universität Konstanz.

Heusinger, Klaus von & Georg A. Kaiser. 2011. Affectedness and differential object marking in Spanish. *Morphology* 21. 593–617.

Himmelmann, Nikolaus P. 1997. *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Tübingen: Niemeyer.

Himmelmann, Nikolaus P. 2004. Lexicalization and grammaticization: opposite or orthogonal? In Walter Bisang, Nikolaus P. Himmelmann & Björn Wiemer (eds.), *What makes grammaticalization - a look from its components and its fringes*, 21–42. Berlin: de Gruyter.

Himmelmann, Nikolaus P. & Beatrice Primus. 2015. Prominence beyond prosody – a first approximation. In Amedeo De Dominicis (ed.), *Prominences in Linguistics. Proceedings of the pS-prominenceS International Conference*, 38–58. Viterbo: DISUCOM Press.

Holisky, Dee Ann. 1987. The case of the intransitive subject in Tsova-Tush (Batsbi). *Lingua* 71. 103–132.

Hoop, Helen de. 2015. *Against affectedness*. In Sander Lestrade, Peter de Swart & Lotte Hogeweg (eds.), *Addenda. Artikelen voor Ad Foolen*, 169–176. Nijmegen: Radboud University.

Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2). 251–299.

Hualde, Jose Ignacio. 1992. *Catalan*. London: Routledge.

Iemmolo, Giorgio. 2009. La marcatura differenziale dell'oggetto in siciliano antico. *Archivio Glottologico Italiano* 94(2). 185–225.

Iemmolo, Giorgio. 2010a. Topicality and differential object marking. Evidence from Romance and beyond. *Studies in Language* 34(2). 239–272.

Iemmolo, Giorgio. 2010b. La marcatura differenziale dell'oggetto in siciliano: un'analisi contrastiva. In Maria Iliescu, Paul Danler & Heidi Siller (eds.), *Actes du XXV Congrès International de Linguistique et de Philologie Romanes*, vol. 2, 341–350. Berlin: de Gruyter.

Iemmolo, Giorgio. In preparation. *Differential object marking*. Oxford: Oxford University Press.

Institut d'Estadística de Catalunya (IDESCAT). 2018. *Survey on Language Uses of the Population* (https://www.idescat.cat/pub/?id=eulp&lang=en, last accessed 30 June 2023).

Jacob, Daniel. 2004. Generalisierte semantische Rollen und freie Linkingprinzipien. In Rolf Kailuweit & Martin Hummel (eds.), *Semantische Rollen*, 104–117. Tübingen: Narr.

Jones, Michael Allan. 1995. The prepositional accusative in Sardinian: its distribution and syntactic repercussions. In John Charles Smith & Martin Maiden (eds.), *Linguistic theory and the Romance languages*, 37–75. Amsterdam: Benjamins.

Jones, Michael Allan. 2003. *Sintassi della lingua sarda*. Cagliari: Condaghes.

Jónsson, Jóhannes Gísli. 2009. Verb classes and dative objects in Insular Scandinavian. In Jóhanna Barðdal & Shobhana Chelliah (eds.), *The Role of Semantic, Pragmatic and Discourse Factors in the Development of Case*, 203–224. Amsterdam: John Benjamins.

Kabatek, Johannes. 2016. Wohin strebt die differentielle Objektmarkierung im Spanischen? *Romanistisches Jahrbuch* 67. 211–239.

Kabatek, Johannes, Philipp Obrist & Albert Wall (eds.). To appear. *Differential Object Marking in Romance – The third wave*. Berlin: de Gruyter.

Kailuweit, Rolf. 2004. Protorollen und Makrorollen. In Rolf Kailuweit & Martin Hummel (eds.), *Semantische Rollen*, 84–104. Tübingen: Narr.

Kailuweit, Rolf. 2005. *Linking: Syntax und Semantik französischer und italienischer Gefühlsverben*. Tübingen: Niemeyer.

Kailuweit, Rolf & Martin Hummel (eds.). 2004. *Semantische Rollen*. Tübingen: Narr.

Kaufmann, Ingrid. 1995. What is an (im-)possible verb? Restrictions on Semantic Form and their consequences for argument structure. *Folia Linguistica* 29. 67–103.

Kehler, Andrew, Laura Kertz, Hannah Rohde & Jeffrey L. Elman. 2008. Coherence and coreference revisited. *Journal of Semantics 25*(1). 1–44.

Klein, Nina & Silvia Kutscher. 2002. Psych Verbs in German and the Principle of Lexical Economy. In *Arbeitspapiere des SFB 282 Theorie des Lexikons* 122. Düsseldorf: Universität.

Kloss, Heinz. 1967. 'Abstand languages' and 'Ausbau languages'. *Anthropological Linguistics* 9(7). 29–41.

Kloss, Heinz. 1987. Abstandsprache und Ausbausprache. In Ulrich Ammon, Norbert Dittmar & Klaus J. Mattheier (eds), *Sociolinguistics: An International Handbook of the Science of Language and Society / Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft*, vol. 1, 302–308. Berlin: de Gruyter.

König, Ekkehard & Shigeshiro Kokutani. 2006. Towards a typology of reciprocal constructions: Focus on German and Japanese. *Linguistics* 44(2). 271–302.

Kotthoff, Helga & Damaris Nübling. 2018. *Genderlinguistik. Eine Einführung in Sprache, Gespräch und Geschlecht*. Tübingen: Narr.

Kretzschmar, Franziska, Tim Graf, Markus Philipp & Beatrice Primus. 2018a. What is a sentient agent? In *Pre-Proceedings of the International Conference of Linguistic Evidence 2018*, 47–49. Tübingen: Universität.

Kretzschmar, Franziska, Tim Graf, Markus Philipp & Beatrice Primus. 2018b. *Manual of agentivity features and verb classification. Sentience verbs in German*. SFB 1252 Prominence in Language [Project B07]. Universität zu Köln. Unpublished manual.

Kretzschmar, Franziska, Tim Graf, Markus Philipp & Beatrice Primus. 2019. An experimental investigation of agent prototypicality and agent prominence in German. In Anja Gattnar, Robin Hörnig, Melanie Störzer & Sam Featherston (eds.), *Proceedings of Linguistic Evidence 2018: Experimental Data Drives Linguistic Theory*, 101–123. Tübingen: Universität.

Krifka, Manfred. 1989. *Nominalreferenz und Zeitkonstitution. Zur Semantik von Massentermen, Pluraltermen und Aspektklassen*. München: Fink.

Kutas, Marta & Steven A. Hillyard. 1980. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science* 207(4427). 203–205.

Kutscher, Silvia. 2009. *Kausalität und Argumentrealisierung. Zur Konstruktionsvarianz bei Psychverben am Beispiel europäischer Sprachen*. Tübingen: Niemeyer.

Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82(13). 1–26.

La Fauci, Nunzio. 1991. L'oggetto con preposizione nei Confessionali siciliani antichi. Risultati di uno spoglio sistematico. In Luciano Giannelli, Nicoletta Maraschio, Teresa Poggi Salani & Massimo Vedovelli (eds.), *Tra Rinascimento e strutture attuali, Atti del Primo Convegno SILFI*, 387–398. Torino: Rosenberg & Sellier.

Labov, William. 1966. *The Social Stratification of English in New York City*. Washington D. C.: Center for Applied Linguistics.

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Laca, Brenda. 1995. Sobre el uso del acusativo preposicional en español. In Carmen Pensado (ed.), *El complemento directo preposicional*, 61–91. Madrid: Visor.

Laca, Brenda. 2006. El objeto directo. La marcación preposicional. In Concepción Company Company (ed.), *Sintaxis histórica del español. Primera parte: La frase verbal*, vol. 1, 423–475. México: Universidad Nacional Autónoma de México.

Lamiroy, Béatrice & Michel Charolles. 2011. Des formes aux sens: les clitiques dans la construction causative en *faire*. In Gilles Corminboeuf & Marie-José Béguelin (eds.), *Du système linguistique aux actions langagières*, 41–59. Louvain la Neuve: De Boeck-Duculot.

Lazard, Gilbert. 1990. Caractéristiques actancielles de l'"européen moyen type". In Johannes Bechert, Giuliano Bernini & Claude Buridant (eds.), *Toward a typology of European languages*, 241–253. Berlin: de Gruyter.

Ledgeway, Adam. 2009. *Grammatica diacronica del napoletano*. Tübingen: Niemeyer.

Ledgeway, Adam. 2016a. Italian, Tuscan, and Corsican. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 206–227. Oxford: Oxford University Press.

Ledgeway, Adam. 2016b. The dialects of southern Italy. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 246–269. Oxford: Oxford University Press.

Ledgeway, Adam & Martin Maiden (eds.). 2016. *The Oxford Guide to the Romance Languages*. Oxford: Oxford University Press.

Ledgeway, Adam, Norma Schifano & Giuseppina Silvestri. 2019. Differential Object Marking and the properties of D in the dialects of the extreme south of Italy. *Glossa* 4(1). 1–25.

Leonetti, Manuel. 2004. Specificity and differential object marking in Spanish. *Catalan Journal of Linguistics* 3. 75–114.

Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago, IL: University of Chicago Press.

Levin, Beth & Malka Rappaport Hovav. 2005. *Argument realization*. Cambridge: Cambridge University Press.

Levinson, Steven C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Levinson, Steven C. 1994. *Pragmatik*. Ins Deutsche übersetzt von Ursula Fries. 2nd ed. Tübingen: Niemeyer.

Levinson, Steven C. 2000. *Presumptive meanings. The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.

Löbner, Sebastian. 1985. Definites. *Journal of Semantics* 4(4). 279–326.

Macaulay, Ron K. S. 1978. Variation and Consistency in Glaswegian English. In Peter Trudgill (ed.), *Sociolinguistic Patterns in British English*, 132–143. London: Arnold.

Maiden, Martin. 2016. Romanian, Istro-Romanian, Megleno-Romanian, and Aromanian. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 91–125. Oxford: Oxford University Press.

Malchukov, Andrej. 2005. Case pattern splits, verb types, and construction competition. In Mengistu Amberber & Helen de Hoop (eds.), *Competition and Variation in Natural Languages*, 73–117. London: Elsevier.

Malchukov, Andrej. 2008. Animacy and Asymmetries in Differential Case Marking. *Lingua* 118(2). 203–221.

Maling, Joan. 2001. Dative: The heterogeneity of the mapping among morphological case, grammatical functions, and thematic roles. *Lingua* 111. 419–464.

Marcellesi, Jean-Baptiste. 1986. Le 'complement d'objet direct' en Corse: à + SN de GV, Ø + SN de GV. In *Actes du XVIIe Congrès International de Linguistique et Philologie Romanes* 4, 127–138. Aix-en-Provence: Université de Provence.

Mardale, Alexandru. 2015. Differential Object Marking in the First Original Romanian Texts. In Virginia Hill (ed.), *Formal Approaches to DPs in Old Romanian*, 200–245. Leiden: Brill.

Melis, Chantal. 1995. El objeto directo personal en el 'Cantar de Mio Cid'. In Carmen Pensado (ed.), *El complemento directo preposicional*, 133–161. Madrid: Visor.

Mensching Guido & Eva-Maria Remberger. 2016. Sardinian. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 270–291. Oxford: Oxford University Press.

Montrul, Silvina. 1996. Clitic-doubled dative 'subjects' in Spanish. In Karen Zagona (ed.), *Grammatical theory and Romance languages. Selected papers from the 25th Linguistic Symposium on Romance Languages*, 183–195. Amsterdam: John Benjamins.

Montrul, Silvina. 2013. Differential object marking in Argentine Spanish: An experimental study. In Laura Colantoni & Celeste Rodríguez Louro (eds.), *The handbook of Argentine Spanish*, 207–228. Frankfurt am Main: Vervuert Iberoamericana.

Montrul, Silvina. 2014. Structural changes in Spanish in the United States: Differential object marking in Spanish heritage speakers across generations. *Lingua* 151(Part B). 177–196.

Montrul, Silvina, Rakesh Bhatt & Roxana Girju. 2015. Differential object marking in Spanish, Hindi, and Romanian as heritage languages. *Language* 91(3), 564–610.

Müller, Bodo. 1971. Das morphemmarkierte Satzobjekt der romanischen Sprachen (Der sogenannte präpositionale Akkusativ). *Zeitschrift für romanische Philologie* 87(5–6). 477–519.

Myers-Scotton, Carol. 1993. *Social Motivations for Codeswitching: Evidence from Africa*. Oxford: Claredon.

Næss, Åshild. 2004. What markedness marks: The markedness problem with direct objects. *Lingua* 114. 1186–1212.

Næss, Åshild. 2007. *Prototypical transitivity*. Amsterdam: John Benjamins.

Neuburger, Kathrin Anne & Elisabeth Stark. 2014. Differential Object Marking in Corsican: Regularities and triggering factors. *Linguistics* 52(2). 365–389.

Nieuwland, Mante S. & Jos J.A. Van Berkum. 2006. When Peanuts Fall in Love: N400 Evidence for the Power of Discourse. *Journal of Cognitive Neuroscience* 18(7). 1098–1111.

Nunberg, Geoffrey. 1979. The non-uniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy* 3. 143–184.

Pană Dindelegan, Gabriela (ed.). 2013. *The Grammar of Romanian*. Oxford: Oxford University Press.

Parry, Mair. 2003. L'oggetto preposizionale nel Ligure medievale. *Verbum* V(1). 113–126.

Pensado, Carmen (ed.). 1995a. *El complemento directo preposicional*. Madrid: Visor.

Pensado, Carmen. 1995b. El complemento directo preposicional. Estado de la cuestión y bibliografía comentada. In Carmen Pensado (ed.), *El complemento directo preposicional*, 11–59. Madrid: Visor.

Pensado, Carmen. 1995c. La creación del complemento directo preposicional y la flexión de los pronombres personales en las lenguas románicas. In Carmen Pensado (ed.), *El complemento directo preposicional*, 179–233. Madrid: Visor.

Pesetsky, David. 1995. *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.

Pineda, Anna. 2013. Double Object Constructions and dative/accusative alternations in Spanish and Catalan: a unified account. *Borealis* 2(1). 57–115.

Pineda, Anna. 2015. Del datiu a l'acusatiu. Un canvi sintàctic en procés en llengües romàniques i basc. *Llengua & Literatura* 25. 73–97.

Pineda, Anna. 2021. The development of DOM in the diachrony of Catalan: (dis)similarities with respect to Spanish. In Johannes Kabatek, Philipp Obrist & Albert Wall (eds.), *Differential Object Marking in Romance – The third wave*. Berlin: de Gruyter.

Pineda, Anna. In press. El complement directe. In Josep Martines & Manuel Pérez-Saldanya (eds.), *Gramàtica del català antic*. Amsterdam: Benjamins.

Pineda, Anna & Carles Royo. 2017. Differential indirect object marking in Romance (and how to get rid of it). *Revue Roumaine de Linguistique* 62(4). 445–462.

Pottier, Bernard. 1968. L'emploi de la préposition a devant l'objet en Espagnol. *Bulletin de la Société de Linguistique* 1. 83–95.

Pountain, Christopher J. 2016. Standardization. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 634–643. Oxford: Oxford University Press.

Prieto Entrialgo, Clara Elena. 2010. Reflexones sobre la marcación preposicional del OD n'asturianu. Dellos datos medievales. *Lletres Asturianes* 103. 17–34.

Primus, Beatrice. 1999a. *Cases and thematic roles. Ergative, accusative and active*. Tübingen: Niemeyer.

Primus, Beatrice. 1999b. Rektionsprinzipien. In Heide Wegener (ed.), *Deutsch – kontrastiv. Typologisch vergleichende Untersuchungen zur deutschen Grammatik*, 135–170. Tübingen: Stauffenburg.

Primus, Beatrice. 2002. Proto-roles and case selection in Optimality Theory. *Arbeiten des SFB 282 "Theorie des Lexikons"* 122. 1–39.

Primus, Beatrice. 2004. Protorollen und Verbtyp: Kasusvariation bei psychischen Verben. In Martin Hummel & Rolf Kailuweit (eds.), *Semantische Rollen*, 377–401. Tübingen: Narr.

Primus, Beatrice. 2006. Hierarchy mismatches and the dimensions of role semantics. In Ina Bornkessel, Matthias Schlesewsky, Bernard Comrie & Angela D. Friederici (eds.), *Semantic role universals and argument linking. Theoretical, typological and psycholinguistic perspectives*, 53–88. Berlin: de Gruyter.

Primus, Beatrice. 2012a. Animacy, generalized semantic roles and differential object marking. In Monique Lamers & Peter de Swart (eds.), *Case, word order, and prominence: Interacting cues in language production and comprehension*, 65–90. Dordrecht: Springer.

Primus, Beatrice. 2012b. *Semantische Rollen*. Heidelberg: Carl Winter.

Putzu, Ignazio. 2008. Per uno studio dell'accusativo preposizionale in sardo antico: emergenze dallo spoglio del Condaghe di San Pietro di Silki. In Romano Lazzeroni, Emanuele Banfi, Giuliano Bernini, Marina Chini & Giovanna Marotta (eds.), *Diachronica et Synchronica. Studi in onore di Anna Giacalone Ramat*, 397–428. Pisa: ETS.

Rispoli Mattana, Anna. 1996. *Alcamo Canta*. Alcamo: Campo.

Roegiest, Eugeen. 1979. A propos de l'accusatif prépositionnel dans quelques langues romanes. *Vox Romanica* 38. 37–54.

Roegiest, Eugeen. 2003. Argument structure of perception verbs and actance variation of the Spanish direct object. In Giuliana Fiorentino (ed.), *Romance objects: Transitivity in Romance languages*, 299–322. Berlin: de Gruyter.

Rohlfs, Gerhard. 1971. Autour de l'accusatif prépositionnel dans les langues romanes: con-cordances et discordances. *Revue de linguistique romane* 35. 312–334.

Rohlfs, Gerhard. 1973. Panorama de l'accusatif prépositionnel en Italie. *Studii şi cercetări lingvistice* 24. 617–621.

Romagno, Domenica. 2007. Canonical and non-canonical marking of core arguments in European languages. A typological approach. In Paolo Ramat & Elisa Roma (eds.), *Europe and the Mediterranean as Linguistic Areas: Convergences from a Historical and Typological Perspective*, 289–315. Amsterdam: John Benjamins.

Romero Heredero, Diego. 2021. Telicity and Differential Object Marking in the history of Spanish. In Johannes Kabatek, Philipp Obrist & Albert Wall (eds.), *Differential Object Marking in Romance – The third wave*. Berlin: de Gruyter.

Romero Heredero, Diego. 2022. *Marcado diferencial de objeto y semántica verbal en español*. Berlin: de Gruyter.

Romero Heredero, Diego & Marco García García. 2023. Differential object marking in Spanish: The effect of affectedness. *Caplletra* 74(1). 259–285.

Rosch, Eleanor H. 1973. On the Internal Structure of Perceptual and Semantic Categories. In Timothy E. Moore (ed.), *Cognitive Development and the Acquisition of Language*, 111–144. New York: Academic Press.

Rosenbach, Anette. 2008. Animacy and grammatical variation – Findings from English genitive variation. *Lingua* 118. 151–171.

Royo, Carles. 2017. Verbs psicològics catalans: el cas de *agradar* i *interessar*. *Caplletra* 62. 65–88.

Ruaix i Vinyet, Josep. 1994. *Observacions crítiques i pràctiques sobre el català d'avui/1*. Moià: Ruaix.

Sancho Cremades, Pelegrí. 1995. *La categoria preposicional*. Valencia: Universitat de Valencia.

Sancho Cremades, Pelegrí. 2002. La preposició i el sintagma preposicional. In Joan Solà, M. Rosa Lloret, Joan Mascaró & Manuel Pérez Saldanya (eds.), *Gramàtica del català contemporani*. Amb la col·laboració de Gemma Rigau, vol. 2, Sintaxi, 1689–1796. Barcelona: Empúries.

Schumacher, Petra B. 2011. The hepatitis called …: Electrophysiological Evidence for Enriched Composition. In Jörg Meibauer & Markus Steinbach (eds.), *Experimental Pragmatics/Semantics*, 199–219. Amsterdam: John Benjamins.

Schumacher, Petra B. 2018. On Type Composition and Agentivity. *Theoretical Linguistics* 44(1-2). 81–91.

Seefranz-Montag, Ariane. 1983. *Syntaktische Funktionen und Wortstellung*. München: Fink.

Seres, Daria & M. Teresa Espinal. 2018. Psychological verbs and their arguments. *Borealis* 7(1). 27–44.

SIL International. *Ethnologue: Languages of the World* (https://www.ethnologue.com/, last accessed 30 June 2023).

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Robert M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aboriginal Studies.

Solà, Joan. 1994. *Sintaxi normativa: estat de la questio,* 2nd ed. Barcelona: Empuries.

Solstad, Torgrim & Oliver Bott. 2017. Causality and Causal Reasoning in Natural Language. In Michael R. Waldmann (ed.), *The Oxford Handbook of Causal Reasoning*, 619–644. Oxford: Oxford University Press.

Sornicola, Rosanna. 1997. L'oggetto preposizionale in siciliano antico e in napoletano antico. Considerazioni su un problema di tipologia diacronica. *Italienische Studien* 18. 66–80.

Sornicola, Rosanna. 2000. Processi di convergenza nella formazione di un tipo sintattico: la genesi ibrida dell'oggetto preposizionale. In Annick Englebert (ed.), *Les nouvelles ambitions de la linguistique diachronique*, 419–428. Berlin: de Gruyter.

Sornicola, Rosanna. 2011. Romance Linguistics and Historical Linguistics: Reflections on Synchrony and Diachrony. In Martin Maiden, John Charles Smith & Adam Ledgeway, *The Cambridge History of the Romance Languages*, 1–49. Cambridge: Cambridge University Press.

Strawson, Peter Frederick. 1950. On referring. *Mind* 59. 320–344.

Swart, Peter de. 2007. *Cross-linguistic variation in object marking*. Utrecht: LOT.

Swart, Peter de. 2014. Prepositional inanimates in Dutch: A paradigmatic case of differential object marking. *Linguistics* 52(2). 445–468.

Swart, Peter de & Helen de Hoop. 2018. Shifting animacy. *Theoretical Linguistics* 44(1-2). 1–23.

Torrego Salcedo, Esther. 1998. *The dependencies of objects*. Cambridge, MA: MIT Press.

Torrego Salcedo, Esther. 1999. El complemento directo preposicional. In Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, 1780–1805. Madrid: Espasa-Calpe.

Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Tsunoda, Tasaku. 1981. Split case-marking patterns in verb-types and tense/aspect/mood. *Linguistics* 19. 389–438.

Tsunoda, Tasaku. 1985. Remarks on transitivity. *Journal of Linguistics* 21(2). 385–396.

Vallverdú Albornà, Teresa. 2002. Fenòmens en grups vocàlics. In Joan Solà, M. Rosa Lloret, Joan Mascaró & Manuel Pérez Saldanya (eds.), *Gramàtica del català contemporani*, vol. 1, 125–167. Barcelona: Empúries.

Van Valin, Robert D. Jr. 1999. Generalized Semantic Roles and the Syntax-Semantics Interface. In Francis Corblin, Carmen Dobrovie-Sorin & Jean-Marie Marandin (eds.), *Empirical Issues in Formal Syntax and Semantics* 2, 373–389. The Hague: Thesus.

Vendler, Zeno. 1957. Verbs and Times. *Philosophical Review* 56. 143–60.

Verhoeven, Elisabeth. 2014. Thematic prominence and animacy asymmetries. Evidence from a cross-linguistic production study. *Lingua* 143. 129–161.

Vogels, Jorrig, Emiel Krahmer & Alfons Maes. 2013. When a stone tries to climb up a slope: The interplay between lexical and perceptual animacy in referential choices. *Frontiers in Psychology* 4(154). 1–15.

Volo, Valentina. 2010. L'oggetto preposizionale nel siciliano del trecento. *Bollettino di Palermo* 22. 171–202.

Weissenrieder, Maureen. 1985. Exceptional uses of the accusative *a*. *Hispania* 68(2). 393–398.

Weissenrieder, Maureen. 1991. A functional approach to the accusative *a*. *Hispania* 74(1). 146–156.

Weissenrieder, Maureen. 1990. Variable uses of the direct-object marker *a*. *Hispania* 73(1). 223–231.

Wheeler, Max W., Alan Yates & Nicolau Dols. 1999. *Catalan. A comprehensive grammar*. London: Routledge.

Witzlack-Makarevich, Alena & Ilja A. Seržant. 2018. Differential argument marking. Patterns of variation. In Ilja A. Seržant & Alena Witzlack-Makarevich (eds.), *The diachronic typology of differential argument marking*, 1–40. Berlin: Language Science Press.

Yamamoto, Mutsumi. 1999. *Animacy and Reference. A Cognitive Approach to Corpus Linguistics*. Amsterdam: John Benjamins.

Appendix

List of appendices

Appendix 1. Acceptability judgement study on Sicilian 261
1.1 Test items 261
1.2 Version A and B of the questionnaire 264
1.3 Form of sociodemographic data 264
1.4 Results of the individual items (ordered by verbs) 266
Appendix 2. Acceptability judgement study on Catalan 268
2.1 Test items (Barcelona) 268
2.2 Test items (Girona) 270
2.3 Version A and B of the questionnaire (Barcelona and Girona) 271
2.4 Form of sociodemographic data 271
2.5 Results of the individual items (ordered by verbs) (Barcelona and Girona) 273

Appendix 1. Acceptability judgement study on Sicilian

1.1 Test items

Test item (with DOM)	**Verb [communicated agentivity for the object]**	**Direct object NP**
1. *Totò cu lu so' malu lamentu avìa affinnutu a lu so' capu.* 'Totò offended his boss with his evil accusations.'	*affènniri* 'to offend' [entailed agentivity]	definite human NP
2. *Nuccia cu paroli tinti avìa affinnutu a 'na cumpagna.* 'Nuccia offended a friend with bad words.'	*affènniri* 'to offend' [entailed agentivity]	indefinite human NP
3. *Mattè cu li so' vuci avìa nuciutu a lu prufissuri.* 'Mattè disturbed the teacher with his screaming.'	*nuciri* 'to bother' [entailed agentivity]	definite human NP
4. *Aspanu arre' nni la bibliuteca avìa nuciutu a 'na picciuttedda.* 'Aspanu again disturbed a girl in the library.'	*nuciri* 'to bother' [entailed agentivity]	indefinite human NP
5. *Dina cu paroli duci avìa cunsulatu a l'omu.* 'Dina comforted the man with sweet words.'	*cunsulari* 'to console' [entailed agentivity]	definite human NP
6. *Adele cu amuri avìa cunsulatu a un picciriddu.*	*cunsulari* 'to console' [entailed agentivity]	indefinite human NP

'Adele comforted a boy with love.'		
7. *Vicenzu cu li so' 'nciurii avìa fattu siddiari a lu sbirru.* 'Vicenzo provoked the police-man with his insults.'	*fari siddiari* 'to provoke' [entailed agentivity]	definite human NP
8. *Sarvaturi cu lu so' tintu fari avìa fattu siddiari a 'na fimmina.* 'Sarvaturi provoked a woman with his bad behaviour.'	*fari siddiari* 'to provoke' [entailed agentivity]	indefinite human NP
9. *Peppe p'addivèrtisi avìa fattu scantari a la vecchia.* 'Peppe frightened the old woman for fun.'	*fari scantari* 'to frighten' [entailed agentivity]	definite human NP
10. *Saridda pi ririri avìa fattu scantari a n'amicu.* 'Saridda frightened a friend to laugh at him.'	*fari scantari* 'to frighten' [entailed agentivity]	indefinite human NP
11. *Dinu 'nmenzu la strata avìa arristatu a lu latru.* 'Dino stopped the thief in the middle of the street.'	*arristari* 'to stop' [presupposed agentivity]	definite human NP
12. *Cicca fora la cantina avìa arristatu a un 'mmriacuni.* 'Cicca stopped a drunk outside the wine cellar. '	*arristari* 'to stop' [presupposed agentivity]	indefinite human NP
13. *Giacuminu di 'nmenzu lu focu avìa sarvatu a lu so' vicinu.* 'Giacuminu saved his neighbour out of the fire.'	*sarvari* 'to save' [presupposed agentivity]	definite human NP
14. *Luca di 'nmenzu lu sciumi avìa sarvatu a un picciriddu.* 'Luca saved a child out of the river.'	*sarvari* 'to save' [presupposed agentivity]	indefinite human NP
15. *Marcu fora la putia avìa acchiappatu a lu latru.* 'Marco catched the thief outside the store.'	*acchiappari* 'to catch' [presupposed agentivity]	definite human NP
16. *Pinu a lu mercatu avìa acchiappatu a un maliutu.* 'Pino catched the criminal on the market.'	*acchiappari* 'to catch' [presupposed agentivity]	indefinite human NP
17. *Sara 'na lu iardinu avìa sintutu a la so' vicina.* 'Sara heard her neighbour in the garden.'	*sentiri* 'to hear' [presupposed agentivity]	definite human NP
18. *Agata 'nmenzu la strata avìa sintutu a 'na picciridda.* 'Agata heard a little girl in the street.'	*sentiri* 'to hear' [presupposed agentivity]	indefinite human NP
19. *Antrea cu la spisa avìa aiutatu a lu vecchiu.*	*aiutari* 'to help' [presupposed agentivity]	definite human NP

'Antrea helped the old man with the shopping.'		
20. *Nanà cu li compiti avìa aiutatu a n'amicu.* 'Nanà helped a friend with his homework.'	*aiutari* 'to help' [presupposed agentivity]	indefinite human NP
21. *Pippinu nni lu cunsigghiu cumunali avìa ammazzatu a lu sinnacu.* 'Pippinu killed the mayor during the meeting.'	*ammazzari* 'to kill' [potential agentivity]	definite human NP
22. *Pitrinu 'nmenzu la chiazza avìa ammazzatu a un maliutu.* 'Petrinu killed a criminal in the middle of the square.'	*ammazzari* 'to kill' [potential agentivity]	indefinite human NP
23. *Maria in du' simani avìa curatu a lu picciutteddu.* 'Maria cured the boy in two weeks.'	*curari* 'to cure' [potential agentivity]	definite human NP
24. *Titidda cu la so' midicina avìa curatu a un malatu*. 'Titidda cured a patient with her medecine.'	*curari* 'to cure' [potential agentivity]	indefinite human NP
25. *Ninetta fora la scola avìa pigghiatu a paroli a lu so' prufissuri.* 'Ninetta insulted his teacher outside the school.'	*pigghiari a paroli* 'to insult' [potential agentivity]	definite human NP
26. *Gina a lu mercatu avìa pigghiatu a paroli a un vinnituri.* 'Gina insulted a seller on the market.'	*pigghiari a paroli* 'to insult' [potential agentivity]	indefinite human NP
27. *Cola cu pugni e cavuci avìa struppiatu a lu sbirru.* 'Cola injured the policemen with punches and kicks.'	*struppiari* 'to cripple' [potential agentivity]	definite human NP
28. *Cuncittina cu lu cuteddu avìa struppiatu a un ospiti.* 'Cuncittina injured a guest with a knife.'	*struppiari* 'to cripple' [potential agentivity]	indefinite human NP
29. *Ciccu 'nmenzu la strata avìa firutu a la so' zita.* 'Ciccu hurt his girlfriend in the street.'	*firiri* 'to injure' [potential agentivity]	definite human NP
30. *Carlu cu lu vastuni avìa firutu a n'amicu.* 'Carlu hurt a friend with a stick.'	*firiri* 'to injure' [potential agentivity]	indefinite human NP
31. *Brizzita a la staziuni avìa vistu a lu prisidenti.* 'Brizzita saw the president at the train station.'	*viriri* 'to see' [unspecified for agentivity]	definite human NP
32. *Pitrina a lu mercatu avìa vistu a un politicu.*	*viriri* 'to see' [unspecified for agentivity]	indefinite human NP

'Pitrina saw a politician on the market.'		
33. *Ciccina tutta priata avìa taliatu a l'atturi.* 'Ciccina looked at the actor with admiration.'	*taliari* 'to watch' [unspecified for agentivity]	definite human NP
34. *Nzinu cu 'nteressi avìa taliatu a un scarparu*. 'Nzinu looked at/observed a shoemaker with interest.'	*taliari* 'to watch' [unspecified for agentivity]	indefinite human NP
35. *Lucia nni lu varveri avìa truvatu a lu so' zitu.* 'Lucia found her boyfriend at the barber's.'	*truvari* 'to find' [unspecified for agentivity]	definite human NP
36. *Nzina a lu cafè avìa truvatu a un cumpagnu*. 'Nzina accidentally found/met a friend at the bar.'	*truvari* 'to find' [unspecified for agentivity]	indefinite human NP
37. *Stefanu da tantu tempu canusci a lu parrinu.* 'Stefanu knows the priest for a long time.'	*canusciri* 'to know' [unspecified for agentivity]	definite human NP
38. *Pasquali da deci anni canusci a un avvucatu.* 'Pasquale knows a lawyer for ten years.'	*canusciri* 'to know' [unspecified for agentivity]	indefinite human NP
39. *Ancilina di lu primu mumentu un pò viriri a lu so' patruni di casa.* 'Ancilina hates the landlord from the first moment on.'	*'un putiri viriri* 'to hate' [unspecified for agentivity]	definite human NP
40. *Pippina di lu primu iornu un pò viriri a un vicinu.* 'Pippina hates a neighbour from the first day on.'	*'un putiri viriri* 'to hate' [unspecified for agentivity]	indefinite human NP

1.2 Version A and B of the questionnaire

The versions A and B of the questionnaire, recorded in MP3 audio format, are available here: https://uni-koeln.sciebo.de/s/kmQyHw8DHzUsg0t.

1.3 Form of sociodemographic data

Dati Personali

Anno di nascita: _____ Sesso: ☐ Femmina ☐ Maschio

Livello di studi:

☐ Scuola media

☐ Scuola superiore/liceo

☐ Università

☐ Altro: _____

In caso di studi universitari, campo di studi: _____

Luogo di nascita (*comune e provincia*): _____

Luogo di residenza (*comune e provincia*): _____

Da quando? (*anno*): _____

Lingua madre: _____

Seconda/e lingua/e: _____

Di solito Lei che lingua parla a casa?

☐ Italiano

☐ Siciliano

☐ Entrambi

☐ Altre: _____

Le viene meglio parlare in italiano o in siciliano?

☐ In italiano

☐ In siciliano

☐ Non c'è differenza

☐ Non so

1.4 Results of the individual items (ordered by verbs)

Agentivity	3a (def. NP & DOM)	3b (def. NP & no-DOM)	3c (indef. NP & DOM)	3d (indef. NP & no-DOM)
(i) entailed				
affènniri 'to offend'	3.917	3.545	4.000	4.583
cunsulari 'to console'	4.667	3.727	3.636	4.667
fari scantari 'to frighten'	4.500	3.727	4.182	4.833
fari siddiari 'to provoke'	4.000	4.750	3.083	3.818
nuciri 'to bother'	4.455	4.250	4.167	4.000
(ii) presupposed				
acchiappari 'to catch'	4.182	4.917	4.583	4.454
aiutari 'to help'	3.909	4.417	4.500	3.636
arristari 'to stop'	4.000	4.667	3.083	3.455
sarvari 'to save'	4.417	3.727	4.091	4.167
sentiri 'to hear'	4.833	3.545	3.273	4.417
(iii) potential				
ammazzari 'to kill'	3.250	4.273	4.909	4.000
curari 'to cure'	3.727	4.583	4.667	3.909
firiri 'to injure'	2.500	4.091	2.727	3.833
pigghiari a paroli 'to insult'	4.833	4.455	4.000	4.583
struppiari 'to cripple'	3.727	3.667	3.917	3.818
(iv) unspecified				
canusciri 'to know'	4.917	3.273	3.909	4.833
taliari 'to watch'	3.833	3.909	3.545	2.583
truvari 'to find'	4.364	4.667	4.750	4.000
'un putiri viriri 'to hate'	3.636	3.583	4.833	4.000
viriri 'to see'	3.727	3.417	4.833	3.818

Table 1. Mean acceptability values of the individual items in the four conditions 3a-d, ordered by verbs

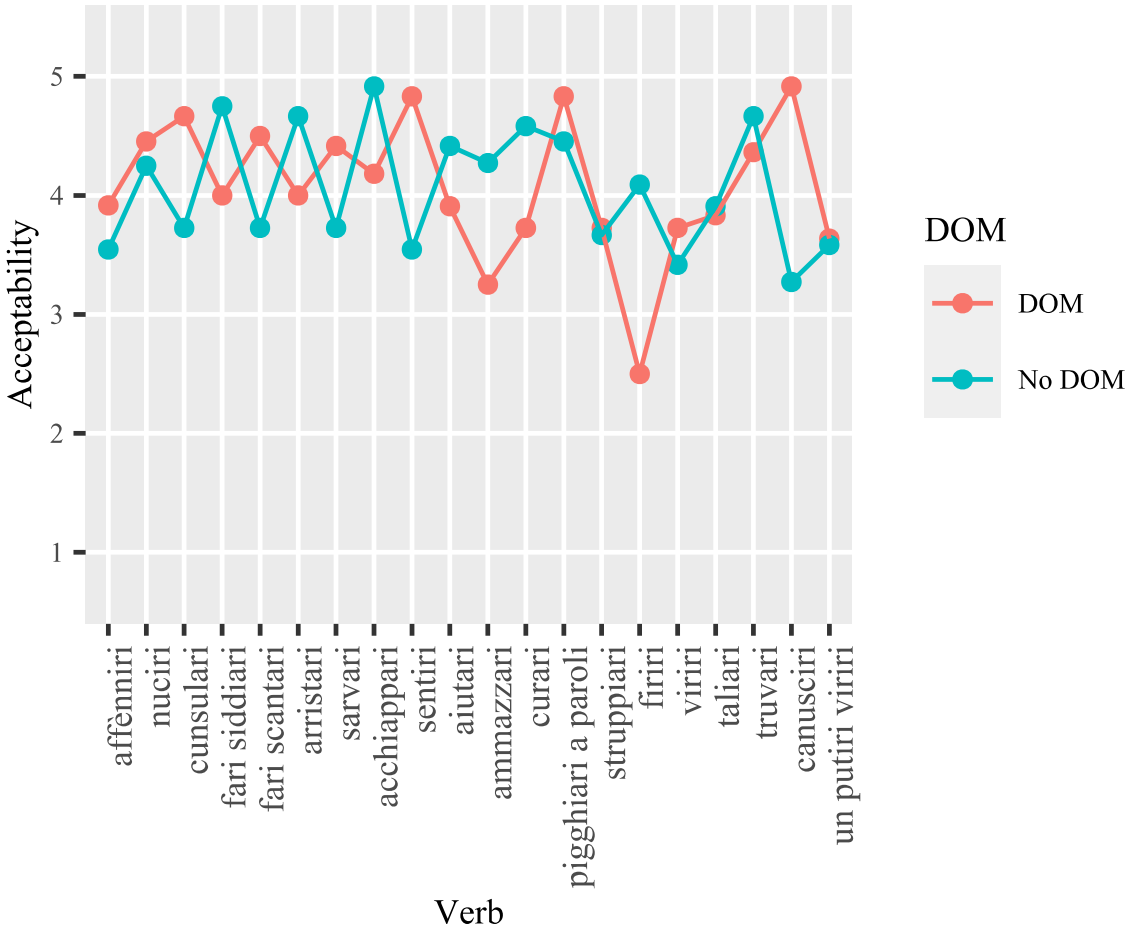


Figure 1. Mean acceptability values of the individual items in condition 3a (*Agentivity & definite NP & DOM*) and 3b (*Agentivity & definite NP & no-DOM*), ordered by verbs

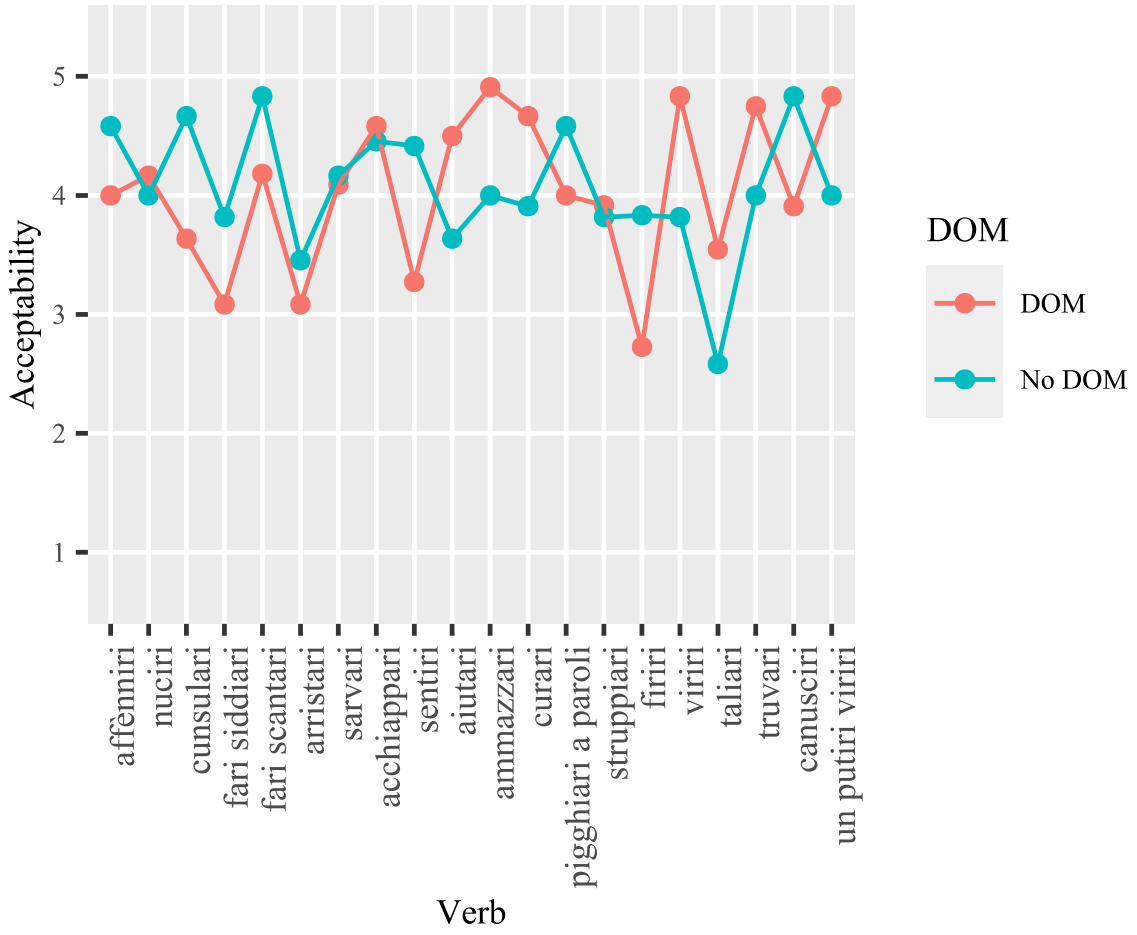


Figure 2. Mean acceptability values of the individual items in condition 3c (*Agentivity & indefinite NP & DOM*) and 3d (*Agentivity & indefinite NP & no-DOM*), ordered by verbs

Appendix 2. Acceptability judgement study on Catalan

2.1 Test items (Barcelona)

Test item (with DOM)	**Verb [communicated agentivity for the object]**
1. *El Guillem ha ofès a la patrona durant la reunió.* 'Guillem offended his boss during the meeting.'	*ofendre* 'to offend' [entailed agentivity]
2. *El Josep ha molestat a la companya de treball durant tot el dia.* 'Josep bothered his collegue all day long.'	*molestar* 'to bother' [entailed agentivity]
3. *La Laia ha consolat a la nena després de la seva caiguda.*	*consolar* 'to console' [entailed agentivity]

'Laia consoled the little girl after she had fallen.'	
4. *El Jordi ha provocat a l'agent de policia amb insults.* 'Jordi provoked the police officer with insults.'	*provocar* 'to provoke' [entailed agentivity]
5. *El Cesc ha espantat a la dona per diversió.* 'Cesc frightened the woman for fun.'	*espantar* 'to frighten' [entailed agentivity]
6. *El Marc ha aturat a la delinqüent davant de la botiga.* 'Marc stopped the criminal in front of the store.'	*aturar* 'to stop' [presupposed agentivity]
7. *El Carles ha salvat a la companya del foc.* 'Carles saved the collegue out of the fire.'	*salvar* 'to save' [presupposed agentivity]
8. *El Pau ha retingut a la lladre a l'estació de tren.* 'Agata caught the thief at the train station.'	*retener* 'to catch' [presupposed agentivity]
9. *La Sara ha sentit a la veïna al jardí.* 'Antrea heard the neighbour in the garden.'	*sentir* 'to hear' [presupposed agentivity]
10. *L'Èric ha ajudat a la jardinera al treball.* 'Èric helped the gardener at work.'	*ajudar* 'to help' [presupposed agentivity]
11. *El Pere ha matat a la criminal al parc.* 'Peter killed the criminal in the park.'	*matar* 'to kill' [potential agentivity]
12. *La Maria ha curat a la pacient en dues setmanes.* 'Maria cured the patient in two weeks.'	*curar* 'to cure' [potential agentivity]
13. *La Núria ha insultat a la companya a classe.* 'Núria insulted the classmate in class.'	*insultar* 'to insult' [potential agentivity]
14. *El Vicent ha lesionat a l'agent de policia amb un ganivet.* 'Vicent crippled the police officer with a knife.'	*lesionar* 'to cripple' [potential agentivity]
15. *La Judit ha ferit a l'adversària al camp de futbol.* 'Judit injured her opponent on the football field.'	*ferir* 'to injure' [potential agentivity]
16. *La Marta ha vist a la presidenta a l'aeroport.* 'Marta saw the president at the airport.'	*veure* 'to see' [unspecified for agentivity]
17. *La Paula ha contemplat a la ballarina amb admiració.* 'Paula looked at the dancer with admiration.'	*contemplar* 'to look at' [unspecified for agentivity]
18. *La Maria ha trobat a la nena al parc infantil.* 'Maria found the girl in the playground.'	*trobar* 'to find' [unspecified for agentivity]
19. *El Víctor coneix a l'escriptora de l'escola.*	*conéixer* 'to know' [unspecified for agentivity]

'Víctor knows the writer from school.'	
20. *La Marina ha odiat a la veïna des del primer dia.* 'Marina hated the neighbour from the first day on.'	*odiar* 'to hate' [unspecified for agentivity]

2.2 Test items (Girona)

Test item (with DOM)	**Verb** **[communicated agentivity for the object]**
1. *En Guillem ha ofès a la patrona durant la reunió.* 'Guillem offended his boss during the meeting.'	*ofendre* 'to offend' [entailed agentivity]
2. *En Josep ha molestat a la companya de treball durant tot el dia.* 'Josep bothered his collegue all day long.'	*molestar* 'to bother' [entailed agentivity]
3. *La Laia ha consolat a la nena després de la seva caiguda.* 'Laia consoled the little girl after she had fallen.'	*consolar* 'to console' [entailed agentivity]
4. *En Jordi ha provocat a l'agent de policia amb insults.* 'Jordi provoked the police officer with insults.'	*provocar* 'to provoke' [entailed agentivity]
5. *En Cesc ha espantat a la dona per diversió.* 'Cesc frightened the woman for fun.'	*espantar* 'to frighten' [entailed agentivity]
6. *En Marc ha aturat a la delinqüent davant de la botiga.* 'Marc stopped the criminal in front of the store.'	*aturar* 'to stop' [presupposed agentivity]
7. *En Carles ha salvat a la companya del foc.* 'Carles saved the collegue out of the fire.'	*salvar* 'to save' [presupposed agentivity]
8. *En Pau ha retingut a la lladre a l'estació de tren.* 'Agata caught the thief at the train station.'	*retener* 'to catch' [presupposed agentivity]
9. *La Sara ha sentit a la veïna al jardí.* 'Antrea heard the neighbour in the garden.'	*sentir* 'to hear' [presupposed agentivity]
10. *L'Èric ha ajudat a la jardinera al treball.* 'Èric helped the gardener at work.'	*ajudar* 'to help' [presupposed agentivity]
11. *En Pere ha matat a la criminal al parc.* 'Peter killed the criminal in the park.'	*matar* 'to kill' [potential agentivity]
12. *La Maria ha curat a la pacient en dues setmanes.* 'Maria cured the patient in two weeks.'	*curar* 'to cure' [potential agentivity]

13. *La Núria ha insultat a la companya a classe.* 'Núria insulted the classmate in class.'	*insultar* 'to insult' [potential agentivity]
14. *En Vicent ha lesionat a l'agent de policia amb un ganivet.* 'Vicent crippled the police officer with a knife.'	*lesionar* 'to cripple' [potential agentivity]
15. *La Judit ha ferit a l'adversària al camp de futbol.* 'Judit injured her opponent on the football field.'	*ferir* 'to injure' [potential agentivity]
16. *La Marta ha vist a la presidenta a l'aeroport.* 'Marta saw the president at the airport.'	*veure* 'to see' [unspecified for agentivity]
17. *La Paula ha contemplat a la ballarina amb admiració.* 'Paula looked at the dancer with admiration.'	*contemplar* 'to look at' [unspecified for agentivity]
18. *La Maria ha trobat a la nena al parc infantil.* 'Maria found the girl in the playground.'	*trobar* 'to find' [unspecified for agentivity]
19. *En Víctor coneix a l'escriptora de l'escola.* 'Víctor knows the writer from school.'	*conéixer* 'to know' [unspecified for agentivity]
20. *La Marina ha odiat a la veïna des del primer dia.* 'Marina hated the neighbour from the first day on.'	*odiar* 'to hate' [unspecified for agentivity]

2.3 Version A and B of the questionnaire (Barcelona and Girona)

The versions A and B of the questionnaire, recorded in MP3 audio format, are available here: https://uni-koeln.sciebo.de/s/kmQyHw8DHzUsg0t.

2.4 Form of sociodemographic data

Dades personals

Any de naixement: _____ Sexe: ☐ femení ☐ masculí

Nivell d'estudis:

☐ Escola primària

☐ Escola secundària

☐ Batxillerat

☐ Formació professional

☐ Universitat

☐ Altres: _____

En cas d'estudis universitaris, camp d'estudi:

_____

Lloc de naixement (*municipi i província*):

_____

Lloc de residència (*municipi i província*):

_____

Des de quan? (*any*): _____

Llengua(es) materna(es):

☐ Castellà

☐ Català

☐ Altres: _____

Llengües estrangeres:

_____

Quina llengua utilitzes normalment a casa?

☐ Castellà

☐ Català

☐ Altres: _____

En quina llengua et sents més a gust?

☐ Castellà

☐ Català

☐ M'és igual

2.5 Results of the individual items (ordered by verbs) (Barcelona and Girona)

Agentivity	Barcelona 2a (DOM)	Barcelona 2b (no-DOM)	Girona 2a (DOM)	Girona 2b (no-DOM)
(i) entailed				
ofendre 'to offend'	3.481	4.028	4.129	4.407
consolar 'to console'	3.407	3.944	4.032	3.704
espantar 'to frighten'	3.259	4.222	4.000	4.444
molestar 'to bother'	2.806	3.593	3.667	4.548
provocar 'to provoke'	3.583	4.074	3.481	4.419
(ii) presupposed				
ajudar 'to help'	1.972	2.185	2.259	2.484
aturar 'to stop'	3.250	4.370	3.815	4.645
retener 'to catch'	3.250	4.000	3.185	4.323
salvar 'to save'	3.308	4.541	3.613	4.259
sentir 'to hear'	3.346	4.324	3.968	4.481
(iii) potential				
curar 'to cure'	3.333	4.259259	3.296296	4.419355
ferir 'to injure'	3.000	4.222222	4.064516	4.148148
insultar 'to insult'	1.963	1.638889	4.225806	4.333333
lesionar 'to cripple'	3.306	3.740741	3.518519	4.483871
matar 'to kill'	2.741	4.027778	3.870968	4.000000
(iv) unspecified				
conéixer 'to know'	3.519	4.222	4.226	4.889
contemplar 'to watch'	3.370	4.333	3.968	4.481
trobar 'to find'	3.167	4.556	4.000	4.710
odiar 'to hate'	3.639	3.926	3.704	4.452
veure 'to see'	3.389	4.593	3.852	4.710

Table 2. Mean acceptability values of the individual items in condition 2a and 2b for Barcelona and Girona, ordered by verbs

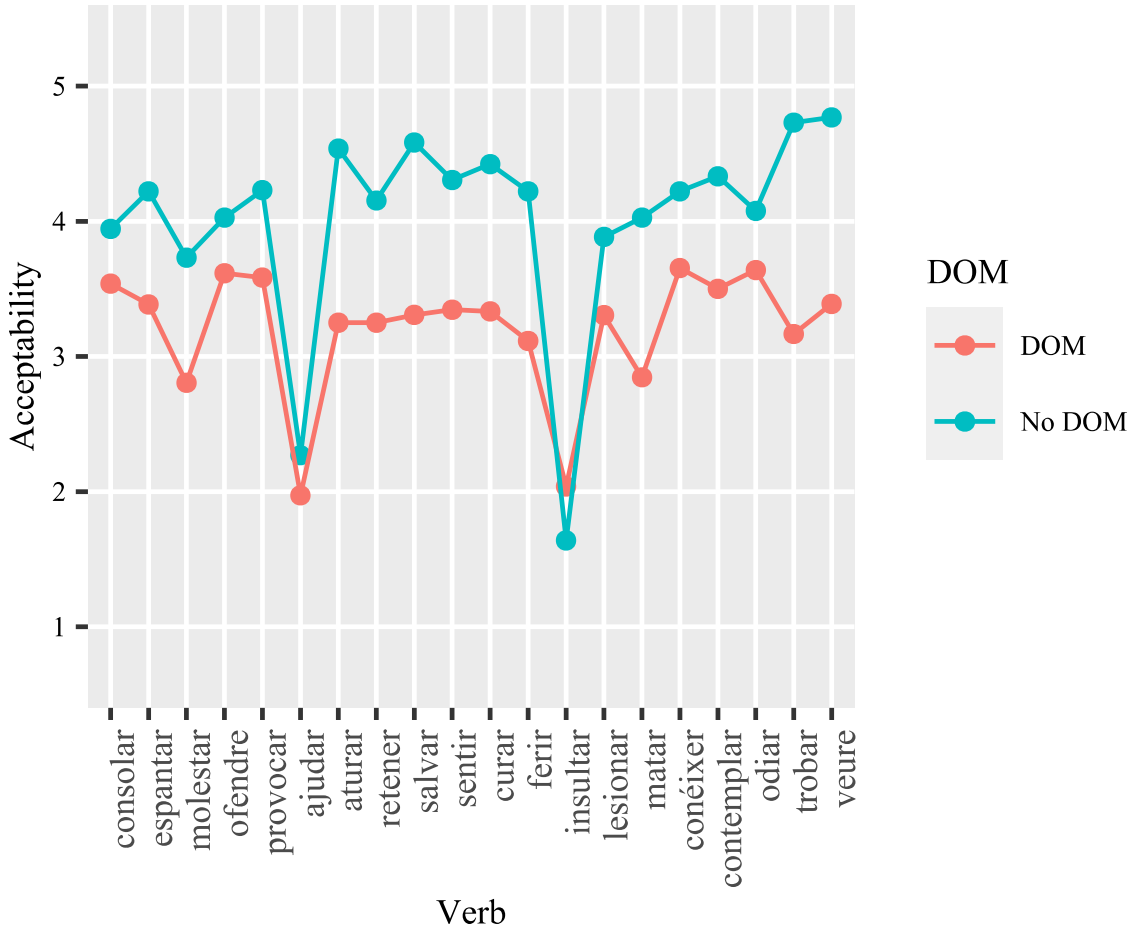


Figure 3. Mean acceptability values of the individual items in condition 2a (*Agentivity & DOM*) and 2b (*Agentivity & no-DOM*) for Barcelona, ordered by verbs

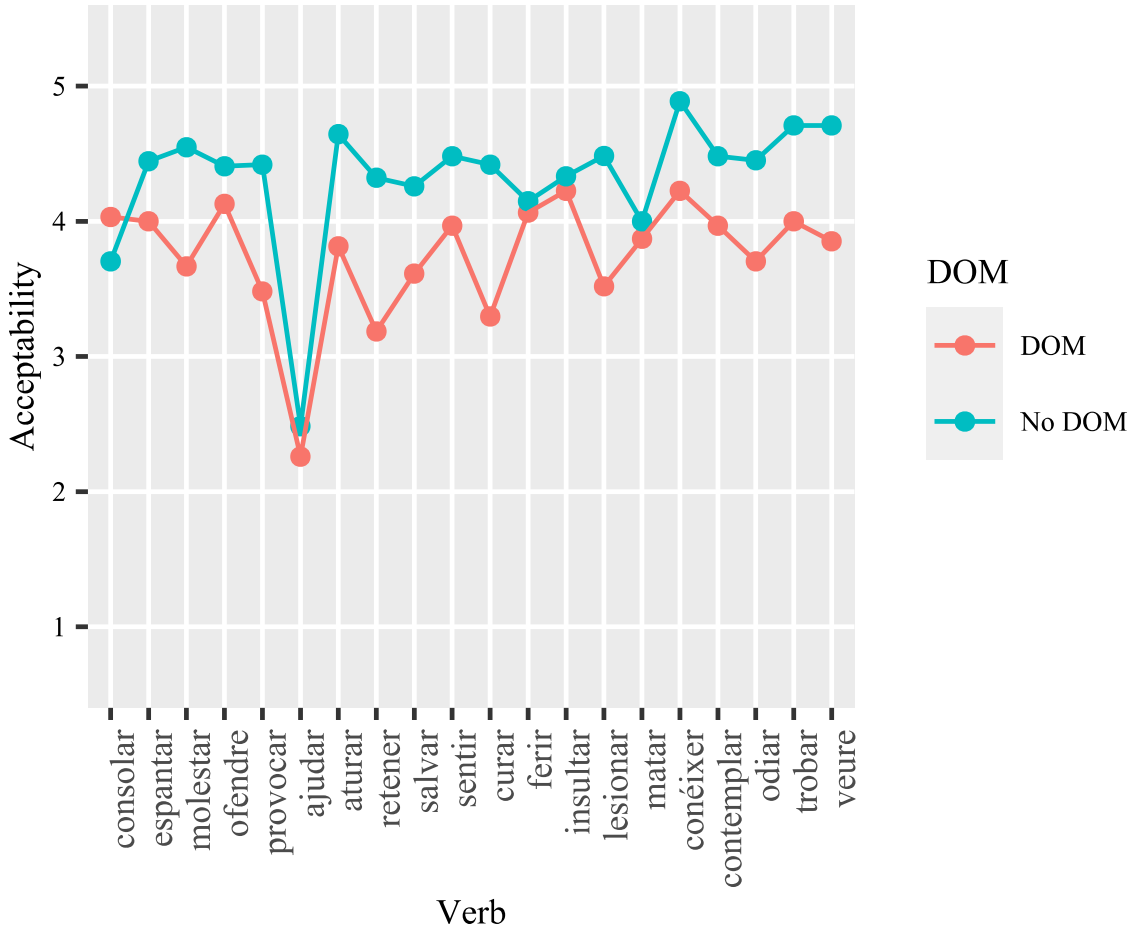


Figure 4. Mean acceptability values of the individual items in condition 2a (*Agentivity & DOM*) and 2b (*Agentivity & no-DOM*) for Girona, ordered by verbs