



Mia Batinić Angster, Marco Angster (Eds)

# The verbal kaleidoscope

Perspectives on the syntax  
and semantics of verbs



Mia Batinić Angster, Marco Angster (Eds)

# The verbal kaleidoscope

## Perspectives on the syntax and semantics of verbs

Zadar, 2024



University of Zadar  
Universitas Studiorum  
Jadertina | 1396 | 2002 |

**The verbal kaleidoscope  
Perspectives on the syntax  
and semantics of verbs**

NAKLADNIK  
Sveučilište u Zadru

ZA NAKLADNIKA  
Josip Faričić, rektor

POVJERENSTVO ZA IZDAVAČKU  
DJELATNOST SVEUČILIŠTA U ZADRU  
Lena Mirošević, predsjednica

UREDNICI  
Mia Batinić Angster  
Marco Angster

RECENZENTI  
Zrinka Kolaković  
Ivica Peša Matracki

LEKTURA  
Antonio Oštarić

FOTOGRAFIJA S NASLOVNICE  
Mia Batinić Angster

OBRADA FOTOGRAFIJE S NASLOVNICE  
Marco Angster

GRAFIČKO OBLIKOVANJE I PRIJELOM  
Grafikart d.o.o.

**The verbal kaleidoscope  
Perspectives on the syntax  
and semantics of verbs**

PUBLISHER  
University of Zadar

FOR THE PUBLISHER  
Josip Faričić, rector

PUBLISHING COMMITTEE OF  
THE UNIVERSITY OF ZADAR  
Lena Mirošević, chair

EDITORS  
Mia Batinić Angster  
Marco Angster

REVIEWERS  
Zrinka Kolaković  
Ivica Peša Matracki

PROOFREADING  
Antonio Oštarić

COVER PHOTO  
Mia Batinić Angster

COVER PHOTO EDITING  
Marco Angster

GRAPHIC DESIGN AND LAYOUT  
Grafikart d.o.o.

ISBN 978-953-331-488-4

# Table of contents

Table of contents .....	iii
List of contributors .....	iv
Preface .....	v
<b>1. Introduction</b> .....	1
Mia Batinić Angster, Marco Angster	
<b>2. An innovative approach to the study of verbs at the syntax-semantics interface</b> .....	7
Elisabetta Ježek	
<b>3. Using event structures to model symmetric and asymmetric verb meanings in Croatian aspectual pairs</b> .....	28
Costanza Marini	
<b>4. Monoclausal configurations with infinitives in Croatian</b> .....	63
Mia Batinić Angster	
<b>5. Croatian middle <i>se</i>-constructions</b> .....	93
Branimir Belaj	
<b>6. A diachronic change of the <i>reflexiva tantum</i> psychological verbs in Croatian</b> .....	133
Lidija Milković	
<b>7. Alternating constructions with <i>mittō</i> ‘send’: the encoding of third arguments</b> .....	172
Martina Giuliani	
<b>8. The encoding of the manipulee in Spanish: the competition between dative and accusative</b> .....	200
Bojana Mikelenić, Metka Bezljaj	
<b>9. Tests for distinguishing arguments and adjuncts and their application to Croatian locative phrases</b> .....	226
Matea Birtić, Ivana Brač	
<b>10. On the argument status of Instruments in Italian</b> .....	261
Alice Suozzi, Anna Cardinaletti, Gianluca E. Lebani	
<b>11. For a cognitive continuum underlying the distinction between arguments and adjuncts</b> .....	302
Alessandro Buttinelli	

## List of contributors

Marco Angster, Department of Linguistics, University of Zadar

Mia Batinić Angster, Department of Linguistics, University of Zadar

Branimir Belaj, Faculty of Humanities and Social Sciences, Josip Juraj  
Strossmayer University of Osijek

Metka Bezlaj, Department of Linguistics, University of Zadar

Matea Birtić, Institute of Croatian Language

Ivana Brač, Institute of Croatian Language

Alessandro Buttinelli, Roma Tre University

Anna Cardinaletti, Ca' Foscari University of Venice

Martina Giuliani, University of Pavia/University of Bergamo

Elisabetta Ježek, Department of Humanities, University of Pavia

Gianluca E. Lebani, Ca' Foscari University of Venice

Costanza Marini, University of Pavia/University of Bergamo

Bojana Mikelenić, Faculty of Humanities and Social Sciences,  
University of Zagreb

Lidija Milković, Faculty of Humanities and Social Sciences,  
University of Zagreb

Alice Suozzi, Ca' Foscari University of Venice

## Preface

The idea of this edited volume originates from the international conference *8<sup>th</sup> Zadar Linguistic Forum (ZLF2022) – Predicate: from polysemy to arguments* held at the University of Zadar in April 2022 and co-organized by the Department of Linguistics of the University of Zadar and the project *SARGADA – Syntactic and semantic analysis of arguments and adjuncts in Croatian* (Institute of Croatian Language). It collects contributions based on a selection of papers presented at the conference and is further enriched by a paper submitted via a call opened to the members of the Scientific board.

We thank the members of the international Scientific board of the conference who, along with us, participated in the double-blind review process of the proposals submitted for the conference: Boban Arsenijević (Karl-Franzens-University of Graz, Institute for Slavic Studies), Branimir Belaj (Department of Croatian Language and Literature, Faculty of Humanities and Social Sciences, Josip Juraj Strossmayer University of Osijek), Matea Birtić, Ivana Brač (Institute of Croatian Language, Zagreb), Carlo Geraci (CNRS, Institut Jean-Nicod & École Normale Supérieure, Paris), Elisabetta Ježek (Department of Humanities, University of Pavia), Ranko Matasović, Ida Raffaelli (Department of Linguistics, Faculty of Humanities and Social Sciences, University of Zagreb), Irena Zovko Dinković (Department of English, Faculty of Humanities and Social Sciences, University of Zagreb).

We would like to express our gratitude to Metka Bezljaj (Department of Linguistics, University of Zadar), Matea Birtić, Siniša Runjaić (Institute of Croatian Language, Zagreb), Costanza Marini (University of Pavia/ University of Bergamo) who took part in the organisation of the conference.

We are grateful to the University of Zadar and the Croatian Ministry of Science and Education which, along with the project *SARGADA*, financially supported the *8<sup>th</sup> Zadar Linguistic Forum*.

The papers submitted for this volume have been evaluated and categorized individually through a double-blind review process by peers affiliated with the following institutions: Institute of Croatian Language in Zagreb, Josip Juraj Strossmayer University of Osijek, Old Church Slavonic Institute in Zagreb, University of Turin, University of Zadar, University

of Zagreb. In addition, the whole volume has been reviewed by Ivica Peša Matracki (University of Zagreb) and Zrinka Kolaković (University of Klagenfurt/University of Hamburg). We would like to thank both the reviewers of the individual papers and of the entire volume for their help and valuable feedback.

MIA BATINIĆ ANGSTER

MARCO ANGSTER

# 1. Introduction

MIA BATINIĆ ANGSTER, MARCO ANGSTER

Verbs are often considered a major word class and a central part of the sentence, while the distinction between verbs and nouns is considered to be «one of the few apparently universal parts-of-speech distinctions» (Schachter & Shopen 2007: 5). However, the perspectives on how to distinguish verbs and nouns, i.e. on how to define each of them, are numerous and diverge from one another depending on the formal and functional criteria taken into account. Privileging functional criteria, Hopper & Thompson (1984) advocate for a discourse basis in the distinction between verbs and nouns. In a functional-typological perspective, Croft (2001) identifies “UNMARKED VERBS” in his conceptual space model of parts of speech as the portion of the conceptual space at the crossing of the propositional act of predication and of the semantic class of actions. On the other hand, other approaches to the definition of verbs highlight their formal features, such as the fact of being the head of a predicate (Dixon 2010: 38) or of a verb phrase seen as «the core of the sentence» and a syntactic projection (D’Alessandro *et al.* 2017: xvii).

Verbs as heads of predicates typically display a series of categories such as tense, aspect, mood, modality and evidentiality (Dixon 2010: 52–53). These categories can be marked on the verb as inflectional exponents – along with the agreement features of person, gender, and number – or via complex grammatical constructions – e.g. periphrastic tenses, modal verb constructions, evidential markers. Verbal categories such as voice and actionality are on the boundary between lexicon and grammar and can have crucial consequences on syntactic structure and semantic interpretation. At the crossroads between syntax, semantics, and the lexicon a central role is played by argument structure, a property of verbs that is reflected in their syntactic behaviour. A single verb can display multiple argument structures which in turn can, but need not be linked to different senses and display different values of voice and actionality.

This short and incomplete survey of the possible perspectives on verbs, of their features and of the different domains relevant for their analysis showcases their multifaceted nature. In the same way as when we look into a kaleidoscope the complex patterns that we see change as soon as we rotate one extremity, the perspective on verbs and the interpretation of their behaviour changes if we adopt different approaches or methodologies. The contributions included in this book constitute a sample of approaches to verbs, their arguments, and the constructions in which they participate.

## Overview of the volume

The study of verbs involves on the one hand accounting for how they interact with the syntactic level, especially as far as argument structure and valency changes are concerned. On the other hand, verbs make a crucial contribution to the meaning of sentences, constituting the core of predication.

The study of verbs at the syntax-semantics interface has thus been a prominent line of research which has been dominated by syntactically-oriented studies at least from the '90s, starting from Levin's (1993) seminal work on English verb alternations. Ježek in Chapter 2 surveys syntactically-oriented approaches to the study of verbs and highlights their limits. After this critical assessment she advocates a semantic-oriented approach that is more promising in clarifying the distinction between arguments and adjuncts by providing a semantically-grounded account for the optionality of the realization of arguments in syntax.

Marini's contribution (Chapter 3) continues on the lines of Ježek's semantically-grounded perspective on verbs to focus on the relationship between aspect and *Aktionsart* in the analysis of Croatian aspectual verb pairs. Starting from the asymmetrical verb pairs extracted from the CroaTPAS resource (Marini & Ježek 2019; Marini 2022a), she models the perfective/imperfective meaning (a)symmetries using the formalism of event structures (Pustejovsky 1991) showing that *Aktionsart* is relevant for explaining some of the observed meaning asymmetries.

In Chapter 4, Batinić Angster takes into consideration verbs embedding infinitives in Croatian. After presenting their classification into semantic classes, she groups the  $V_{\text{finite}} + V_{\text{non-finite}}$  configurations in Croatian into mono- and biclausal structures grounding her analysis on Grano's (2015) and Cinque's (2004, 2006) account of restructuring. Applying a series of tests

to Croatian, she finds that modal, aspectual and implicative verbs behave monoclausally with their embedded infinitives, whereas other subject-oriented/subject-control verbs such as factives and desideratives participate in biclausal configurations.

In the following two chapters the issue of the interplay between syntax and semantics is considered in the domain of voice. Chapter 5 is focused on middle *se*-constructions in Croatian. Branimir Belaj offers a typology of the uses of *se*, the light form of the reflexive pronoun *sebe*, and advocates for the replacement of the term *reflexive verb* with *middle verb*. His survey of middle *se*-constructions covers single-participant as well as reciprocal multi-participant middle verbs which are discussed in the light of Cognitive Grammar. Middle verbs, which show a low degree of participant distinguishability and of event elaboration (Kemmer 1993), are recognized as occupying a position on the transitivity hierarchy between transitive reflexive constructions and prototypical one-participant intransitive constructions. Additionally, middle verbs also appear to be close to passives and impersonal constructions in that they show aspects of process topicalization.

The domain of reflexive marking and its link to the transitive and the intransitive pole is considered from a diachronic point of view in Lidija Milković's contribution on Croatian *reflexiva tantum* verbs (Chapter 6). In her contribution the author focuses on psychological verbs and investigates their diachronic development drawing on their attestations in historical and etymological dictionaries of Croatian, but also of other Slavic languages. Considering a set of parameters relevant for valency changes, she observes that these parameters interact hierarchically. Beside this, *reflexiva tantum* psychological verbs seem to have emerged in Croatian only recently, while Croatian verbs in general appear to have been more transitive in the earlier history of the language than now.

In the remaining five chapters of the book the focus is directed to arguments, their encoding, alternations, and the argument-adjunct distinction. In Chapter 7 Giuliani considers the ditransitive Latin verb *mittō* 'send'. In her corpus-based analysis of the occurrences of this verb in Latin texts ranging from the 3rd century BCE to the 5th century CE she focuses her attention to the encoding of the third argument of this verb, i.e. a recipient-like argument, that can occur most prominently as a PP (*ad* + ACC: *ad me* 'to me') or as a bare dative (*mihi* 'to me'). The data are analysed taking into account three parameters: 1) the animacy of the theme

argument, 2) the semantic role, and 3) the affectedness of the recipient-like argument. The first results suggest that *ad* + ACC co-occurs with animate theme-arguments, marks (animate) directions and less affected recipients. The dative has a tendency to occur with inanimate theme arguments, true recipients which make use of the themes (usually gifts).

Mikelenić & Bezljaj in their paper (Chapter 8) take a closer look at the encoding of the manipulee in Spanish. They analyse the alternation between dative and accusative clitics, also known under the name of *leísmo*, in different Spanish-speaking areas. To do so, the authors compare corpus-extracted data on nine manipulation verbs that take a prepositional clitic object taking into consideration the number and the gender of the clitics and the type of complement clause. The use of clitics is found to be more conservative in Peninsular Spanish, favouring dative clitics, than in Latin American Spanish. Tendencies and area- and verb-specific variation observed in the competition between different strategies seem to favour García-Miguel's (2015) proposal of the absence of clear boundaries between direct and indirect object.

Birtić & Brač (Chapter 9) address the question of whether the locative phrases expressed as third complements in three-place verbs or as second complements in two-place verbs can be proven to be arguments or adjuncts. To this end, they apply a selection of tests for argumenthood issued from different frameworks (dependency and generative grammar) to a series of such verbs in Croatian. A thorough application of the selected tests leads the authors to conclude the following: only the third complements of *baciti* 'throw, cast' and *složiti* 'compose, put together, arrange, set' are confirmed as optional arguments; the results of the test for other verbs are inconsistent: different tests point to the optionality or the obligatoriness of their arguments; the replacement test supports the adjunct status of locative phrases; the *wh*-island test and the iterativity test when applied to locative phrases do not prove as useful diagnostic tools in determining their status.

Chapter 10 by Suozzi, Cardinaletti & Leboni investigates instruments in Italian. Instruments are characterized by a mismatch between the semantic and the syntactic level: they are often required by the verb, so they constitute a semantic argument, but are recognized as syntactically optional. The authors discuss three semantic and five syntactic criteria for argumenthood applied to instruments in Italian showing that they share more properties with arguments than adjuncts. Furthermore, in the vein

of Pustejovsky's (1995) proposal of a three-fold subdivision of arguments in Shadow, Default and Open arguments, they distinguish different types of instruments which can be omitted depending on features such as their semantic recoverability or their incorporation in the meaning of the verb. They claim that the syntactic omission of instruments is to be seen as a case of argument omission. The pattern of production/omission of instruments is confirmed in an analysis of spontaneous speech issued from three Italian spoken corpora.

Buttinelli (Chapter 11) advocates for a cognitive continuum underlying the argument-adjunct distinction. With the aim of proposing a reinterpretation of their relation in terms of a continuum of the representation of the sentence – of objective and subjective construal – rather than as a dichotomy, the author focuses on the behaviour of one Italian adverbial preposition of limitation – *riguardo (a)* NP 'regarding NP'. Considering its freedom of movement and optionality, this marginal adverbial behaves as adjunct, but it is found to cover three different linguistic metafunctions (Halliday 1970; Hengeveld 1989; Dik *et al.* 1990) characterized by different degrees of subjectivity: ideational, interpersonal, and textual. It is the closest to the argument domain when exhibiting the ideational metafunction as the most objective one. In that case it is in the scope of negation, affected by the question and is part of the focus.

The contributions to this volume cover a range of perspectives on the syntactic and semantic behaviour of verbs, pursuing different theoretical approaches and methods. They make up a composite view on the verb that can stimulate the readers to further deepen their knowledge in one of the directions delineated by the contributors. By collecting these papers, we hope to offer an encouragement to researchers to embark on the study of the verbal kaleidoscope.

## Appendix

### Acknowledgments

This work has been partly supported by the Croatian Science Foundation under the projects *Syntactic and Semantic Analysis of Arguments and Adjuncts in Croatian – SARGADA* (IP-2019-04-7896) and *Semantic-Syntactic Classification of Croatian Verbs – SEMTACTIC* (IP-2022-10-8074).

## References

- Cinque, G. (2004). Restructuring and functional structure. In: Belletti, A. (ed) *Structures and Beyond: The Cartography of Syntactic Structures. Volume 3*, 132–191. Oxford: Oxford University Press.
- Cinque, G. (2006). *Restructuring and functional heads: The cartography of syntactic structures. Volume 4*. Oxford: Oxford University Press.
- Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford University Press.
- D'Alessandro, R., Franco, I. & Gallego, Á. J. (eds) (2017). *The verbal domain*. Oxford: Oxford University Press.
- Dik, S. C., Hengeveld, K., Vester, E. & Vet, C. (1990). The hierarchical structure of the clause and the typology of adverbial satellites. In: Nuyts, J., Machtelt Bolkestein, A. & Vet, C. (eds) *Layers and levels of representation in language theory*, 25–70. Amsterdam/Philadelphia: John Benjamins.
- Dixon, R. M. (2010). *Basic linguistic theory. Volume 2: Grammatical topics*. Oxford: Oxford University Press.
- García-Miguel, J. M. (2015). Variable coding and object alignment in Spanish: A corpus-based approach. *Folia Lingüística*, 49(1), 205–256.
- Grano, T. (2015). *Control and restructuring*. Oxford: Oxford University Press.
- Halliday, M. A. K. (1970). Functional diversity in language as seen from a consideration of modality and mood in English. *Foundations of Language*, 6(3), 322–365.
- Hengeveld, K. (1989). Layers and operators in Functional Grammar. *Journal of Linguistics*, 25, 127–157.
- Hopper, P. J. & Thompson, S. A. (1984). The discourse basis for lexical categories in universal grammar. *Language*, 60(4), 703–752.
- Kemmer, S. (1993). *The Middle Voice*. Amsterdam: John Benjamins.
- Levin, B. (1993). *English Verb Classes and Alternations. A Preliminary Investigation*. Chicago: The University of Chicago Press.
- Marini, C. (2022a). CroaTPAS: uspoređivanje značenja vidskih parnjaka s namjerom istraživanja odnosa između vida, aktionsarta i glagolske polisemije u hrvatskom. *Croatica: časopis za hrvatski jezik, književnost i kulturu*, 46(66), 67–80.
- Marini, C. & Ježek, E. (2019). CROATPAS: A Resource of Corpus-Derived Typed Predicate Argument Structures for Croatian. In: Bernardi, R., Navigli, R. & Semeraro, G. (eds) *Proceedings of the 6th Italian Conference on Computational Linguistics (CLiC-It)*.
- Pustejovsky, J. (1991). The Syntax of Event Structure. *Cognition*, 41, 47–81.
- Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge (MA): MIT Press.
- Schachter, P. & Shopen, T. (2007). Parts-of-speech systems. Language typology and syntactic description. In: Shopen, T. (ed) *Language typology and syntactic description. 2nd Edition. Volume I: Clause structure*, 1–60. Cambridge: Cambridge University Press.

## 2. An innovative approach to the study of verbs at the syntax- semantics interface

ELISABETTA JEŽEK

Keywords: *verb class; argument structure; lexical semantics; semantic composition; argument incorporation*

In this contribution we start by critically evaluating the pros and cons of syntactically-oriented studies that were published from the '90s onward, with the goal of pinning down the semantic properties of verbs. First, we highlight the strengths of this approach. Second, we address its two major weaknesses: the inability to clearly distinguish arguments from adjuncts and with the shortcomings in accounting for the optionality in the syntactic realization of arguments. We argue that these weaknesses may be overcome by adopting a semantic-oriented approach from the start. Specifically, we propose that, unlike adjuncts, arguments are complements that play a role in influencing the meaning that the verb acquires in its contexts of use, and that syntactic optionality is licensed by semantic factors, in particular the degree to which an argument is incorporated in the verb on which it depends.

*Original scientific paper*

### 1. Introduction

Syntax-based approaches are an established model in theoretical linguistics for the study of verbs at the syntax- semantics interface. These approaches generally start by identifying subcategorization frames and their alternations for each verb, and then analyse them with the goal of identifying correlations between the syntactic behaviour and the semantic properties (Levin 1993). Despite the results that have been achieved, these approaches have shown not to be without problems. In this contribution we propose a complementary approach to the study of verbs at the syntax- semantic interface, in which the analysis is performed starting from the semantics of the verbs and that of their arguments. After reviewing the nature of the meaning of verbs and its primary constituents (section 2), we critically

evaluate the syntactically-oriented studies that have been systematically pursued from the '90s onward, with the goal of pinning down the semantic properties of verbs (section 3). First, we highlight the strengths of this approach. Second, we address its two major weaknesses: the inability to clearly distinguish arguments from adjuncts and the difficulty to account for the optionality in the syntactic realization of arguments (section 4). We argue that these weaknesses may be overcome by adopting a semantic-oriented approach from the start (section 5). In particular, we propose that, unlike adjuncts, arguments are complements that play a role in influencing the meaning that the verb acquires in the context of use, and that syntactic optionality is licensed by semantic factors, in particular the degree to which an argument is incorporated in the verb on which it depends.

## 2. The meaning of verbs

The denotation of a verb includes at least three dimensions which, together, constitute different aspects of the same object, namely the event that the verb expresses.<sup>1</sup> These are: the temporal structure of the event, the participants of the event (the arguments or valency of the verb), and the inherent nature of the event. Together, these components constitute the three main dimensions of analysis that have been pursued in linguistic studies – jointly or individually – to classify verbs. We review them briefly in the next subsections.

### 2.1. The temporal structure of the event

The event expressed by the verb is placed in time and, if it is endowed with duration, it has an internal structure that can be broken down into phases or intervals. The analysis of these phases is generally implemented through the use of linguistic tests such as the adverbial modification introduced by the expressions 'in/at *x* time' and 'for *x* time'. The applicability of the first would reveal that the event expressed by the verb includes a telic or culminative component; that of the second would exclude it. Based on this test, it is possible to distinguish verbs expressing *states* (*possess*, *remain*),

---

<sup>1</sup> In the following, we use the term *event* to refer to what Emmon Bach (1986: 6) called *eventualities*, i.e. static and dynamic situations.

*processes* (*work, sleep, walk*), all admitting ‘for x time’ but not ‘in x time’; punctual or culminative events (*find, arrive, burst*) that admit ‘at x time’ but not ‘for x time’; incremental or gradual events (degree achievements, incremental theme- or multi-scalar verbs: *increase, cool down*), which admit both adverbials. The event receives a specific temporal encoding when it becomes the denotation of a verb (i.e. its meaning). In principle, there may be a mismatch between an event “as it is observed in the world” and the way it is encoded in the meaning of verbs, although this is not the norm.<sup>2</sup>

## 2.2. Event participants

The event expressed by the verb usually has participants, i.e. animate/inanimate, concrete/abstract entities (including places and time periods) that hold a relation of *participation* with the event itself, playing different roles: the agent, the experiencer, the patient, the receiver, the origin, the destination, the instrument, etc. The linguistic encoding of the event makes some of these participants grammatically relevant: these are the participants called *arguments* or *actants*, which, contrary to other complements (the so-called *adjuncts* or *circumstantials*; for the latter, cf. Tesnière 1959), must be mandatorily expressed in the syntax because they complete the meaning of the verb. Depending on the number of arguments it is possible to distinguish zero-argument (*snow, flash*), one-argument (*sleep*), two-argument (*abolish, inhabit, dwell*), up to three-argument (*give, dedicate*) verbs. The existence of verbs with four arguments (*translate*) is a matter of controversy (Levin & Rappaport 2005). When adopting the notion of argument, an interesting question arises: which aspect(s) of the event may become an argument? Typically, participants or properties of participants in the event (including other events) may become arguments (such as the intentionality of the agent, the causes of the event, instruments, psychological states). Moreover, in the linguistic encoding of the event, some participants may be *incorporated* in the verb semantics, whereas other participants may be located in the background, an issue to which we return later.

Besides bearing a role, which is assigned to them by the verb, the arguments display semantic properties, in particular they belong to a

---

<sup>2</sup> For the inherent temporal properties of the event expressed by verbs, often referred to as *Aktionsart* or *Actionality*, see the taxonomies of Vendler (1967), Dowty (1979), and subsequent ones inspired by those.

semantic class (Physical object, Human, Institution, Animal, Artifact, Abstract entity etc.). This class must satisfy the *selectional restriction* (Chomsky 1965) or *preference* (Wilks 1975) that the verb requires: for example, the It. verb *abitare* ‘inhabit, live in/at’ only admits Animates in subject position, preferably Humans and less typically Animals, and rules out Institutions, as the example in (1) shows:

- (1) a. *Luca abita a Milano.*  
       ‘Luca lives in Milan.’  
       b. \**L’ONU abita a New York.*  
       ‘The UN lives in New York.’

Finally, arguments require a specific syntactic realization (Subject, Direct Object, Prepositional Complement, etc.) which together constitutes the subcategorization frame for the verb, for example, for *abitare* the subcategorization frame of (1) is <Subj> *abita* <PP>.

### 2.3. The inherent meaning

The inherent nature of the event, once it is linguistically encoded, allows one to identify verb classes such as motion verbs (*go, walk*), verbs of manner (*slip*), verbs of perception (*feel, hear*), verbs of cognition (*understand, grasp*), or, on a more abstract level, verbs indicating reciprocity, reflexivity, and so forth. Several attempts to define a list of verb classes on semantic grounds have been attempted.<sup>3</sup> For example, in the WordNet Project (Fellbaum 1998), a list of 15 classes for verbs (called *supersenses*) has been defined, in order to provide the collaborators working on the project with a broad initial classification for lexicon items. The proposed classes are listed in Table 1, where the first column reports the name of the class and the second column a set of English examples:

---

<sup>3</sup> Batinić Angster (*this volume*, Chapter 4) for example classifies Croatian verbs taking an infinitive in different semantic classes showing how the semantics of a verb affects its syntactic behaviour.

Table 1. Supersenses for Verbs in WordNet (Fellbaum 1998)

Verb Class	Examples
Body	grooming, dressing and bodily care
Change	size, temperature change, intensifying
Cognition	thinking, judging, analysing, doubting
Communication	telling, asking, ordering, singing (animal sounds)
Competition	fighting, athletic activities
Consumption	eating and drinking
Contact	touching, hitting, tying, digging
Creation	sewing, baking, painting, performing (cooking)
Emotion	feeling
Motion	walking, flying, swimming
Perception	seeing, hearing, feeling
Possession	buying, selling, owning
Social	political and social activities and events
Stative	being, having, spatial relations
Weather	raining, snowing, thundering

The proposed classes raise a number of issues from a linguistic perspective: what is the use of the supersenses? What do they tell us about the syntactic behaviour of verbs, the inferences they give rise to, the participants they involve, the manner in which the event takes place, the purpose of the event? These questions together lead to a broader question: on what grounds is it possible to group verbs in semantic classes?

The operation of identifying semantic classes, that is, groupings of verbs with similar meanings, and organise them in a hierarchy that distinguishes for example between types of motion verbs, types of perception verbs, and so on, is intuitively simple but technically complex (Ježek 2016: 128). In general, the difficulties encountered can be said to derive primarily from the fact that the meaning of verbs consists of a bundle of features with different semantic prominence. Consider for instance the case of verbs describing a change of position, like *sit*: although the action of sitting involves the motion of the person performing the action (a motion we may characterize as “internal”), this is not the

prominent feature in the meaning of the verb, and it would appear odd to classify *sit* as a verb of motion on a par with *enter* and *exit* (it is, in fact, categorized as a ‘verb of assuming a position’ in Levin 1993). But how is semantic prominence to be defined? Several scholars contend that prominent features are those that are transparent in the syntax, i.e. those that influence a verb syntactic behaviour (Levin & Rappaport Hovav 2005: Chapter 1); nevertheless, linking an observed syntactic behaviour to the appropriate semantic component is a notoriously arduous task.

Another problem is that verbs encode only some aspects of the event they denote while presupposing others. This distinction between denoted and presupposed information is by no means easy to make, although there are clear cases. *Arrive*, for example, presupposes motion but encodes, in fact, the result/effect of such motion, consisting in the fact that the person or thing arrived is located in a place which differs from the one it was located in before the arriving event took place. On this ground, it is reasonable to question whether *arrive* should be considered a verb of change of location rather than a verb of motion, and which criteria are eligible to distinguish between the two.

### 3. Traditional approaches: from syntax to semantics

The traditional classification of verbs rests on the distinction between transitive verbs and intransitive verbs, i.e. about the ability of a verb to be accompanied or not by a direct object. A more refined classification is based on the observation of the syntactic alternations allowed by the verbs, i.e. of the range of syntactic configurations that each verb can present. This may or may not include the number of arguments (cf. section 2.2). For example, in Italian, some verbs only allow transitive uses (*abolire* ‘abolish’, *indossare* ‘wear’, *affittare* ‘rent’), others only allow intransitive uses (*camminare* ‘walk’, *russare* ‘snore’, *arrivare* ‘arrive’, *cadere* ‘fall’), and others allow both (*affondare* ‘sink’, *guarire* ‘heal’). Looking into those alternations, and following the influential work of Levin (1993), in Ježek (2003) we proposed to classify Italian verbs based on the presence or absence of four different syntactic realizations: transitive use (TR), intransitive use with auxiliary *to have* (INTR AV, unergative), intransitive use with auxiliary *to be* (INTR ES, unaccusative), intransitive use with pronominal marker *si* (and *to be* as

auxiliary, INTR PRON).<sup>4</sup> The result is a list of 15 classes, which exhausts all the possible combinations of these four syntactic realizations, as shown in Table 2, which provides an example of a verb for each class.

Table 2. Italian verb classes and syntactic alternations (Ježek 2003)

Class	Verb	TR	INTR AV	INTR ES	INTR PRON
1	<i>abolire</i> 'abolish'	x	–	–	–
2	<i>camminare</i> 'walk'	–	x	–	–
3	<i>cadere</i> 'fall'	–	–	x	–
4	<i>pentirsi</i> 'repent'	–	–	–	x
5	<i>squillare</i> 'ring'	–	x	x	–
6	<i>appropfitare</i> 'take advantage'	–	x	–	x
7	<i>ammuffire</i> 'get mouldy'	–	–	x	x
8	<i>sedimentare</i> 'settle'	–	x	x	x
9	<i>mangiare</i> 'eat'	x	x	–	–
10	<i>affondare</i> 'sink'	x	–	x	–
11	<i>alzare</i> 'raise'	x	–	–	x
12	<i>ingiallire</i> 'yellow'	x	–	x	x
13	<i>continuare</i> 'continue'	x	x	x	–
14	<i>chiudere</i> 'close'	x	x	–	x
15	<i>bruciare</i> 'burn'	x	x	x	x

To understand the table, consider the verb *abolire* 'to abolish' in (2), which presents only a transitive use and therefore falls into class 1, and the verb *russare* 'to snore' in (3) that exhibits only an intransitive use with the auxiliary *avere*, and falls into class 2:

- (2) *abolire* TR            *Il governo ha abolito la pena di morte.*  
                                   'The Government has abolished the death penalty.'  
                                   INTR AV    \**Il governo ha abolito.* \*'The Government has abolished.'  
                                   INTR ES    \**Il governo è abolito.* \*'The Government is abolished.'<sup>5</sup>  
                                   INTR PRON \**Il governo si è abolito.* \*'The Government PRON is abolished.'

<sup>4</sup> The classification was resumed, among others, in Salvi & Vanelli (2004).

<sup>5</sup> Note that 'The government is abolished' is grammatical in the stative reading (as in 'This may mean that the government is abolished') but not in the dynamic reading, whereas the latter is required by intransitive uses.

- (3) *russare* TR            \**Il raffreddore ha russato Gianni.* \*‘The cold snored Gianni.’  
 INTR AV            *Gianni ha russato.* ‘Gianni has snored.’  
 INTR ES            \**Gianni è russato.* \*‘Gianni is snored.’  
 INTR PRON        \**Gianni si è russato.* \*‘Gianni PRON is snored.’

The case of *bruciare* ‘to burn’ is more complex, as it admits all four uses, as shown in (4) (class 15):

- (4) *bruciare* TR            *L’incendio ha bruciato la casa.* ‘The fire has burned the house.’  
 INTR AV            *La casa ha bruciato per ore.* ‘The house has burned for hours.’  
 INTR ES            *La casa è bruciata.* ‘The house is burned.’<sup>6</sup>  
 INTR PRON        *La casa si è bruciata.* ‘The house PRON is burned.’<sup>7</sup>

The classification we performed based on a large sample of Italian verbs (cf. Ježek 2003 for details on the dataset) shows that the properties mostly represented in the syntax of the verbs (particularly, in relation to the syntactic alternations that the verbs exhibit) are those related to the *event type*, that is, the aspectual or actional properties (process vs. state vs. change of state, cf. section 2.1), *agency*, and the presence or absence of *causality*. For example, TR verbs (that is, verbs of class 1 that only admit a transitive use) tend to be lexically agentive (*abolire* ‘abolish’), i.e. they do not occur with unintentional subjects.<sup>8</sup> Verbs that alternate between TR and INTR AV (class 9), on the other hand, mainly denote processes (*mangiare* ‘eat’, *bere* ‘drink’)

<sup>6</sup> See the previous note concerning stative vs. dynamic readings.

<sup>7</sup> It is well known that the pronominal marker *si* may occur in Italian in several constructions with different semantic values along the parameter of voice (reflexive, middle and passive respectively). In the current study we consider only intransitive constructions, where *si* marks spontaneity/lack of causalness, event completion, or both (middle uses). We exclude reflexive constructions in which *si* signals coreference with the subject (*Luca si è vestito* ‘Luca got dressed’) and passive constructions (*si vedono le stelle* ‘the stars are seen’, *si abolirono le tasse* ‘taxes were abolished’). For an overview of pronominal uses in Italian see Ježek (2003: 130–141). On the passive reinterpretation of the *si* + V construction in Italian, see Cennamo (2001).

<sup>8</sup> Note that most verbs that are generally classified as agentive, in certain contexts allow for non-agentive uses, such as *ferire* ‘injure’ (involuntarily). This is not the case of TR verbs: unlike *abolire*, *ferire* does not belong to this latter class as it admits also an INTR PRON use (*ferirsi*).

instead of states or happenings. Finally, verbs that alternate between TR and INTR ES and/or PRON (classes 10 to 12) describe predominantly changes of state (*ingiallire* ‘become yellow’).

## 4. Problems with syntax-first approaches

As stated above, the syntactic alternation model represents a refined approach with respect to the model of analysis based exclusively on transitivity/intransitivity. If we also add the number of arguments, we obtain an effective procedure to analyse the linguistic encoding of the participants in the event expressed by the verb. However, there are at least two weaknesses, which are connected to each other. The first is the well-known difficulty in drawing a distinction between arguments and adjuncts with respect to a specific verb. The second is the optionality of some arguments, that questions the very notion of argument, which is grounded on the idea that arguments are those complements which require mandatory syntactic realization. The first of these problems is exemplified in (5) where the adverbial ‘in front of the hotel’ has an uncertain status (argument or adjunct) with respect to the verb *arrive*, while the second is exemplified in (6), in which the direct object of the verb *drink*, expressed in (6a), is absent in (6b):<sup>9</sup>

- (5) *Gianni rilesse quel che aveva scritto fin quando il taxi non arrivò davanti all'albergo.*  
 ‘Gianni reread what he had written until the taxi arrived in front of the hotel.’
- (6) a. *I ragazzi bevono birra al pub.*  
 ‘The boys are drinking beer in the pub.’  
 b. *Il vecchio si riposò, bevve e fu assalito da uno strano pensiero.*  
 ‘The old man rested, drank, and was haunted by a strange thought.’

In the following sections, we examine two solutions we propose for solving these two problems. Our proposal highlights the role of semantics in determining the distinction between arguments and adjuncts and in licensing syntactic optionality.

<sup>9</sup> All examples are derived and adapted from the ItTenTen 2020 corpus (about 13 billion words; Jakubíček *et al.* 2013), queried through the Sketch Engine online platform (<https://app.sketchengine.eu>; Kilgarriff *et al.* 2004).

## 5. Innovative Approach: Starting from the Semantics

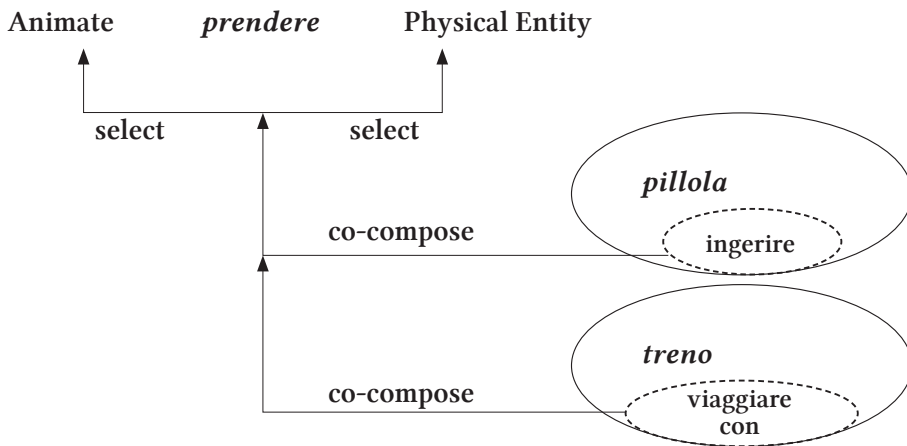
In contrast with the common practice that tackles the problems highlighted in section 4 starting from a syntactic perspective, we propose that those phenomena are better understood if the analysis is performed starting from the semantic level, particularly from the assumptions that, i.) the meanings of the arguments play an active role in the semantic composition with verb (contrary to adjuncts), and that ii.) arguments may be incorporated at different degrees into the inherent meaning of the verb. According to this proposal, the syntactic level is secondary and dependent on the semantic one. We first review our semantic approach, and then apply it to the problems raised in section 4.

### 5.1. Enriched composition in semantics

Our proposal departs from the traditional semantic account, according to which polysemy is a list of pre-defined senses stored in each lexical entry, and embraces a contextual approach, in which words have a core meaning that is flexible and varies from occurrence to occurrence as a function of the interaction with the other words they combine with, and of the situation of utterance (Recanati 2012). Accordingly, our approach departs from the principle of compositionality outlined in formal semantics (Partee 1995: 153), which claims that the meaning of a complex linguistic expression is determined by the sum of the meanings of its component parts, and the way they are put together. In contrast to this view, we assume that the meanings of words not only add up but also influence each other when they combine: from a theoretical point of view, we move from composition seen as an additive operation, to *enriched composition* (Jackendoff 1997) and *co-composition* (Pustejovsky 2012). Specifically, according to the principle of co-composition, the meaning of a complex expression is the result of the mutual adjustments that take place between the meanings of the combining words. Under this line of reasoning, one can argue that in the combination of a verb with its argument, the meaning of the verb is co-dependent on the meaning of its arguments, and that the variation in the semantics of arguments co-causes a variation in the verbal meaning. An example applied to the light verb *prendere* ‘to take’ is provided in Figure 1, where we graphically represent the co-composition with two different direct object arguments, i.e. *pillola* ‘pill’ and *treno* ‘the train’, resulting in two different

meanings for the verb ('ingest' and 'travel with' respectively). The example is based on the assumption that the meanings of nouns, represented by the oval shapes in the figure, include in their sublexical information the typical action in which the denoted entity is used.<sup>10</sup> When the verb selects its arguments (the direct object Physical Entity in our case), these actions (*ingest* for *pill* and *travel with* for *train*) are activated by the co-composition mechanism, and fused with the meaning of the verb. This operation results in two meanings for the verb *prendere* dependent on the meaning of the argument, namely *ingerire* 'ingest' and *viaggiare con* 'travel with'.

Figure 1. Co-composition of *prendere* + direct object



Co-composition is a unifying principle that applies to different syntactic roles (Subjects, Direct Objects, etc.) and to a wide range of semantic phenomena, including the shift from a literal to a figurative meaning of verbs, as in the examples below, that can be all accounted for in co-composition terms as illustrated in Figure 1.<sup>11</sup>

<sup>10</sup> This idea has been formalized in different ways in the semantic literature: see the notion of *hidden event* and *telic quale* proposed in Pustejovsky (1995).

<sup>11</sup> In (7) we provide a synonym for each sense that the verb acquires in combination with the different arguments.

- (7) a. *divorare la pasta* ‘devour the pasta (=eat)’ | *divorare un romanzo* ‘devour a novel (=read)’  
 b. *aprire il libro* ‘open the book (=unfold)’ | *aprire un dibattito* ‘open a debate (=begin)’

In conclusion, by adopting the view proposed in this section, the perspective of analysis of verbs and their arguments is reversed: instead of claiming that the number and type of arguments varies according to the meaning of the verb, as in the traditional account, we claim that the meaning of the verb varies according to the semantics of the argument(s).<sup>12</sup> In the next sections, we take this background into account and apply it to the analysis of the two issues raised in section 4.

## 5.2. The argument-adjunct distinction

In section 5.1 we moved away from the traditional view according to which, in the semantic composition, a word behaves either as an active functor or as a passive argument, and embraced the idea that the arguments can both complete and influence the meaning of the verb they combine with. This move has important consequences on the way one can account for the argument-adjunct distinction.<sup>13</sup> In particular, we can assume that the ability of complements to co-compose with the verb does not apply to the adjuncts, which are instead *inactive* on verbal semantics and never play a role in determining the sense that the verb assumes in its context of use (Ježek 2021). According to this line of reasoning, the verb *partire* ‘to leave’ has a single argument in the example in (8a) (the subject *la macchina* ‘the car’) not for syntactic reasons but because that single argument is sufficient for the verb to take on the meaning of ‘to start (off)’. As it happens, that single argument is essential to ensure that the verb acquires this meaning and not another possible meaning it could acquire (e.g. ‘go away, move away from a place’). Note that in (8a), the adverbial expressions that follow the

<sup>12</sup> A reviewer asked an interesting question whether arguments – which influence the verb meaning – also influence the selection of the verb. In this regard, we refer the reader to a different contribution where the selection of verbs by nouns is discussed using the notion of collocational meaning (Ježek 2014).

<sup>13</sup> For a survey of tests for argumenthood applied to Croatian, see Birtić & Brač (*this volume*, Chapter 9).

verb are instead adjuncts, as they neither complete nor modify the meaning of the verb, which is assigned to the verb by the subject argument alone. In example (8b), instead, the verb *partire* ‘to start’ takes two arguments, as long as both the subject *dolore* ‘pain’ and the expression indicating the source (*dal braccio* ‘from the arm’ / *a livello del collo* ‘at the neck level’) are necessary to complete the meaning that the verb takes on in the context (‘originate from’). In other words, according to our account, the unacceptability of the expression \*‘the pain starts’ is the result of a semantic inconsistency and not a syntactic one, or rather, the syntactic inconsistency is due to semantic factors:

- (8) a. *La macchina non parte più / mai al primo colpo.* (1 argument)  
 ‘The car no longer starts / never starts the first time.’  
 b. *Il dolore parte dal braccio / a livello del collo.* (2 arguments)  
 ‘The pain starts from the arm / at the neck level.’

The assumption that the distinction between arguments and adjuncts originates in the semantic composition of the verb and its complements, and that only the arguments act as functors in this process, is fully supported by the empirical observation that the valence of the verb varies according to its meaning. When one attempts to draw the distinction between arguments and adjuncts, what must be observed, therefore, is not the verb but the meaning it acquires (that is, the event that it expresses) in the specific context of use.

We applied the approach described above to identify verb senses and to tell apart arguments from adjuncts in the construction of the T-PAS resource (Ježek *et al.* 2014), an inventory of argument structures for Italian verbs focused on the semantics of arguments.<sup>14</sup> In T-PAS, for each verb meaning, a specific typed predicate-argument structure (T-PAS, informally called *pattern*) is provided. An example of pattern for the verb *guidare* ‘to drive’ in its ‘operate’ sense is [Human] *guida* [Road Vehicle]. Arguments are defined in terms of semantic classes notated between square brackets, called *semantic types*. Patterns are corpus-derived, i.e. they are acquired through manual clustering and annotation of corpus instances, following the Corpus Pattern Analysis (CPA) methodology (Hanks 2013). Several studies

<sup>14</sup> The T-PAS project was developed at the Department of Humanities of University of Pavia, with the technical support of Lexical Computing Ltd. The resource can be freely accessed and downloaded at <https://tpas.unipv.eu>.

performed on T-PAS data<sup>15</sup> have shown that co-composition is a useful principle for discriminating verb senses and in telling apart arguments from adjuncts. Also, T-PAS data allow to define semantic classes for verbs based on the semantics of the arguments (subject, direct object, prepositional complement, also in combination). A list of four classes is reported in Table 3: for example, **vehicle\_verbs** are verbs that take Vehicles as Subject or Object, and so forth.

Table 3. Verb classes based on the semantic class of their arguments

CLASS	EXAMPLES
vehicle_verbs	<i>accelerare, frenare, parcheggiare, guidare, svoltare, rallentare, sterzare, sbandare</i>
hair_verbs	<i>annodare, pettinare, sistemare, spazzolare, spettinare, tagliare, tingere, tirare</i>
money_verbs	<i>accettare, accreditare, attirare, cedere, corrispondere, depositare, distribuire, emettere, guadagnare, incassare, pagare, prestare, riscattare, scommettere, vincere</i>
plants_verbs	<i>fiorire, crescere, germinare, germogliare, riprodursi, nascere, sbocciare, appassire</i>

Note that the verbs in the classes may vary considerably as regards their semantic spectrum, and consequently their distribution. For example, in the case of **hair\_verbs**, *tirare* ‘pull’ has a broad semantic spectrum and one might arguably claim that it is unusual or far less frequent with an argument like *hair* than with other arguments. This is confirmed by corpus analysis, according to which the typicality score of the combination of *capelli* ‘hair’ with *tirare*, calculated through the Sketch Engine platform<sup>16</sup> is lower (7.4) than that of *capelli* with the narrow spectrum verb *pettinare* ‘comb’ (9.0). At the same time, *tirare* is more typical for *hair* than *annodare* ‘tie’ (4.6), and less typical for *hair* than *tagliare* (9.4). We conclude that each class has members that are more or less prototypical for the class and that the typicality score

<sup>15</sup> Currently, T-PAS contains 1160 analyzed verbs, 5529 patterns and ca. 200,000 annotated corpus instances. The references to the studies performed using T-PAS data are listed at <https://tpas.unipv.it/publications/>.

<sup>16</sup> The typicality score in the Sketch Engine platform is based on *logDice*, which is a metrics that indicates how strong the collocation of two words is. The higher the score, the stronger the collocation. A low score means that the words in the collocation also frequently combine with many other words.

can help define the internal organization of the classes. It has been noted that the classes in Table 3 show similarities with the *frames* (Fillmore 1982/2006) contained in the FrameNet repository (Fillmore *et al.* 2002),<sup>17</sup> although the methodologies adopted in the two resources differ. For example, the frames listed in FrameNet, although they are stipulated with the help of corpus data annotation, are not induced by manual clustering of corpus concordances based on the semantics of arguments, as are the patterns identified in T-PAS. The study of the similarities and differences between semantic patterns and cognitive frames represents an interesting direction of future research that could valuably contribute to our understanding of how semantic classes for verbs are organized and can be identified.

### 5.3. Types of arguments and compositional operations

As referenced in section 2.2, before influencing the meaning of verbs, arguments complete it, providing an essential semantic component of the linguistically encoded event, namely the participants: as a result, they are syntactically mandatory. Consider for example the case of the direct object of the verb *sbarrare* ‘to bar’ in (9), which is mandatory in all the occurrences of the verb (see the agrammaticality in (9b)):

- (9) a. *Luca ha sbarrato accuratamente tutte le porte.*  
 ‘Luca barred all doors.’  
 b. \**Luca ha sbarrato.*  
 \*‘Luca barred.’

It has been noted, however, that «sometimes an argument is obligatorily left out of the surface structure because it is subsumed as a part of the meaning of the predicate» (Fillmore 1969: 119). According to Pustejovsky (1995), this is a particular type of argument (*shadow* argument) that can be expressed only making it more specific. For example, the shadow argument *con il telefono* ‘with the phone’ can co-occur with the verb *telefonare* ‘to phone’ only if the type of phone is specified, as in (10b), and not in (10a):

- (10) a. \**Luca ha telefonato a Luisa con il telefono.*

<sup>17</sup> The FrameNet resource can be accessed at <https://framenet.icsi.berkeley.edu/>. I am thankful to the participants in the Workshop on Argument Structure held at Roma Tre University in September 2022 for pointing out this aspect.

- \*‘Luca phoned Luisa with the phone.’  
 b. *Luca ha telefonato a Luisa con il suo telefono portatile.*  
 ‘Luca phoned Luisa with his mobile phone.’

Interestingly, shadow arguments do not occur only with denominal verbs such as *telefonare* ‘to phone’ but also with verbs involving body parts, such as *camminare* ‘to walk (\*by foot)’, *vedere* ‘to see (\*with the eyes)’, and others (Ježek 2018):

- (11) a. \**Luca stava camminando a piedi.*  
 \*‘Luca was walking by foot.’  
 b. *Luca stava camminando a piedi scalzi.*  
 ‘Luca was walking barefoot.’
- (12) a. \**Luca ha visto una stella cadente con gli occhi.*  
 \*‘Luca saw a shooting star with his eyes.’  
 b. *Luca ha visto una stella cadente con i propri occhi, a occhio nudo.*  
 ‘Luca saw a shooting star with his own eyes, with the naked eye.’

It was also noted that there exist arguments, called *hidden*, that are not expressed at all in the syntax of the verb. An example is provided by the arguments of verbs of implicit creation (Bisetto & Melloni 2007; Ježek 2014), that is, verbs that denote the creation of something through the representation of a source (Ježek & Pustejovsky 2019), such as *fotografare* ‘photograph’, *copiare* ‘copy’, and *dipingere* ‘paint’. Implicit creation verbs denote the coming into being of a new entity as a result of the event itself, which, unlike explicit creation verbs such as *costruire* ‘to build’, does not appear superficially as the direct object of the verb; instead, the direct object expresses the source of the representation (*representation source theme* in Dowty’s 1979 terminology). This is the case of the direct object in (13): the created entity (the photograph) is not expressed nor expressible as an argument of the verb in the syntax.

- (13) *Luisa ha fotografato la targa.* (=created entity → photograph)  
 ‘Luisa photographed the license plate.’

The resulting entity, although not expressed, can be referred to indirectly via anaphoric expressions. In (14), for example, the clitic pronoun establishes an anaphoric relation with the unexpressed created entity, the resulting photograph:

- (14) *Luisa ha fotografato la targa e l'ha inviata a Luca.*<sup>18</sup>  
 'Luisa photographed the licence plate and sent it to Luca.'

The anaphoric test supports the hypothesis that, with these verbs, the participant that expresses the result of the action is an argument, but it is *hidden*, not expressible in the syntax of the verb.

Finally, it is well known that some arguments show optionality in the syntax, that is, they can be dropped or omitted (cf. section 4). In Ježek (2018), we argue that syntactic optionality is licensed by either *pragmatic* or *semantic* factors. A pragmatic omission occurs when an argument remains unexpressed because the referent is known to the interlocutors. In this case, the saturation of the valency of the verb takes place through the retrieval of this information from the situational context. This is the case of the example in (15), in which the dots in the brackets highlight the information that must be reconstructed for the expressions to be semantically complete and interpretable:

- (15) *Gianni iniziò (...) alle 8.30 e finì (...) alle 15.*  
 'Gianni started (...) at 8.30 and finished (...) at 3pm.'

The omissions in (15) have a pragmatic and/or textual nature (Fillmore 1986). The verbs *iniziare* and *finire* still denote a relationship between two participants, even if the second argument is not realized in the syntax, and force the listener to find something in context to saturate their meaning. If an element is not contextually assigned to the omitted argument, the sentence remains semantically incomplete.

A different phenomenon is at play when syntactic optionality is licensed by semantic factors (semantic omission). In this case, an argument may be dropped because it can be predicted by the verb semantics. Specifically, the implicit information that needs to be reconstructed to saturate the verb is the class of entities that the verb typically selects, for example the FOOD

---

<sup>18</sup> It should be noted that in a different example such as *Luisa ha fotografato il tramonto e l'ha mandato a Luca* 'Luisa photographed the sunset and send it to Luca' the clitic shows agreement with the noun *tramonto* (masculine) and not with the noun *fotografia* (feminine). This may be accounted for by assuming that the clitic refers metonymically to the information (the sunset) transferred through the photograph. In any case, the transfer cannot take place in the absence of a resulting object on which the information is recorded: this holds also for (14). See Ježek & Pustejovsky (2019) for further details.

class in the case of eating in (16a) and the DOCUMENT class in the case of reading in (16b):

- (16) a. *Luisa ha mangiato alle 18.00.*  
       ‘Luisa ate at 6pm.’  
       b. *Luca legge spesso in treno.*  
       ‘Luca often reads in the train.’

In other words, the target of the omission in (16a) and (16b) is the particular class of objects intended by the verb for that argument, and not a specific referent of the class, as in the case of the pragmatic omission in (15). Only under the condition that the omitted object is understood generically as a class and the focus is on the action being performed, the omission can be interpreted as semantic. If the target of the omission is a specific referent of the class (for example, a pizza for FOOD; a letter for DOCUMENT), then the omission should be interpreted as pragmatic and not as semantic.

#### 5.4. On the incorporation of arguments

After examining the different properties of arguments and providing a semantic account of the variability in their syntactic behaviour in 5.3, we finally propose that the relation between the verb and its arguments can be understood and represented as a relation of greater or lesser incorporation. The notion of incorporation has been used primarily in syntactic studies (Mithun & Corbett 1999). Following a suggestion that dates back to Gruber (1962), in Ježek (2018) we applied it to semantics. According to our proposal, the arguments exhibit different degrees of semantic incorporation with respect to the inherent meaning of the verb on which they depend. This is represented in (17), where the verbal *root* is meant not as the root in morphological terms but as the root in semantic terms (Levin & Rappaport 2005), coinciding with what in section 2 we have called the inherent meaning of verbs (motion, perception, cognition, etc.).

- (17) [[[[V<sub>root</sub> arg<sub>1</sub>] arg<sub>2</sub>] arg<sub>3</sub>]]

The most incorporated arguments (arg<sub>1</sub>) are the arguments that reside in the root. They can be extracted and expressed in the syntax only if they add new information to the information already introduced by the root, that is, if they are ‘informative’ (as in (10b)). Partially incorporated arguments

(arg<sub>2</sub>), on the other hand, are less built-in in the root, and are therefore more easily projected into the syntax, but they can be omitted, because they are still predictable from the verbal meaning. This is the case of arguments that can be omitted based on semantic factors (s. (16)). Finally, arguments that are not incorporated to any degree (arg<sub>3</sub>) cannot be omitted, except for pragmatic reasons (that is, unless the referent they introduce can be reconstructed pragmatically; s. (15)). According to our proposal, only arg<sub>3</sub> are outside the verbal root: this justifies their syntactic mandatoriness in all contexts (as in (9)). Whether the type of argument discussed in relation to creation verbs (cf. (13)) shall be accounted for in terms of incorporation, remains a matter of debate.

## 6. Concluding observations

After reviewing the three main components of the meaning of verbs, in this contribution we have emphasized in the first place the usefulness of syntactic approaches to the classification of verbs and their arguments at the syntax-semantic interface, especially if they take into account the syntactic alternation patterns in which each verb occurs. Secondly, we have highlighted that these approaches leave two critical issues unsolved, namely the argument-adjunct distinction and the syntactic optionality of arguments. For both we have proposed a solution based on semantic principles. Particularly, we have shown that adopting a semantic perspective as a starting point in analysing the relationship between the verb and its arguments, and assuming that, unlike adjuncts, arguments play an active role in the semantic composition with the verb, contributes to account for phenomena that cannot be explained by the traditional syntax-first approach. Moreover, we have shown that by assuming that arguments may be semantically incorporated to different degrees in the verbal meaning, one can successfully account for their syntactic optionality, when it is not due to a pragmatic factor.

## Appendix

### Acknowledgments

We would like to thank two reviewers for their insightful comments on an earlier version of the paper.

### Abbreviations

arg	argument
INTR AV	intransitive with auxiliary <i>avere</i> ‘to have’
INTR ES	intransitive with auxiliary <i>essere</i> ‘to be’
INTR PRON	intransitive with reflexive marking
TR	transitive

## References

- Bach, E. (1986). The Algebra of Events. *Linguistics and Philosophy*, 9, 5–16.
- Bisetto A. & Melloni, C. (2007). Result Nominals: A Lexical-Semantic Investigation. In: Booij, G., Ducceschi, L., Fradin, B., Guevara, E., Ralli, A. & Scalise, S. (eds) *On-line Proceedings of the Fifth Mediterranean Morphology Meeting (MMM5), September 15–18, 2005, Fréjus, France*, 393–412. (Available online: <https://geertbooij.files.wordpress.com/2014/02/mmm5-proceedings.pdf>).
- Cennamo, M. (2001). Classi verbali e cambiamento sintattico: la reinterpretazione passiva del costruito riflessivo. In: Salvi, G. (ed) *Semantica e lessicologia storiche, Atti del XXXII Congresso Internazionale della Società di Linguistica Italiana*, 225–242. Roma: Bulzoni.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge (MA): The MIT Press.
- Dowty, D. R. (1979). *Word Meaning and Montague Grammar*. Dordrecht: Reidel.
- Fellbaum, C. (ed) (1998). *Wordnet: An Electronic Lexical Database*. Cambridge (MA): The MIT Press.
- Fillmore, C. J. (1969). Types of lexical information. In: Kiefer, F. (ed) *Studies in Syntax and Semantics*, 109–137. Dordrecht/ Boston: Reidel.
- Fillmore, C. J. (1986). Pragmatically Controlled Zero Anaphora. In: V. Nikiforidou, V., VanClay, M., Niepokuj, M. & Feder, D. (eds) *Proceedings of the XII Annual Meeting of the Berkeley Linguistics Society*, 95–107. Berkeley: Berkeley Linguistic Society.
- Fillmore, C. J. (1982/2006). Frame Semantics. *Cognitive Linguistics. Basic Readings*, 34, 373–400.
- Fillmore, C. J., Baker, C. F. & Sato, H. (2002). The FrameNet Database and Software Tools. In *Proceedings of the Third International Conference on Language Resources and Evaluation. Volume IV, Las Palmas. LREC*.
- Gruber, J. S. (1965). *Studies in lexical relations*. PhD thesis, Massachusetts Institute of Technology.

- Hanks, P. (2013). *Lexical Analysis: Norms and Exploitations*. Cambridge (MA): The MIT Press.
- Jackendoff, R. (1997). *The Architecture of the Language Faculty*. Cambridge (MA): The MIT Press.
- Jakubiček, M., Kilgarriff, A., Kovář, V., Rychlý, P. & Suchomel, V. (2013). The TenTen Corpus Family. In: *7th International Corpus Linguistics Conference CL*, 125–127.
- Ježek, E. (2003). *Classi di verbi tra semantica e sintassi*. Pisa: ETS Edizioni.
- Ježek, E. (2014). *Classes of Creation Verbs*. In: Simone, R. & Masini, F. (eds) *Word Classes: Nature, Typology, Computational Representations*, 37–50. Amsterdam/Philadelphia: John Benjamins.
- Ježek, E. (2014). Esistono le collocazioni? Denotazione vs. significato collocazionale. In: Danler, P. & Konecny C. (eds) *Dall'architettura della lingua italiana all'architettura linguistica dell'Italia*, 409–419. Frankfurt a. M.: Peter Lang.
- Ježek, E. (2018). Partecipanti impliciti nella struttura argomentale dei verbi. In: Dallabrida, S. & Cordin, P. (eds.) *La grammatica delle valenze. Spunti teorici, strumenti e applicazioni*, 55–71. Firenze: Franco Cesati.
- Ježek, E. (2022). La teoria della struttura argomentale dei verbi: problemi sintattici e proposte semantiche. *Storie e linguaggi. Rivista di studi umanistici*, 8(1), 47–66.
- Ježek, E. & Pustejovsky J. (2019). Dynamic Interpretation of Predicate-Argument Structure. *Lingue e Linguaggio*, 18, 179–207.
- Kilgarriff, A., P. Rychlý, P. Smrz, & D. Tugwell (2004). The Sketch Engine. In *Proceedings of the Eleventh EURALEX International Congress. Lorient, France*.
- Levin, B. (1993). *English Verb Classes and Alternations. A Preliminary Investigation*. Chicago: The University of Chicago Press.
- Levin, B. & Rappaport Hovav, M. (2005). *Argument Realization*. Cambridge: Cambridge University Press.
- Mithun, M. & Corbett, G. (1999). *The Effect of Noun Incorporation on Argument Structure*. In: Mereu, L. (ed) *Boundaries of Morphology and Syntax*, 49–71. Amsterdam-Philadelphia: John Benjamins.
- Partee, B. (1995). Lexical semantics and compositionality. In: Osherson, D. (ed) *Invitation to Cognitive Science (2<sup>nd</sup> ed.)*. Part I: *Language (editors: Gleitman, L. & Liberman, M.)*, 311–360. Cambridge (MA): The MIT Press.
- Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge (MA): MIT Press.
- Pustejovsky, J. (2012). Co-compositionality in grammar. In: Werning, M., Hinzen, W. & Machery, E. (eds) *The Oxford Handbook of Compositionality*, 371–384. Oxford: Oxford University Press.
- Recanati, F. (2012). Compositionality, flexibility and context-dependence. In: Werning, M., Hinzen, W. & Machery, E. (eds) *The Oxford Handbook of Compositionality*, 175–191. Oxford: Oxford University Press.
- Salvi, G. & Vanelli, L. (2004). *Nuova grammatica italiana*. Bologna: Il Mulino.
- Tesnière, L. (1959) *Éléments de syntaxe structural*. Paris: Klincksieck.
- Vendler, Z. (1967). Verbs and Times. In: *Linguistics in philosophy*, 97–121. Ithaca (NY): Cornell University Press.
- Wilks, Y. (1975). A preferential, pattern-seeking semantics for natural language interference. *Artificial Intelligence*, 6, 1, 53–74.

### 3. Using event structures to model symmetric and asymmetric verb meanings in Croatian aspectual pairs

COSTANZA MARINI

Keywords: *event structure; Aktionsart; aspect; Croatian; verb polysemy*

Croatian aspectual verb pairs allow to express «the different ways of viewing the internal temporal constituency of a situation» (Comrie 1976: 3) by choosing between two morphologically related lexical variants: a perfective and an imperfective one (Babić 2002: 516–531; Barić *et al.* 1995; Browne 1993: 331–332; Kordić 1997: 41; Matovac 2022: 225–231; Silić & Pranjković 2007). However, not all verb meanings displayed by one variant are always available to the other – not even in the most derivationally prototypical of pairs. Pairs lacking one or more meanings that are present in the sense inventory of their aspectual counterpart are called *asymmetrical aspectual pairs* (Polančec 2022: 127–129) and their asymmetry is believed to be caused by constraints posed by the *Aktionsart* of the meanings involved (Kolaković 2018: 95; Polančec 2020: 160). Based on a small selection of corpus-extracted (*hrWaC* corpus, Ljubešić & Klubička 2014) and made-up examples, a comparison will be carried out between the meaning inventories of a selection of symmetric and asymmetric verb pairs available in the CroaTPAS resource (Marini & Ježek 2019; Marini 2022a). Their *Aktionsarten* will be modelled using a customized version of the formalism of event structures (Pustejovsky 1991) to shed more light on the interaction between lexical and grammatical aspect in Croatian.

*Original scientific paper*

#### 1. Introduction

The first scholar drawing a clear line between the often-misconstrued concepts of aspect and *Aktionsart* was Sigmund Agrell (1908), who stated that while aspect is what differentiates the imperfective from the perfective form in Slavic verb pairs, *Aktionsart* expresses how the events denoted by verbs unfold. Different *Aktionsarten* (or simply event types) have since then been identified and considered either as a lexical property of verbs (Dowty

1979; Vendler 1957) or a compositional one resulting from the interaction with their syntactic context (Verkuyl 1972).

In this paper, the focus will be on verb polysemy and on the fact that different verb meanings can be associated with different event types (Rappaport Hovav & Levin 1998). More precisely, a comparison will be carried out between the meaning inventories of a selection of Croatian aspectual verb pairs available in CroaTPAS (Marini & Ježek 2019; Marini 2022a), a bilingual Croatian/English pattern dictionary tailor-made to portray verbal polysemy and based on the *hrWaC* corpus (Ljubešić & Klubička 2014). The different *Aktionsarten* expressed by the verb meanings involved are going to be modelled using the formalism of event structures proposed by Pustejovsky (1991), in order to identify recurring traits in the internal structure of the modelled events to justify why a verb may lack a meaning which is available in its aspectual counterpart, as well as to model proper aspectual pairs.

## 2. A short history of aspect

The verb and its categories have always drawn the interests of scholars since ancient times. The first ones to distinguish between the three main verbal tenses (past, present, and future) were Aristotle and Plato, soon followed by the Stoics who added *aorist* and *imperfect*. Marcus Terentius Varro (116 – 27 BC) is commonly considered to be the first to identify the category of aspect (Binnick 1991: 34), which will however be properly characterized only many centuries afterwards.

The first appearance of the term *vid* (Rus. *vidy*, from the Greek *eidōs*, Eng. *aspect*) was made in the Russian grammars of the 15<sup>th</sup> and 16<sup>th</sup> century, where it was used as a synonym of “verbal type” (Gojmerac 1980: 9). The first time the term *vid* was translated with the term *aspect* was in the French translation of Greč’s Russian Grammar of 1827 by Karl Philipp Reiff. After that, the term entered several Western European languages throughout the 19<sup>th</sup> century and started to intertwine with the notion of *Aktionsart*, which was originally chosen as the German translation of *vid*, but soon began to be used to refer to both the perfective vs. imperfective distinction, as well as «the manner in which the event expressed by the verb unfolds»<sup>19</sup> – a definition provided by Brugmann (1904: 492–493).

---

<sup>19</sup> «Aktionsart ist, im Gegensatz zur Zeitstufe, die Art und Weise, wie die Handlung des Verbums vor sich geht.»

It was not until the pivotal contribution *Aspektänderung und Aktionsartbildung beim polnischen Zeitworte* (1908) by the Swedish Slavicist Sigurd Agrell that a clear line was drawn between the concept of *Aktionsart* and aspect in their current understanding:

«I do not use the label *Aktionsart* to refer to the two main Slavic verb categories, namely the unterminative and the terminative forms (the Imperfective and the Perfective) – those I call *aspects*. With the expression *Aktionsart*, I refer to the until now mostly unobserved – and even less classified – *meaning functions of affixed verbs* [...], that better express how the action is carried out, the way its development unfolds.»<sup>20</sup> (Agrell 1908: 78)

Although further distinctions should be drawn between *Aktionsarten* and the semantic functions of Slavic prefixes,<sup>21</sup> the importance of this work cannot be stressed enough. According to Gojmerac (1980: 18), Agrell is the first to define aspect as the category allowing for different ways of looking at the events expressed by verbs, which is absolutely in line with later definitions such as Comrie's (1976), who famously defines aspect as the «different ways of viewing the internal temporal constituency of a situation» (Comrie 1976: 3). Comrie also adds that aspect is the category encoding the difference between *he was reading* and *he read* (which English grammars insist on referring to as different tenses), and that «perfectivity indicates the view of a situation as a single whole, without distinction of the various separate phases that make [it] up; while the imperfective pays essential attention to the internal structure of the situation» (Comrie 1976: 16). Moreover, Comrie criticizes the idea that the basic function of the perfective should be that of representing events as momentary or punctual. On the contrary, he points out that – rather than reducing an event to a single point – «a more helpful metaphor would perhaps be to say that the perfective reduces a situation to a *blob*» (Comrie 1976: 18). Indeed, being a three-dimensional object with potential internal complexity and clearly

<sup>20</sup> «Unter *Aktionsart* verstehe ich [...] nicht die beiden Hauptkategorien des slavischen Zeitwortes, die unvollendete und die vollendete Handlungsform (das Imperfektivum und das Perfektivum) – diese nenne ich *Aspekte*. Mit dem Ausdruck *Aktionsart* bezeichne ich bisher fast gar nicht beachtete – geschweige denn klassifizierte – *Bedeutungsfunktionen der Verbalkomposita* [...], die genauer ausdrücken, wie die Handlung vollbracht wird, die Art und Weise ihrer Ausführung markieren».

<sup>21</sup> For more on this, see Gorbov (2020).

circumscribed limits makes a blob a better representation of the perfective view of a situation than a point.

Following Koschmieder's works in the 1930s, the two concepts of aspect and *Aktionsart* started not only to develop independently from one another, but also to become more and more associated with grammaticalization (the former) and lexicalisation (the latter), to the point that they sometimes go by the name of *grammatical aspect* and *lexical aspect*:

«Aspect is a fully grammaticized, obligatory, systematic category of languages, operating with general oppositions such as that of perfective and non-perfective, while *Aktionsarten* are purely lexical categories, nongrammatical, optional, and unsystematic.» (Binnick 1991: 170)

Aspectual class, aspectual character, situation type, situation aspect, action, verb class, lexical aspect, inherent aspect, and eventuality type are among the many labels *Aktionsarten* can be referred to (Polančec 2020: 41). In the rest of this contribution, *Aktionsart* will be the term used to refer to the phenomenon.

### 3. *Aktionsarten*

The first traces of a discussion surrounding *Aktionsarten* (Dowty 1979) can be found in Aristotle, who in *Metaphysics*, book 9, section 1048b distinguished between *kineseis* 'movements' and *energiai* 'actualities' to contrast actions that necessarily involve an end or result (such as that of *building*) with others which are complete in themselves (such as *being happy*).

The Oxford philosopher Gilbert Ryle (1949) later came back to Aristotle's classes and coined the term *achievement* for resultative verbs (such as *winning*, *unearthing* or *finding*) to be distinguished from irresultative activities (such as *listening* or *hunting*) that usually require a longer period of time to unfold. But it was Zeno Vendler, who put forth, in his paper *Verbs and Times* (1957), the traditional classification of verbs in four different lexical classes he calls *time schemata*, namely:

- Activity terms – *run*, *push a cart*
- Accomplishment terms – *run a mile*, *draw a circle*
- Achievement terms – *recognize*, *spot*
- State terms – *know*, *believe*

The three features Vendler bases his classification on are the following: (a) dynamicity (the presence of change), (b) duration, and (c) telicity (the presence of a goal or *télos*) (see Table 1).

Table 1. Vendler's (1957) verb classes

Verb classes	Dynamicity	Duration	Telicity	Examples
States	–	+	–	<i>know</i>
Activities	+	+	–	<i>run</i>
Accomplishments	+	+	+	<i>fix</i>
Achievements	+	–	+	<i>spot</i>

State terms, such as *know* in (1a), are not considered dynamic since they contain no elements of change. They can be seen as homogeneous situations characterized by a given duration but lacking a specific goal. Activity terms, such as *run* in (1b), on the other hand, express dynamic events lasting a certain amount of time, which still do not point at any *télos*. Accomplishment terms, such as *fix* in (1c), express events which have a duration, are dynamic and are characterized by a gradual progression towards a precise goal which must be reached for the event to have occurred. Finally, achievement terms, such as *spot* in (1d), express events without any duration culminating in an instant.

- (1) a. *I **know** the truth.*  
 b. *Claire **runs** in the park.*  
 c. *I **fixed** the thermostat.*  
 d. *I **spotted** Mary in the crowd.*

This last class contains several verbs (such as *cough*, *flash*, or *blink*) that are able to express not only single achievements, as in (2a), but also iterated instantaneous events, as in (2b). Verbs such as these are called semelfactives (from Latin *semel* 'once') or *points* (Moens & Steedman 1988; Smith 1991) and provide a prime example of *aspectual coercion*. As defined by Moens & Steedman (1988: 17), aspectual coercions consist in the «change in the aspectual type of a proposition under the influence of modifiers like tenses, temporal adverbs and auxiliaries». In (2b) for instance, it would be the combination of a punctual verb like *cough* with a durative modifier such as *for five minutes* to induce an aspectual mismatch and trigger an iterative interpretation. On the other hand, in a sentence like (2c), it is the use of the progressive that coerces once again a semelfactive into an iterative process.

- (2) a. *Blake coughed once to say yes.*  
 b. *Charlotte coughed for five minutes.*  
 c. *Sally is coughing.*

## 4. Representing events

*Aktionsarten* have been represented in many ways over the years. Here we wish to provide a short overview starting with the feature decomposition representation put forth by Van Valin & LaPolla (1997: 93), which is based on the three binary features [ $\pm$ static], [ $\pm$ telic], [ $\pm$ punctual] (s. Table 2).

Table 2. Feature decomposition of Vendler's verb classes  
(Van Valin & LaPolla 1997: 93, with modifications)

Verb class	Feature decomposition representations
State	[+static], [-telic], [-punctual]
Activity	[-static], [-telic], [-punctual]
Accomplishment	[-static], [+telic], [-punctual]
Achievement	[-static], [+telic], [+punctual]

Another way to represent events is that of decomposing them into semantic primitives (Dowty 1979; Rappaport Hovav & Levin 1998; Van Valin & LaPolla 1997). In line with Dowty (1979), Van Valin & LaPolla (1997: 102) call their decomposed lexical representations of sentences *logical structures*, which are made up of different elements: verbal constants (e.g. *do*), modifiers (BECOME, INGR) and variables. In Table 3, we can find four English examples chosen to be representative of the four Vendler classes and their respective logical structures.

Table 3. Logical structures representing Vendler's verb classes  
(Van Valin & LaPolla 1997: 102, with modifications)

Verb class	Example	Logical Structure
State	<i>Sarah is beautiful.</i>	be' (Sarah, [beautiful'])
Activity	<i>Mark is singing.</i>	do' (Mark, [sing' (Mark)])
Accomplishment	<i>Ava baked a cake.</i>	BECOME baked' (Ava, cake)
Achievement	<i>The window broke.</i>	INGR broken' (window)

Achievements require the presence of the modifier INGR, which stands for *ingressive* and encodes instantaneous changes. Accomplishments, on the other hand, need the modifier BECOME, which encodes change over some temporal span.

Similarly, Rappaport Hovav & Levin (1998: 107) call their decomposed representations of verbs *lexical semantic templates* or *event structure templates*. In their framework, predicate decomposition involves two major types of components: primitive predicates making up for the structural aspect of verb meaning and constants to take care of the idiosyncratic elements. Table 4 provides their four basic templates, one for each of Vendler's verb classes.

Table 4. Event structure templates of Vendler's classes  
(Rappaport Hovav & Levin 1998: 108, with modifications)

Verb class	Event structure template
<i>State</i>	[x <STATE>]
<i>Activity</i>	[x ACT <sub>&lt;MANNER&gt;</sub> ]
<i>Accomplishment</i>	[[x ACT <sub>&lt;MANNER&gt;</sub> ] CAUSE [BECOME [y <STATE>]]]
<i>Achievement</i>	[BECOME [y <STATE>]]]

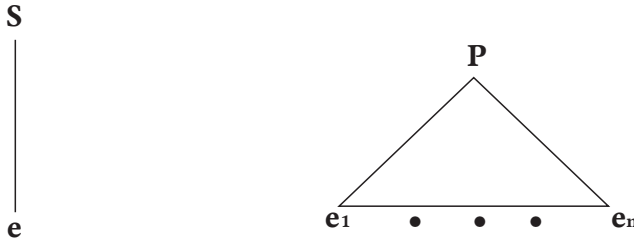
Next to feature decomposition and semantic primitives, however, several scholars decided to model Vendler's *Aktionsart* classes using graphical formalisms which focused on the internal configuration of the events being described: it is the case of Moens & Steedman's *event nucleus* (1988) and Pustejovsky's *event structure* (1991). As in the aforementioned models, also these formalisms rely on the idea that events can be broken down into smaller phases or subevents.

According to Moens & Steedman (1988), the event nucleus is constructed from a preparatory phase, a culminating event, and a consequent phase. Activities only refer to the preparatory phase, states only to the consequent phase, achievements to the culminating event and to the consequent phase and finally accomplishments to all three parts of the event nucleus.

First introduced by Pustejovsky (1991), event structures represent a well-attested formalism aimed at modelling the *Aktionsart* encoded in predicates by decomposing the events (e) they denote into smaller subevents. Following Vendler's classification, Pustejovsky divides events into *states* (S), *processes* (P), and *transitions* (T), further subdivided into *achievements*, and *accomplishments*.

From a structural point of view, *states* are the only event type consisting of a single event – static, atelic, and lasting – while *processes* (also called *activities*) consist of a dynamic series of identical and temporally ordered subevents (Figure 1).

Figure 1. Event structures for *states* (S) and *processes* (P) (Pustejovsky 1991: 40, with modifications)



There are language-specific tests that help identify the *Aktionsart* encoded in predicates. For instance, in Croatian, *states* are tested using the question *Što se događa?* ‘What is happening?’, which should elicit a negative answer (Polančec 2015: 79). On the other hand, *activities* can be tested using adverbials such as *za jedan sat* ‘in an hour’, which they do not allow, because – being atelic – they do not have an inherent completion point (Polančec 2015: 82).

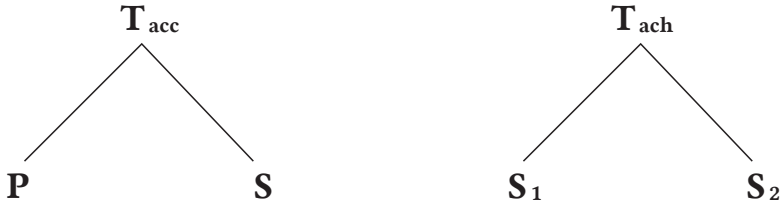
In light of this, the highlighted predicates in sentences (3a) and (3b), taken from Polančec and already tested for their *Aktionsart* (2015: 81–82), can be modelled using the event structures in Figure 1. The event of *hoping* from sentence (3a) can be modelled as a *state*, while that of *dealing with a festival* from (3b) can be modelled as a *process*. Indeed, they are both atelic and lasting, but the first one is not dynamic, i.e. it cannot be conceptualized as a series of different subevents like the second.

- (3) a. **Nadamo se pozitivnoj odluci Vijeća.**  
 ‘We hope for a positive decision from the Council.’  
 b. **Festivalima se bave samo kulturnjaci.**  
 ‘Only culture specialists deal with festivals.’

On the other hand, *transitions* require an opposition of two subsequent events (Figure 2). According to Pustejovsky (1991: 42), when it is a *process* (P)

bringing about the change, we are dealing with an *accomplishment*, whereas *achievements* make no explicit reference to the activity being performed and involve only two subsequent states ( $S_1$  and  $S_2$ ).

Figure 2. Event structures for the two types of transitions: *accomplishments* and *achievements* (Pustejovsky 1991: 40, with modifications)



In terms of tests that help distinguish between them in Croatian, Polančec (2015: 82–83) states that both *accomplishments* and *achievements* pass the test of telicity, but while accomplishments accept adverbials with a longer time span such as *za jedan sat* ‘in an hour’, achievements require shorter ones, like *u minutu* ‘in a minute’.

This is the case of sentences (4a) and (4b) – also taken from Polančec (2015: 82–83). The event denoted by the verb *naučiti* in (4a) can be modelled as an *accomplishment* since, in order to acquire knowledge, a certain amount of time must be devoted to learning the subject matter. On the other hand, the event expressed in sentence (4b) is an abrupt change of state, which can be modelled as an *achievement* as portrayed by the event structure on the right in Figure 2.

- (4) a. **Naučio je sve u jedan sat.**  
 ‘He learnt everything in an hour.’  
 b. *A sad će me ovaj led baciti nekoliko godina unatrag.*  
 ‘But now this ice will set me back several years.’

Moreover, Polančec (2015: 82–83) points out that, most of the verbs that pass the accomplishment test are perfective verbs belonging to aspectual pairs whose imperfective counterparts pass the durativity test with adverbials like *jedan sat* ‘for an hour’. More will be said on the role of aspectual pairs and their relationship with *Aktionsart* in the following section.

## 5. Croatian aspectual verb pairs

Let us now specifically focus on the Croatian verb system. Indeed, the Croatian language – like all Slavic languages – adds an element of complexity to the modelling of *Aktionsarten* if we try to account for the role played by verbal aspect (s. section 2).

Quoting Gojmerac (1980: 21), «according to Agrell [(1908)], verbal aspect is a perspective on the event, which becomes grammaticalized in Slavic languages and represents an actual grammatical category of the verb. Two are the possible ways of looking at events, each of them corresponding to one of the members of an aspectual opposition.»<sup>22</sup>

With these words, Gojmerac (1980) helps bringing together Comrie's definition of aspect and the reality of Slavic languages, including Croatian, where perfective and imperfective actions are lexicalized in two different forms, also called *aspectual pairs* (Babić 2002: 516–531; Barić *et al.* 1995; Browne 1993: 331–332; Kordić 1997: 41; Matovac 2022: 225–231; Silić & Pranjković 2007: 288).

For instance, let us look at the following examples taken from Silić & Pranjković (2007: 288):

- (5) a. *Susjed je prodavao kuću.*  
 'The neighbour was selling the house.'  
 b. *Susjed je prodao kuću.*  
 'The neighbour sold the house.'

As clearly explained by the authors, the imperfective variant *prodavati* (imperfective, 'to sell') is used in sentences like (5a) to focus on an action in which only the process is emphasized, while in sentences like (5b) featuring the perfective variant *prodati* (perfective, 'to sell'), the focus is on a completed action, on the result of a process, on an action which is not characterized by processuality.

In other words, Croatian aspectual pairs allow speakers to express «the different ways of viewing the internal temporal constituency of a situation» (Comrie 1976: 3) by choosing between different (but morphologically related) lexical variants.

<sup>22</sup> «Za Koschmieder je, kao i za Agrella, glagolski vid pogled na radnju, koji se u slavenskim jezicima gramatikalizirao i predstavlja objektivnu gramatičku kategoriju glagola. Moguća su dva načina gledanja, a svakom odgovara jedan član vidske opozicije.»

## 5.1. Asymmetric aspectual verb pairs

According to Croatian literature, the members of an aspectual verb pair are usually related to each other either via prefixation, suffixation, or suppletivity (Isačenko 1968; Silić 1978). For what concerns prefixation, prefixes are always thought to add a new meaning component to that of the base verb (Isačenko 1968: 363), whereas aspectual pairs related via other derivational strategies should show identical lexical meanings (Łazorczyk 2010: 105; Silić 1978: 46).

However, this is not always the case, as pointed out by both Cvikić & Jelaska (2007: 198) and Kolaković (2020: 122), who bring to our attention the pair *odgovarati/odgovoriti* ‘answer’. These two verbs share their most cognitively salient sense of ‘providing an answer’, but despite their not being derived from one another via prefixation, they do not share the meaning from sentence (6) ‘to suit one’s needs’, which is only characteristic of the imperfective variant *odgovarati*.

- (6) *Termin mi ne odgovara.*  
 ‘The appointment does not suit me.’

Polančec (2018) spots this phenomenon in several verb pairs, including *odgovarati/odgovoriti*. He focuses specifically on what he calls *osamostaljeni izvedeni nesvršeni glagoli* ‘orphan secondary imperfectives’ – i.e. «verbs that have the morphological structure of a derived imperfective verb, but lack their perfective partner in one or all of their meanings» (Polančec 2018: 114)<sup>23</sup> – and provides a new label to refer to their pairs, namely *asimetrični vidski parovi* ‘asymmetrical aspectual pairs’.

But why do these asymmetries take place? It is already accepted that multiple verb meanings can be associated to multiple event types (Breu 1994; Rappaport Hovav & Levin 1998), but could it be that it is in fact the *Aktionsart* encoded in these “asymmetric verb meanings” – i.e. meanings that do not appear in the sense inventory of a verb’s aspectual counterpart – that do not allow for a different viewing of their internal temporal constituency? In other words, could it be that their *Aktionsart* poses a constraint and cannot be preserved using a different aspectual variant without the verb losing its original meaning? This speculation seems to be in line with Kolaković (2018:

<sup>23</sup> «[Glagoli] koji imaju morfološku strukturu izvedenog nesvršenog glagola, ali nemaju svršeni parnjak bilo u jednom od svojih značenja ili pak u svim svojim značenjima.»

95), who believes that «the possibility of realising an aspectual opposition, namely an aspectual partner, depends also on the lexico-actional functions [of the verb meaning]»<sup>24</sup> and with Polančec (2020: 160), who points out that:

«The main reason for the importance of grammatical aspect in aspect languages lies in the fact that [...] grammatical aspect is an obligatory, paradigmatic category. This entails that, in many contexts where a verb occurs, grammatical aspect marking is present, and actionality is bound to be perceived via aspect morphology».

To provide more evidence on the intertwined relationship of *Aktionsart* and verbal aspect, an attempt will be made in this contribution at adapting the formalism of event structures (s. section 4) to fit the needs of Croatian aspectual verb pairs. Selected examples featuring both symmetric and asymmetric Croatian verb senses will therefore be analysed and their *Aktionsart* modelled using the selected formalism.

By modelling asymmetrical senses, we hope to find recurring traits in their *Aktionsarten* to justify why a verb may lack a meaning which is available in its aspectual counterpart. On the other hand, symmetrical senses will be modelled to provide a clearer graphical portrayal of the relationship between imperfective and perfective variants.

## 6. Methodology

Before modelling the *Aktionsarten* of symmetrical and asymmetrical Croatian verb meanings using the formalism of event structures, said meanings had to be identified and sentences portraying them had to be retrieved and evaluated. The following procedure was followed:

- (1) a selection of asymmetrical Croatian aspectual verb pairs was identified within the pattern inventory of the CroaTPAS resource (Marini & Ježek 2019; Marini 2022a) (section 6.1);
- (2) corpus examples portraying both the asymmetrical and symmetrical meanings of the selected verb pairs were extracted from the *hrWaC*

---

<sup>24</sup> «O leksičko-akcionalnim funkcijama ovisi i mogućnost ostvarivanja aspektne opreke, odnosno aspektnoga partnera.»

corpus (Ljubešić & Klubička 2014) and validated by native speakers (section 6.2);

- (3) made-up examples featuring the asymmetrical meanings in the syntactic context of their non-retrieved aspectual counterparts were created and evaluated by native speakers to confirm the agrammaticality of the non-retrieved options (section 6.3).<sup>25</sup>

## 6.1. CroaTPAS asymmetrical verb pairs

CroaTPAS<sup>26</sup> – i.e. *Croatian Typed Predicate Argument Structures* (Marini & Ježek 2019; Marini 2022a) – is a digital Croatian/English bilingual dictionary containing a collection of corpus-derived Croatian verb valency structures, whose argument slots have been manually annotated with a hierarchy of Semantic Type labels. Like its Italian sister project T-PAS (Ježek *et al.* 2014), CroaTPAS is primarily conceived to frame verb polysemy, since each semantically typed verb argument structure (also called *pattern*) in its inventory corresponds to a different verb sense. The resource currently contains 180 Croatian verbs for a total of 795 different verb senses.

CroaTPAS's annotation methodology is a customized version of Corpus Pattern Analysis (CPA, Hanks 2013), a lexicographic methodology relying on the idea that meaning should be mapped onto its prototypical contexts of use. CPA requires the following four steps:

- (1) 250 corpus concordance lines are randomly sampled for each verb from a reference corpus, in this case, the Croatian Web as Corpus (Ljubešić & Klubička 2014) available on the *Sketch Engine* (Kilgarriff *et al.* 2014);
- (2) the different verb senses are identified by the lexicographer looking at the concordance lines;
- (3) pattern strings are created in a pattern editing environment connected to the online corpus;
- (4) numbers are assigned to the corpus lines exemplifying each pattern, so that they can be retrieved as corpus evidence.

---

<sup>25</sup> Four native speakers between 19 and 32 years of age were involved in the evaluation, 1 man and 3 women, all of them with university degrees in Croatian studies or foreign languages.

<sup>26</sup> Available at: <https://croatpas.baisa.cz/>.

The result of this procedure are verb entries such as the one for the imperfective verb *piti* ‘to drink’ portrayed in Figure 3. In this case, six different senses were retrieved in the 250 randomly selected concordance lines, which correspond to the six patterns below. Each pattern string contains the third person singular of the verb under scrutiny together with its arguments, which are represented only by their Semantic Types, conventionally between square brackets. Semantic Types can be found between round brackets when they qualify as “default arguments”, i.e., arguments that can be linguistically omitted while still being necessary for a correct interpretation of the sentence meaning (Ježek 2018: 59). Lexical items contributing to idiomatic verb patterns can be found in curly brackets.

Figure 3. CroaTPAS patterns for the imperfective verb *piti* ‘to drink’

## **piti**

aspect: **imperfective**, related verb: **popiti**

1	[Animate] <sub>NOMINATIVE</sub> <b>pije</b> ([Liquid] <sub>ACCUSATIVE</sub> ) [Animate] drinks [Liquid]
2	[Human] <sub>NOMINATIVE</sub> <b>pije</b> ([Drug] <sub>ACCUSATIVE</sub> ) [Human] swallows, ingests [Drug]
3	[Human] <sub>NOMINATIVE</sub> <b>pije</b> [Human] drinks [Alcoholik Drink]
4	[Human = <i>Football Player</i> ] <sub>NOMINATIVE</sub> <b>pije</b> {gol ili golove} [Human = Football Player] fails to prevent a goal
5	[Human] <sub>NOMINATIVE</sub> <b>pije</b> [Human] <sub>DATIVE</sub> {krv} (na {slamku}) [Human] is a nuisance to other [Human]
6	[Vehicle] <sub>NOMINATIVE</sub> <b>pije</b> [Fuel] <sub>ACCUSATIVE</sub> [Vehicle] consumes [Fuel]

As we can see, each verb entry is linked to a related verb, in this case the verb *popiti*, which is the perfective aspectual variant of *piti*. This makes

it easy to assess whether both variants possess the same meaning inventory or whether some verb meanings are asymmetrical and pertain to only one of the two.

CroaTPAS contains 17 asymmetrical verb pairings – i.e. pairs of related entries that don't share all their attested meanings. However, not all asymmetrical verb pairings in CroaTPAS are asymmetrical aspectual verb pairs, as shown in Table 5.

Table 5. Asymmetrical verb pairings in CroaTPAS

N.	CroaTPAS verb pairings	Italian equivalent and English translation
1	<i>brojiti/izbrojiti</i>	<i>contare</i> 'to count'
2	<i>dirati/dirnuti</i>	<i>toccare</i> 'to touch'
3	<i>glasati/izglasati</i>	<i>votare</i> 'to vote'
4	<i>gostiti/ugostiti</i>	<i>ospitare</i> 'to host'
5	<i>jesti/pojesti</i>	<i>mangiare</i> 'to eat'
6	<i>odgovarati/odgovoriti</i>	<i>corrispondere</i> 'to correspond'
7	<i>raditi/uraditi</i>	<i>funzionare</i> 'to work'
8	<i>silaziti/sići</i>	<i>scendere</i> 'to descend'
9	<i>štovati/poštovati</i>	<i>rispettare</i> 'to respect'
10	<i>tumačiti/protumačiti</i>	<i>interpretare</i> 'to interpret'
11	<i>tužiti/optužiti</i>	<i>accusare</i> 'to accuse'
12	<i>uputiti/upućivati</i>	<i>avviare</i> 'to start'
13	<i>učiti/naučiti</i>	<i>imparare</i> 'to learn'
14	<i>voditi/provoditi</i>	<i>guidare</i> 'to lead'
15	<i>vući/izvući</i>	<i>trarre</i> 'to draw'
16	<i>zamisliti/zamišljati</i>	<i>immaginare</i> 'to imagine'
17	<i>zvati/pozvati</i>	<i>chiamare</i> 'to call'

This is due to the fact that the choice of Croatian verb pairings to be included in CroaTPAS was made following an Italian-driven procedure, namely selecting the Croatian translational equivalents of a sample of Italian verbs from the fundamental vocabulary of the Italian language – a group of 2,000 words covering about 90% of all Italian written and spoken text (Chiari & De Mauro 2014: 113) – taking them from the Italian/Croatian bilingual dictionary by Špikić (2017), which often offered multiple translation options.

Therefore, some of these pairings (e.g. *vući/izvući*) have been selected to evaluate the semantic impact of prefixes on base verbs rather than based on aspectual opposition.

## 6.2. GDEX corpus example extraction

Given that, in this study, our interest is restricted only to Croatian asymmetrical aspectual verb pairs not derivationally related by prefixations (s. section 5.1), the only CroaTPAS verb pairings that will be analysed are those in Table 6, whose English translations have been adapted to better fit their main attested meaning in Croatian.

Table 6. Asymmetrical non-prefixed aspectual verb pairs in CroaTPAS

N.	CroaTPAS aspectual verb pairs	English translation
1	<i>dirati/dirnuti</i>	‘to touch’
2	<i>odgovarati/odgovoriti</i>	‘to answer’
3	<i>silaziti/sići</i>	‘to descend’
4	<i>uputiti/upućivati</i>	‘to send’
5	<i>zamisliti/zamišljati</i>	‘to imagine’

In order to be able to carry out our analysis, corpus examples portraying both the asymmetrical and symmetrical meanings of these verb pairs were extracted from the *hrWaC* corpus (Ljubešić & Klubička 2014) and validated by native speakers.

Indeed, even though CroaTPAS as a whole has already been evaluated and is able to represent the polysemy of the Croatian verbs it contains (Marini 2022b), since its pattern inventory relies on the manual annotation carried out by a single annotator on a corpus containing a finite number of examples, Croatian informants have been called upon to confirm the asymmetry of the verb pairs from Table 6 and the grammaticality of the corpus examples extracted to represent them.

As a matter of fact, to provide examples of both the symmetric and asymmetric meanings of the verbs from Table 6, only corpus examples scoring at least 0.8 on the GDEX algorithm (Kilgarriff *et al.* 2008) were extracted from the *hrWaC* corpus (Ljubešić & Klubička 2014). GDEX (*Good Dictionary EXamples*) is an algorithm able to identify corpus examples that are well-fitted for lexicographic and teaching purposes. It is embedded in

the *Sketch Engine* platform and automatically assigns corpus sentences a score from 0 to 1 based on their level of lexical and syntactic complexity.

Table 7 provides an overview of all the validated GDEX examples providing information about their aspectual value (column 2), GDEX score (column 4), and whether they are symmetric (S) or asymmetric (A) (column 5).

Table 7. Selection of GDEX examples extracted from the *hrWaC* corpus

Meaning	Aspect	Example	GDEX	S/A
<i>dirati_1</i>	<i>imperf.</i>	<i>Ne diraj me, ne <b>diram</b> te.</i> 'Don't touch me, and I won't be touching you.'	0.871	S
<i>dirnuti_1</i>	<i>perf.</i>	<i>Koliko <b>ste</b> puta loptu <b>dirnuli</b> u tih dvadesetak minuta na travnjaku?</i> 'How many times have you touched the ball in those twenty minutes on the field?'	0.894	S
<i>dirati_se_5</i>	<i>imperf.</i>	<i><b>Diraju se</b> bez ikakvog rizika od trudnoće.</i> 'They masturbate without any risk of pregnancy.'	0.846	A
<i>odgovarati_1</i>	<i>imperf.</i>	<i>Ne želim tebi ovdje ništa sugerirati, samo ti <b>odgovaram</b> na pitanje.</i> 'I do not want to suggest anything here, I am just answering the question.'	0.924	S
<i>odgovoriti_1</i>	<i>perf.</i>	<i>Mene si citirao, pa <b>ću</b> ti i <b>odgovoriti</b>.</i> 'You mentioned me, so I am also going to give you an answer.'	0.925	S
<i>odgovarati_2</i>	<i>imperf.</i>	<i>Firma ne <b>odgovara</b> za neuspjele rezervacije.</i> 'The company does not answer for unsuccessful bookings.'	0.838	A
<i>odgovarati_3</i>	<i>imperf.</i>	<i>Ako ti to ne odgovara - ne <b>odgovara</b> ti ni posao u bankarskom sektoru.</i> 'If this does not suit you – then also a bank job will not suit you.'	0.925	A
<i>odgovarati_4</i>	<i>imperf.</i>	<i>Niti jedan od naših postojećih dizajna ne <b>odgovara</b> Vašim željama?</i> 'Not even one of our existing designs corresponds to your wishes?'	0.925	A
<i>odgovarati_5</i>	<i>imperf.</i>	<i>Glavni tajnik za svoj rad <b>odgovara</b> Predsjedniku.</i> 'The Secretary General answers to the President for his work.'	0.825	A
<i>odgovarati_6</i>	<i>imperf.</i>	<i>Sutra pišem biologiju i <b>odgovaram</b> filozofiju.</i> 'Tomorrow, I have to sit a biology exam and to be orally tested in philosophy.'	0.837	A

<i>silaziti_1</i>	<i>imperf.</i>	<b>Silazim</b> s busa iz Botinca. 'I am getting off of the bus from Botinec.'	0.830	S
<i>sići_1</i>	<i>perf.</i>	Obojica <b>su sišla</b> s traktora i počeli se tući. 'The two got off the tractor and started to hit each other.'	0.905	S
<i>silaziti_6</i>	<i>imperf.</i>	Smiješak joj <b>ne silazi</b> s lijepog lica. 'The smile does not get off of her face.'	0.893	A
<i>upućivati_1</i>	<i>imperf.</i>	Poseban poziv <b>upućujemo</b> bivšim članovima zbora. 'We send a special invitation to former choir members.'	0.95	S
<i>uputiti_1</i>	<i>perf.</i>	Organizator <b>će</b> svim klubovima <b>uputiti</b> pismo zahvale. 'The organizer will send thank you letters to all clubs.'	0.95	S
<i>upućivati_4</i>	<i>imperf.</i>	Taj naziv <b>upućuje</b> na japodsku tradiciju. 'That term refers to a Iapodian tradition'	0.902	A
<i>zamišljati_1</i>	<i>imperf.</i>	Naravno da ponekad u mašti <b>zamišljam</b> one grube situacije. 'Of course, sometimes in my mind I imagine awful situations.'	0.9	S
<i>zamisliti_1</i>	<i>perf.</i>	<b>Zamislite</b> svijet gdje moraš raditi samo što ti kažu. 'Imagine a world in which you have to do only what you are told.'	0.95	S
<i>zamisliti_se_2</i>	<i>perf.</i>	Čovjek <b>se zamisli</b> nad svakom sitnicom koja mu padne na pamet. 'The guy becomes pensive because of every little thing that comes to his mind'	0.95	A

As we can see, symmetric meanings available in both the perfective and imperfective variant of a given aspectual pair bear the same meaning number (*zamisliti\_1* vs. *zamišljati\_1*), which are also their pattern numbers in the CroaTPAS resource. The examples from Table 7 will be dealt with in section 7.

### 6.3. Made-up examples

On the other hand, when no corpus examples portraying a given verb sense were found, we had to make sure that that sense was indeed asymmetric, and that the absence of its aspectual counterpart was not due to insufficient corpus evidence.

Therefore, made-up examples featuring syntactic contexts that would normally trigger the use of the non-retrieved aspectual variants were created and evaluated by three native speakers to confirm the agrammaticality of the non-retrieved options and the subsequent asymmetry of the featured variant. Table 8 contains all the made-up examples that were created to confirm the asymmetric meanings from Table 7.

Table 8. Made-up examples used to confirm the asymmetric meanings from Table 7.<sup>27</sup>

N.	Example	Options	Syntactic cues
1	<i>Nakon što <b>su se</b> _____, vodili su ljubav.</i> 'After engaging in masturbation, they made love.'	<b>dirali</b> /dirnuli	<i>nakon što,</i> <i>perfective-oriented</i>
2	<i>Čim mi posao _____, izgubim interes.</i> 'As soon as a job suits me, I lose interest.'	<b>odgovara</b> /odgovori	<i>čim,</i> <i>perfective-oriented</i>
3	<i>Čim dizajn <b>bude</b> _____ Vašim željama, počet ćemo graditi zgradu.</i> 'As soon as the design corresponds to your wishes, we will start building.'	<b>odgovarao</b> /odgovorio	<i>čim,</i> <i>perfective-oriented</i>
4	<i>Prije nego što <b>je</b> Glavni tajnik za svoj rad _____ Predsjedniku, odgovarao je samo Upravnom odboru.</i> 'Before the Secretary General responded for his work to the President, he answered only to the Steering Committee.'	<b>odgovarao</b> /odgovorio	<i>prije nego što,</i> <i>perfective-oriented</i>
5	<i>Nakon što <b>sam</b> _____ filozofiju, išla sam kući.</i> 'After being orally tested in philosophy, I went back home'	<b>odgovarala</b> /odgovorila	<i>nakon što,</i> <i>perfective-oriented</i>
6	<i>Čim diktator <b>bude</b> _____ za svoje zločine, prihvatit ćemo Vaše uvjete.</i> 'As soon as the dictator answers for his crimes, we will accept your conditions.'	<b>odgovarao</b> /odgovorio	<i>čim,</i> <i>perfective-oriented</i>
7	<i>Čim joj osmijeh _____ s lica, rasplakat će se.</i> 'As soon as the smile leaves her face, she will start crying.'	<b>siđe</b> /silazi	<i>čim,</i> <i>perfective-oriented</i>

<sup>27</sup> The grammatical option is highlighted in bold.

8	<i>Prije nego što je ime na to _____, ne znamo što je značilo.</i> 'Before the meaning referred to that, we don't know what it meant.'	<b>upućivalo</b> /uputilo	<i>prije nego što,</i> <i>perfective-oriented</i>
9	<i>Danima se umjetnik _____.</i> 'The artist was pensive for days.'	zamislilo/zamišljao	<i>danima,</i> <i>imperfective-oriented</i>

As it can be seen in the Table 8, when the missing aspectual variant for a given meaning was the perfective one, a made-up example was created starting off with a subordinative conjunction that would usually require the use of a perfective verb afterwards, e.g. *čim* 'as soon as', *prije nego što* 'before', or *nakon što* 'after' (Čilaš Mikulić *et al* 2015: 59; Matovac 2022: 229). This is the case of example (2) from Table 8, which was created to ascertain the asymmetry of *odgovarati\_3* 'to suit somebody' from Table 7.

Respondents were presented with a fill-in-the-blanks task where they had to choose between the perfective or imperfective variant of the given verb, in this case *odgovara* (imperfective) or *odgovori* (perfective). The option that the respondents agreed on is highlighted in bold. The fact that, in this case, the imperfective variant *odgovara* was still necessary even in a perfective-oriented scenario gives us evidence of the asymmetry of this specific meaning.

Please, take notice that under the influence of the subordinative conjunction *čim* 'as soon as', in sentences (3) and (6), the perfective auxiliary *bude* was used, but in both cases the choice of the lexical verb still falls on the imperfective variant *odgovarati* and not on *odgovoriti*.

When the possibly asymmetrical meaning was a perfective one, an imperfective-oriented syntactic environment was created using the adverb *danima* 'for days' (Čilaš Mikulić *et al* 2015: 59), as in the case of example (9) from Table 8, which was created to assess the asymmetry of the reflexive meaning *zamisliti\_se\_2* 'to become pensive' from Table 7. Considering the fact that neither of the possible answers is highlighted in bold, in this case the respondents found the use of both the perfective and imperfective variant agrammatical, which indeed does make the meaning *zamisliti\_se\_2* 'to become pensive' an asymmetric one. More will be said on this in section 7.2.

## 7. Analysis

In this section, we will analyse the interaction between *Aktionsart* and verbal aspect in Croatian aspectual verb pairs by modelling both the symmetric (section 7.1) and asymmetric verb senses (s. section 7.2) identified in Table 7 using the formalism of event structures (section 4).

We will show how event structures are able to graphically mirror the lexical link between the perfective and imperfective variant in aspectual pairs, while at the same time allowing for a view of perfectivity in line with Comrie's *blob* metaphor (1976: 18). Moreover, event structures will all be useful in the modelling of asymmetrical meanings, allowing us to visualize why a verb may lack a meaning which is available in its aspectual counterpart.

The CroaTPAS aspectual verb pairs that are going to be at the centre of this analysis are those from Table 6: *dirati/dirnuti* 'to touch', *odgovarati/odgovoriti* 'to answer', *silaziti/sići* 'to descend', *uputiti/upućivati* 'to send', and *zamisliti/zamišljati* 'to imagine'.

### 7.1. Modelling symmetric verb meanings

The following examples are taken from Table 7 and are all cases of symmetric meanings, i.e. meanings that can be expressed using both the perfective and imperfective variant of the same aspectual pair. For instance, examples (7a) and (7b) feature the meaning 'touching a physical entity', which can be expressed perfectly by the variant *dirnuti* and imperfectively by the variant *dirati*.

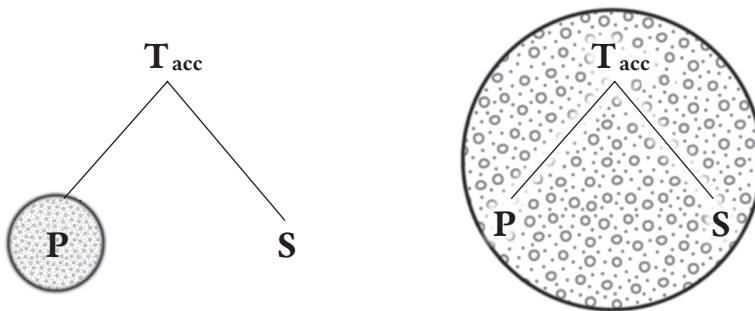
- (7) a. *Ne diraj me, ne **diram** te.*  
 'Don't touch me, and I won't touch you.'
- b. *Koliko **ste** puta loptu **dirnuli** u tih dvadesetak minuta na travnjaku?*  
 'How many times have you touched the ball in those twenty minutes on the field?'

In (7a), *diram* is actually a present tense used to express a future scope, which is why it was translated as 'I won't touch you'. The focus is not on the completed action, but rather on a future scenario when the action is being carried out, something along the lines of 'there will be no touching on my part'. On the other hand, even though in (7b) there is explicit mention of a number of times the ball has been touched, each touching event is seen

as completed, which is why the choice of aspectual variant falls on the perfective one.

In light of this, the best event structures to model the highlighted events in (7a) and (7b) are *transitions*, specifically *accomplishments* ( $T_{acc}$ ), like the ones in Figure 4.

Figure 4. Accomplishment event structures modelling symmetric meanings in Croatian aspectual pairs



The imperfective *touching* event in sentence (7a) can be seen as the *process* phase of a  $T_{acc}$  event structure (Figure 4, to the left), while the structure to the right allows us to consider the *transition* in its whole including the resulting *state* of the ball being touched, as in example (7b).

We believe that  $T_{acc}$  event structures are uniquely able to model symmetric verb senses, because the link between their branches mirrors the link between aspectual variants. As a matter of fact, when an imperfective verb is used, the focus is on the P branch, but the rest of the structure still exists – just like the possibility of using the same meaning perfectly. On the other hand, when the perfective variant is used, the focus is on the  $T_{acc}$  structure as a whole, but we can still catch a glimpse of the P phase within the overall event – just like we know that an imperfective variant still exists.

To further implement this idea, *patterned focus circles* were added to show which event phase to focus on depending on the example. We believe this graphical rendition to be particularly helpful to explain the perfective view of a situation, as well as in line with Comrie's (1976: 18) *blob* metaphor, since the internal complexity of the  $T_{acc}$  is maintained and not reduced to a single point.

The next aspectual pair that will be analysed is *odgovarati/odgovoriti*, whose most salient symmetric meaning is that of ‘providing an answer’, as testified by corpus examples (8a) and (8b):

- (8) a. *Ne želim tebi ovdje ništa sugerirati, samo ti **odgovaram** na pitanje.*  
 ‘I do not want to suggest anything here, I am just answering the question.’  
 b. *Mene si citirao, pa **ću** ti i **odgovoriti**.*  
 ‘You mentioned me, so I am also going to give you an answer.’

In much the same way as in examples (7a) and (7b), were we to assign an *Aktionsart* to the event of ‘providing an answer’, the most fitting one would be that of an *accomplishment*, since we can identify a *process* phase (the answering) bringing about a change, which is the existence of an answer (a new entity) by the end of the *transition*. In example (8a), the use of the imperfective variant *odgovarati* allows us to peek into the structure of the event focusing only on the subevent characterized by processuality, whereas the use of the perfective variant *odgovoriti* in (8b) makes us look at the event as a whole and at its being accomplished – even if in the future. Please, take notice of the choice of verb tenses used in the English translations: a progressive form to render the imperfective view, and the actual lexical presence of an end state (*the answer*) to provide the same perspective given by the perfective variant.

Also in this case, the event structure that we can use to model the events in (8a) and (8b) is a  $T_{acc}$  event structure (Figure 4), since it provides the possibility of singling out and associating its different structural elements to the two aspectual variants of the verb pair, while mirroring the existing link between them.

The focus in (8a) is on the first branch (i.e. on the process phase) because the structure models the use of the imperfective variant *odgovarati* which allows us to view the event as in progress. However, the rest of the structure does not disappear. The event of ‘answering’ still entails a resulting state, which is just not our primary concern when choosing this specific variant. Conversely, in the structure to the right, the focus is on the transition as an undivided whole, as expressed in (8b) by the perfective variant *odgovoriti*.

The same goes for examples (9a–b), (10a–b), and (11a–b):

- (9) a. ***Silazim** s busa iz Botinca.*  
 ‘I am getting off of the bus from Botinec.’  
 b. *Obojica **su sišla** s traktora i počeli se tući.*  
 ‘The two got off the tractor and started to hit each other.’

- (10) a. *Poseban poziv **upućujemo** bivšim članovima zbora.*  
 ‘We are sending a special invitation to former choir members.’  
 b. *Organizator će svim klubovima **uputiti** pismo zahvale.*  
 ‘The organizer will send thank you letters to all clubs.’
- (11) a. *Naravno da ponekad u mašti **zamišljam** one grube situacije.*  
 ‘Of course, sometimes in my mind I imagine awful situations.’  
 b. ***Zamislite** svijet gdje moraš raditi samo što ti kažu.*  
 ‘Imagine a world in which you have to do only what you are told.’

In (9a), (10a), and (11a), we are presented with imperfective views of the events of ‘getting off of a bus’, ‘sending an invitation’, and ‘imagining situations’. However, all of these meanings are actually symmetric, and can also be expressed using the perfective variants of their respective aspectual pair, as in (9b), (10b), and (11b).

All the events expressed by the imperfective variants can be associated with actions that are either portrayed as taking place at the moment of speaking, as in (9a) and (10a), or as habitual but still durative, as in (11a). Therefore, all these events can be seen, and subsequently modelled, as the *process* phases of a  $T_{acc}$  event structure (Figure 4, to the left).

On the other hand, all the events expressed by the perfective variants are presented as a whole – despite the change in verbal tenses: in (9b) we have a past tense (*su sišla*, ‘they got off’), in (10b) a future (*će uputiti*, ‘he will send’), and in (11b) an imperative (*zamislite*, ‘imagine’). In all these cases, the events are considered globally and can thus be modelled as  $T_{acc}$  event structures, where the focus circles surrounding the structures help visualising them as a blob.

Before turning to the asymmetric meanings from Table 7, let us consider the fact that all the symmetric meanings we found in our inventory could be modelled as the different phases of an *accomplishment* event structure ( $T_{acc}$ ). This supports the data from Polančec (2015: 83), who states that most of the Croatian verbs that pass the *accomplishment* test are perfective verbs belonging to aspectual pairs whose imperfective counterparts pass the durativity test.

## 7.2. Modelling asymmetric verb meanings

In the following section, we will encounter only asymmetric verb meanings – i.e. meanings that are absent from the inventory of one of the aspectual variants making up a verb pair. To make sure they were indeed asymmetric, made-up examples with syntactic contexts that require the use of the non-

retrieved aspectual variant were created and validated by native speakers (s. Table 8).

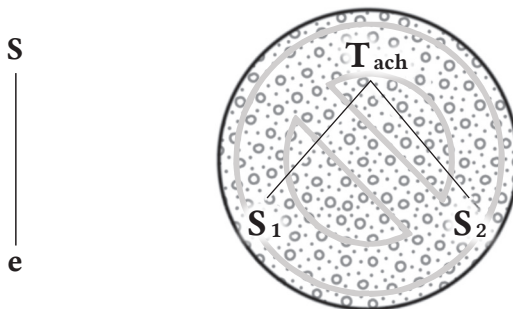
The *Aktionsart* of these validated asymmetrical meanings will now be modelled using the formalism of event structures to provide a graphical representation that may help us better understand their asymmetry.

The first meaning we will focus on belongs to the imperfective verb *odgovarati* and is amiss in its perfective counterpart *odgovoriti*. It is the meaning of ‘suing somebody’ that we find in the corpus-extracted sentence (12a):

- (12) a. *Ako ti to ne odgovara – ne **odgovara** ti ni posao u bankarskom sektoru.*  
 ‘If this does not suit you – then also a bank job will not [suit you].’  
 b. *\*Čim mi nešto **odgovori**, izgubim interes.*  
 ‘As soon as something suits me, I lose interest.’  
 c. *Čim mi nešto **odgovara**, izgubim interes.*  
 ‘As soon as something suits me, I lose interest.’

As shown in (12b), the perfective variant *odgovoriti* cannot be used to express this meaning despite the presence of the conjunction *čim* (‘as soon as’), which typically favours a perfective view on events. Our respondents opted for the use of the variant *odgovarati* both by itself (12c) and introduced by the phase verb *početi* (‘to start’) as in *Čim mi nešto **počne odgovarati**, izgubim interes* ‘As soon as something starts to suit me, I lose interest’. We found this last solution particularly interesting since it effectively solves the problem of being presented a perfective-oriented environment while being forced to use an imperfective form.<sup>28</sup>

Figure 5. Event structures modelling asymmetric state meanings



<sup>28</sup> See Batinić Angster (*this volume*, Chapter 4, section 4.3.5) for a discussion of the temporal uniformity of infinitives embedded by phasal or aspectual verbs, among others.

In terms of *Aktionsart*, both the events from (12a) and (12c) can be categorised as *states*, since in both cases we deal with events that are homogeneous, atelic, and lasting (Kolaković 2018: 95). Indeed, in (12c) we can identify a beginning for this state that is caused by the choice of connector ('as soon as'), but this does not change the fact that the event of 'being suitable' expressed by *odgovarati* is just one single *state* and not a *transition* between *state*<sub>1</sub> and *state*<sub>2</sub>, thus making it impossible to model this meaning using the event structure of an achievement ( $T_{ach}$ ) like the one on the right of Figure 5.

In other words, the lack of telicity in the meaning makes it impossible for the variant *odgovoriti* to be used even in a perfective-oriented context such as that of (12b). We can therefore state that the *Aktionsart* of this meaning poses a constraint to the creation of a perfective view of the event, which makes the meaning asymmetric.

A similar situation is to be found in examples (13a–c), still relating to the pair *odgovarati/odgovoriti*. The asymmetric sense this time is that of 'corresponding to something', which can be expressed only with the imperfective variant *odgovarati*.

- (13) a. *Niti jedan od naših postojećih dizajna ne **odgovara** Vašim željama?*  
 'Not even one of our existing designs corresponds to your wishes?'  
 b. \**Čim dizajn **bude odgovorio** Vašim željama, počet ćemo graditi zgradu.*  
 'As soon as the design corresponds to your wishes, we will start building.'  
 c. *Čim dizajn **bude odgovarao** Vašim željama, počet ćemo graditi zgradu.*  
 'As soon as the design corresponds to your wishes, we will start building.'

Despite the use of the connector *čim* 'as soon as' trying to elicit the use of the perfective variant *odgovoriti* (13b), the meaning of 'corresponding to something' can only be expressed by *odgovarati*, (13c). Only the perfective auxiliary *bude* aligns with the conjunction, but not the verb on a lexical level. Also in this case, the *Aktionsart* of the involved verb meaning is that of *state*, since we are dealing with a homogeneous, atelic event characterized by a certain duration (Kolaković 2018: 95). Even though we can hope there will be a moment in which the design will start to meet the client's expectations and finally correspond to what he or she wants, the verb does not express a *transition*, but simply a *future state*.

Once again, the structures in Figure 5 can model the situation: the state of ‘corresponding to one’s wishes’ cannot be reduced to the starting state ( $S_1$ ) of an achievement event structure ( $T_{ach}$ ) to be viewed globally using a perfective variant because there is no telic element in the meaning to begin with. Also in this case, to put it in Kolaković’s words, it is the actional properties of the verb meaning that are preventing the realisation of the aspectual partner *odgovoriti* (Kolaković 2018: 95).

The same goes for examples (14a–c), (15a–c), as well as (16a–c). The latter triple features the aspectual pair *upućivati/uputiti*, where the meaning of ‘referring to’ results typical only of the imperfective variant *upućivati* and goes amiss in the perfective one *uputiti* because of the constraints posed by the fact that the meaning expresses an atelic *state*. To support this, we can use Polančec’s test and ask the question *Što se događa?* ‘What is happening?’ (2015: 79), which does elicit a negative answer, since nothing is happening if something is simply ‘referring to a tradition’ (16a).

- (14) a. *Glavni tajnik za svoj rad **odgovara** Predsjedniku.*  
 ‘The Secretary General answers to the President for his work.’  
 b. \**Prije nego što **je** Glavni tajnik za svoj rad **odgovorio** Predsjedniku, odgovarao je samo Upravnom odboru.*  
 ‘Before the Secretary General answered to the President for his work, he answered only to the Steering Committee.’  
 c. *Prije nego što **je** Glavni tajnik za svoj rad **odgovarao** Predsjedniku, odgovarao je samo Upravnom odboru.*  
 ‘Before the Secretary General answered to the President for his work, he answered only to the Steering Committee.’
- (15) a. *Firma **ne odgovara** za neuspjele rezervacije.*  
 ‘The company does not answer for unsuccessful bookings.’  
 b. \**Čim diktator **bude odgovorio** za svoje zločine, prihvatit ćemo Vaše uvjete.*  
 ‘As soon as the dictator answers for his crimes, we will accept your conditions.’  
 c. *Čim diktator **bude odgovarao** za svoje zločine, prihvatit ćemo Vaše uvjete.*  
 ‘As soon as the dictator answers for his crimes, we will accept your conditions.’

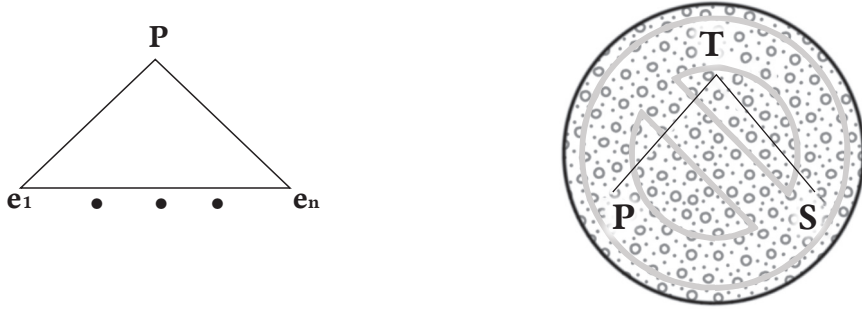
- (16) a. *Taj naziv **upućuje** na japodsku tradiciju.*  
 ‘That term refers to a tradition of Japode.’  
 b. \**Prije nego što **je** ime na to **uputilo**, ne znamo što je značilo.*  
 ‘Before the meaning referred to that, we don’t know what it meant.’  
 c. *Prije nego što **je** ime na to **upućivalo**, ne znamo što je značilo.*  
 ‘Before the meaning referred to that, we don’t know what it meant.’  
 On the other hand, something quite different happens in examples (17a–c):
- (17) a. *Sutra pišem biologiju i **odgovaram** filozofiju.*  
 ‘Tomorrow, I have to sit a biology exam and to be orally tested in philosophy.’  
 b. \**Nakon što **sam odgovorila** filozofiju, išla sam kući.*  
 ‘After being orally tested in philosophy, I went back home.’  
 c. *Nakon što **sam odgovarala** filozofiju, išla sam kući.*  
 ‘After being orally tested in philosophy, I went back home.’

In (17b), the meaning that goes amiss in the perfective variant *odgovoriti* is that of ‘being orally tested, undergoing an oral test’, which is present in the sense inventory of the imperfective variant *odgovarati*. What justifies this asymmetry is the fact that a perfective view on the event would somehow undermine the main actional feature of the meaning itself, which is that of processuality. ‘Being tested’ is an event which has a certain duration and dynamicity (not all its moments are the same), but it lacks a goal, it is not a telic event because the person who is being tested has no control over the end of the activity. Indeed, we can even test whether this meaning qualifies as a *process* (or *activity*) using a durativity test:

- (18) a. *Odgovarala sam filozofiju **sat vremena**.*  
 ‘I have been orally tested in philosophy for an hour.’

Considering all this, we find ourselves in a different predicament than in the previous examples. This time the actional constraint which does not allow for the use of the perfective variant *odgovoriti* is not the *state* character of the event, but its being a *process* which cannot be reduced to the first phase of an *accomplishment* (as portrayed by the agrammatical event structure to the right of Figure 6).

Figure 6. Event structures modelling asymmetric process meanings



The same goes for the meaning exemplified in (19a) and (19c), namely that of ‘engaging in masturbation’. Also in this case, we find ourselves facing an asymmetric meaning (a reflexive one this time), which cannot be expressed by the perfective variant *dirnuti se*, but only by the imperfective one *dirati se*.

- (19) a. ***Diraju se bez ikakvog rizika od trudnoće.***  
 ‘They masturbate without any risk of pregnancy.’  
 b. *Nakon što su se dirnuli, vodili su ljubav.*  
 ‘After having touched each other, they made love.’  
 c. ***Nakon što su se dirali, vodili su ljubav.***  
 ‘After engaging in masturbation, they made love.’

As a matter of fact, if we used *dirnuti se* as in (19b), the sentence would not be ungrammatical, but the sexual meaning would be lost and the new sentence would read something like ‘After having touched each other (with their hands), they made love’. This is because this latter meaning of the pair *dirati/dirnuti* is a symmetrical one whose internal structure can be modelled like an accomplishment. Conversely, the meaning in (19c) cannot be expressed with the perfective variant *dirnuti*, because – even though it does express a completed action – it needs to underline the durativity of said activity.

In terms of event structures, this translates to saying that the best structure to model (19c) is the same as (19a), i.e. that of an autonomous *process* which cannot be traced back to the first phase of an *accomplishment* event like the one to the right of Figure 6.

Let us now turn to the asymmetric reflexive meaning *zamisliti se* ‘becoming pensive’, which lacks an imperfective counterpart. This prompted

us to create an imperfective-oriented context using the adverb *danima* ‘for days’, to try to elicit the use of an imperfective form (Čilaš Mikulić *et al* 2015: 59), as can be seen in sentences (20b–c).

- (20) a. *Čovjek se zamisli nad svakom sitnicom koja mu padne na pamet.*  
 ‘The guy becomes pensive over every little thing that comes to his mind.’  
 b. *\*Danima se umjetnik zamišljao.*  
 ‘The artist was pensive for days.’  
 c. *\*Danima se umjetnik zamislio.*  
 ‘The artist was pensive for days.’  
 d. *Danima je umjetnik bio zamišljen.*  
 ‘The artist was pensive for days.’

‘Becoming pensive’ is a classic example of a verb expressing a change of state (an *achievement* in Vendler’s (1957) terminology), i.e. an event involving two subsequent states ( $S_1$  and  $S_2$ ): the first being a non-pensive state of mind and the second a pensive one (Figure 7, to the left).

Figure 7. Event structures modelling asymmetric achievement meanings



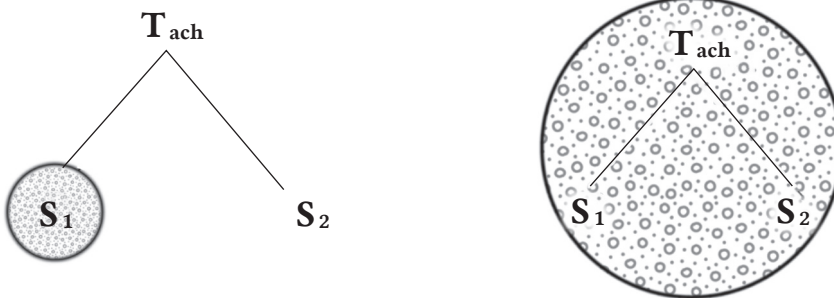
According to our respondents, neither the imperfective variant *zamišljati* (*se*) nor the perfective one *zamisliti* (*se*) can express the beginning state ( $S_1$ ) nor the end state ( $S_2$ ), as can be seen from (20b–c). However, other verbal expressions can be used, such as *biti zamišljen* ‘being pensive’, as can be seen from example (20d) *Danima je umjetnik bio zamišljen* ‘The artist was pensive for days.’ The event structure to the right of Figure 7 tries to model this situation, portraying the impossibility to peer into the different phases of the transition using the same verbal expression used to refer to the transition itself.

Finally, the last meaning that we are going to consider is the one from sentence (21a), namely the idiomatic meaning *smiješak nekome ne silazi s lica* ‘somebody cannot get a smile off of their face’. In this case, respondents did not find this meaning to be asymmetric, since it was enough to leave out the negation from the perfective-oriented scenario to allow the meaning to be used perfectly with the variant *sići*, as can be seen from (21c).

- (21) a. *Smiješak joj ne silazi s lijepog lica.*  
 ‘The smile does not get off of her face.’  
 b. \**Čim joj osmijeh silazi s lica, rasplakat će se.*  
 ‘As soon as the smile leaves her face, she will start crying.’  
 c. *Čim joj osmijeh siđe s lica, rasplakat će se.*  
 ‘As soon as the smile leaves her face, she will start crying.’

In terms of *Aktionsarten*, the meaning in (21a) can be modelled as a state, because nothing is happening if a smile does not get off of somebody’s face. On the other hand, example (21c) expresses a transition, a change of state, from an  $S_1$  (smile is there) to an  $S_2$  (smile has left). The two meanings can therefore be modelled as the beginning state of an achievement (Figure 8, to the left) and as the achievement itself, considered in its entirety, as highlighted by the focus circle around the event structure on the right of Figure 8.

Figure 8. Event structures modelling asymmetric achievement meanings



What can be considered asymmetric is the idiomaticity of the imperfective meaning in (21a), since it is considerably more commonly used than the perfective one from (21c), as testified by corpus data.

## 8. Conclusions

The aim of this contribution was to shed light on the intricate relationship between aspect and *Aktionsart* in Croatian aspectual verb pairs, first by providing a comprehensive historical overview of these concepts (s. section 2 and 3), and secondly by modelling the *Aktionsart* of a selection of Croatian verb meanings using a customized version of the formalism of event structures (section 4).

Special attention was devoted to aspectual pairs lacking one or more meanings that are present in the sense inventory of their aspectual counterpart. These are called *asymmetrical aspectual pairs* (Polančec 2022: 127–29) and their asymmetry is believed to be caused by constraints posed by the *Aktionsart* of the meanings involved (Kolaković 2018: 95; Polančec 2020: 160).

The following methodology was followed: (1) asymmetrical Croatian aspectual pairs from the CroaTPAS resource (Marini & Ježek 2019; Marini 2022a) were identified (section 6.1); (2) corpus examples portraying both the asymmetrical and symmetrical meanings were extracted from the *hrWaC* corpus (Ljubešić & Klubička 2014) (section 6.2); (3) made-up examples featuring the asymmetrical meanings in the syntactic context of their non-retrieved aspectual partners were created and evaluated by native speakers to confirm the asymmetry (section 6.3).

Both the corpus-extracted (Table 7) and made-up examples (Table 8) were analysed and modelled using event structures, which we believe have proven particularly effective on several levels. As far as the symmetrical meanings go, event structures were able to provide a graphical portrayal of the relationship between perfective and imperfective variants (section 7.1), while at the same time allowing for a view of perfectivity in line with Comrie's (1976: 18) *blob* metaphor. Moreover, event structures have also proven useful in the modelling of asymmetrical meanings (section 7.2), allowing us to visualize the actional reasons behind the asymmetries. According to our results, all *Aktionsarten* (*states*, *processes*, and *transitions*) seem to have the potential to pose certain constraints on the aspectual viewing of the events they encode and to prevent verbs from displaying meanings that are available to their aspectual counterparts.

## Appendix

### Acknowledgments

I would like to thank two reviewers for their comments on an earlier version of the paper.

### Abbreviations

ACT	activity
e	event
INGR	ingressive
P	process
S	state
T	transition
T <sub>acc</sub>	accomplishment
T <sub>ach</sub>	achievement

## References

- Agrell, S. (1908). *Aspektänderung Und Aktionsartbildung Beim Polnischen Zeitworte*. Lund: Håkan Ohlssons Buchdruckerei.
- Babić, S. (2002). *Tvorba riječi u hrvatskome književnome jeziku*. Zagreb: Hrvatska akademija znanosti i umjetnosti/Globus nakladni zavod.
- Barić, E., Lončarić, M., Malić, D., Pavešić, S., Peti, M., Zečević, V. & Znika, M. (1995). *Hrvatska gramatika*. Zagreb: Školska knjiga.
- Binnick, R. I. (1991). *Time and the Verb: A Guide to Tense and Aspect*. New York/Oxford: Oxford University Press.
- Browne, W. (1993). Serbo-Croat. In: Comrie, B. & Corbett G. G. (eds) *The Slavonic Languages*, 306–387. London: Routledge.
- Brugmann, K. (1904). *Kurze Vergleichende Grammatik Der Indogermanischen Sprachen*. Strassbourg: Trübner.
- Chiari, I. & De Mauro, T. (2014). The New Basic Vocabulary of Italian as a Linguistic Resource. In: Basili, R. & Lenci, A. (eds) *Proceedings of the 1st Italian Conference on Computational Linguistics*, 113–116. Pisa: Pisa University Press.
- Čilaš Mikulić, M., Gulešić Machata, M. & Udier, S. (2015). *Razgovarajte s nama! Udžbenik hrvatskoga kao drugoga i stranoga jezika za stupanj B2*. Zagreb: FF Press.
- Comrie, B. (1976). *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.
- Cvikić, L. & Jelaska, Z. (2007). Složenost ovladavanja glagolskim vidom u inojezičnome hrvatskome. *Lahor: časopis za hrvatski kao materinski, drugi i strani jezik*, 2(4), 190–216.

- Dowty, D. R. (1979). *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. Dordrecht: D. Reidel Publishing Company.
- Gojmerac, M. (1980). *Glagolski vid u hrvatskome ili srpskom i njemačkom jeziku*. PhD thesis, University of Zagreb.
- Gorbov, A. (2020). Aktionsart. *Encyclopedia of Slavic Languages and Linguistics Online*.
- Hanks, P. (2013). *Lexical Analysis: Norms and Exploitations*. Cambridge (MA): MIT Press.
- Isačenko, A. V. (1968). *Die Russische Sprache Der Gegenwart. Formenlehre*. Niemeyer: Halle.
- Ježek, E. (2018). Partecipanti impliciti nella struttura argomentale dei verbi. In: Dallabrida, S. & Cordin, P. (eds.) *La grammatica delle valenze. Spunti teorici, strumenti e applicazioni*, 55–71. Firenze: Franco Cesati.
- Ježek, E., Bernardo, M., Feltracco, A., Bianchini, A. & Popescu O. (2014). T-PAS: A Resource of Corpus-Derived Typed Predicate-Argument Structures for Linguistic Analysis and Semantic Processing. In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, 890–895.
- Kilgarrieff, A., Vitek, B., Bušta, J., Jakubiček, M., Kovář, V., Michelfeit, J., Rychlý, P. & Suchomel, V. (2014). The Sketch Engine: Ten Years On. *Lexicography*, 1, 7–36.
- Kilgarrieff, A., Husák, M., McAdam, K., Rundell, M. & Rychlý, P. (2008). GDEX: Automatically Finding Good Dictionary Examples in a Corpus. In: *Proceedings of the XIII EURALEX International Congress*, 425–432.
- Kolaković, Z. (2018). *Dvoaspektni glagoli – Razlike između (p)opisa u priručnicima i stanja u korpusu s posebnim osvrtom na uporabu izvornih govornika*. PhD thesis, University of Zagreb/University of Regensburg.
- Kolaković, Z. (2020). Par – Nepar – Aspektni Par\*. *Jezikoslovlje*, 21(2), 103–147.
- Kordić, S. (1997). *Serbo-Croatian*. München: Lincom Europa.
- Łazorzcyk, A. A. (2010). *Decomposing Slavic Aspect: The Role of Aspectual Morphology in Polish and Other Slavic Languages*. PhD thesis, University of Southern California.
- Ljubešić, N. & Klubička, F. (2014). {bs, hr, sr} WaC – Web Corpora of Bosnian, Croatian and Serbian. In: *Proceedings of the 9th Web as Corpus Workshop (WaC-9)*, 29–35.
- Marini, C. (2022a). CroaTPAS: uspoređivanje značenja vidskih parnjaka s namjerom istraživanja odnosa između vida, akionsarta i glagolske polisemije u hrvatskom. *Croatica: časopis za hrvatski jezik, književnost i kulturu*, 46(66), 67–80.
- Marini, C. (2022b). CroaTPAS: A Survey-Based Evaluation. In: *Proceedings of the 18th Joint ACL - ISO Workshop on Interoperable Semantic Annotation within LREC2022*, 76–80.
- Marini, C. & Ježek, E. (2019). CROATPAS: A Resource of Corpus-Derived Typed Predicate Argument Structures for Croatian. In: Bernardi, R., Navigli, R. & Semeraro, G. (eds) *Proceedings of the 6th Italian Conference on Computational Linguistics (CLiC-It)*.
- Matovac, D. (2022). *Basic Croatian Grammar: For Croatian Language Learners*. Zagreb: Hrvatska sveučilišna naklada.
- Moens, M. & Steedman, M. (1988). Temporal Ontology and Temporal Reference. *Computational Linguistics*, 14(2), 15–28.
- Polančec, J. (2015). Što nam Aktionsart glagola može reći o njegovu vidu. In: Udier, S. & Cergol Kovačević, K. (eds) *Višejezičnost kao predmet multidisciplinarnih istraživanja: zbornik radova s međunarodnoga znanstvenog skupa Hrvatskog društva za primijenjenu lingvistiku održanoga od 25. do 27. travnja 2014. godine u Zagrebu*, 75–88. Zagreb: Srednja Europa/Hrvatsko društvo za primijenjenu lingvistiku (HDPL).

- Polančec, J. (2018). Osamostaljeni izvedeni nesvršeni glagoli u hrvatskom jeziku. *Suvremena lingvistika*, 44(85), 113–138.
- Polančec, J. (2020). *A Typology of Aspect-Actionality Interactions*. PhD thesis, University of Zadar.
- Polančec, J. (2022). Vidski neupareni glagoli u hrvatskom jeziku. *Croatica: Časopis za hrvatski jezik, književnost i kulturu*, 46(66), 105–133.
- Pustejovsky, J. (1991). The Syntax of Event Structure. *Cognition*, 41, 47–81.
- Rappaport Hovav, M. & Levin, B. (1998). Building Verb Meanings. In: Butt, M. & Geuder, W. (1998). *The Projection of Arguments: Lexical and Compositional Factors*, 97–134. Stanford: CSLI Publications.
- Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson's University Library.
- Silić, J. (1978). An Approach to the Study of Aspectuality in the Croatian Literary Language. In: Filipović, R. (ed) *Contrastive Analysis of English and Serbo-Croatian*, 42–70. Zagreb: Zavod za lingvistiku.
- Silić, J. & Pranjković, I. (2007). *Gramatika hrvatskoga jezika: za gimnazije i visoka učilišta*. 2<sup>nd</sup> ed. Zagreb: Školska knjiga.
- Smith, C. (1991). *The Parameter of Aspect*. Dordrecht: Springer.
- Špikić, A. (2017). *Croato Compatto: Dizionario Croato/Italiano e Italiano/Croato*. Bologna: Zanichelli.
- Van Valin, R. D. & LaPolla, R. J. (1997). *Syntax: Structure, Meaning, and Function*. Cambridge: Cambridge University Press.
- Vendler, Z. (1957). Verbs and Times. *The Philosophical Review*, 66(2), 143–160.
- Verkuyl, H. J. (1972). *On the Compositional Nature of the Aspects*. Dordrecht: D. Reidel Publishing Company.

## 4. Monoclausal configurations with infinitives in Croatian

MIA BATINIĆ ANGSTER

Keywords: *infinitives; restructuring; control; verbal classes; predicates; complementation*

This paper investigates monoclausal properties of configurations with infinitives in Croatian. These configurations are contrasted with biclausal configurations with infinitives embedded by non-restructuring Croatian predicates. Croatian predicates embedding infinitives are first classified into the verbal classes proposed by Landau (2000). Secondly, following Grano's (2015) account of restructuring inspired mostly by Cinque's cartographic approach (2004), three classes of Croatian predicates that participate in monoclausal configurations are identified: modals, aspectuals, and implicatives. These express meanings below Tense in Cinque's hierarchy (1999, 2004, 2006) and, unlike other subject-oriented predicates that can embed infinitives (e.g. desideratives, factives), give rise to monoclausal configurations displaying properties such as preserved synonymy under passivization, matrix licensing of negative polarity items, unavailability of finite complementation, and temporal uniformity.

*Original scientific paper*

### 1. Introduction

The aim of this paper is to investigate monoclausal properties of configurations with infinitives in Croatian. Infinitival complements have generally been studied as complements that participate in biclausal configurations with covert subject in a complement clause. The unexpressed subject (PRO) can depend on the subject of the matrix clause or can be identified as the object of the matrix predicate. These cases are called subject- and object-control respectively, whereby control is intended as the relation between PRO and its antecedent (cf. Landau 2000: 1). In the previous case PRO is co-referential

with the subject of the matrix clause, while in the latter, called object-control, PRO is co-referential with the object.<sup>29</sup>

However, it has been noted that some  $V_{\text{finite}} + V_{\text{non-finite}}$  configurations form a kind of complex predicate in which the infinitive does not have the status of clausal complement. This phenomenon is known under the name of *clause union* (Aissen & Perlmutter 1976; Roussou 2009) or *restructuring* (Rizzi 1976; Cinque 2004, 2006; Wurmbrand 2003; Grano 2015, 2017). An array of different properties can point to monoclausality of these configurations, e.g. clitic climbing (s. Aissen & Perlmutter 1976; Rizzi 1978; Cardinaletti & Shlonsky 2004; Cinque 2004, 2006), long passivization (s. Wurmbrand 2003), exhaustive control, ban on overt embedded subjects, unavailability of finite complementation, “long-distance” licensing of negative indefinites, inverse scope (s. Grano 2015), etc.

Until recently, a restricted number of studies have been devoted to control and restructuring predicates in Croatian.<sup>30</sup> The aim of this paper is to attempt to fill this gap discussing several arguments for a monoclausal approach to certain configurations with infinitives in Croatian.

I expect that some classes of Croatian verbs participate in monoclausal configurations, therefore I aim at identifying and distinguishing them from verbs lacking monoclausal properties, i.e. verbs that can only involve clausal infinitival complements. Based on previous accounts on monoclausal configurations with infinitives, I will apply a set of tests as diagnostics for discerning between those two groups of verbal classes. I will observe that Croatian modals (e.g. *moći* ‘can’, *morati* ‘must’, deontic *imati* ‘have to’), aspectuals (e.g. *početi* ‘start’, *završiti* ‘end’, *nastaviti* ‘continue’) and implicatives (e.g. *uspjeti* ‘manage’, *izbjeći* ‘avoid’) – unlike other subject-oriented predicates<sup>31</sup> that embed infinitives, subject-control predicates (e.g. *odlučiti* ‘decide’, *obećati* ‘promise’, *ponuditi se* ‘offer’, *očekivati* ‘expect’) – give rise to monoclausal configurations displaying properties such as preserved

<sup>29</sup> Object control verbs in Croatian prefer finite complements introduced by *da* ‘that’, therefore object control constructions are rather rare (Batinić Angster 2019: 283). Object control verbs taking an infinitive are for example the Spanish “manipulation” verbs taken into consideration by Mikelenić & Bezlaj (*this volume*, Chapter 8).

<sup>30</sup> See below section 2.

<sup>31</sup> As a neutral label grouping together subject-control verbs and restructuring but subject-control-like verbs I use “subject-oriented predicates”. This allows me to exclude object-control and object-control-like verbs (s. Batinić Angster 2019).

synonymy under passivization, matrix licensing of negative polarity items (NPI), unavailability of finite complementation and of tense-mismatches.

## 2. Monoclausal $V_{\text{finite}} + V_{\text{non-finite}}$ configurations

One of the most often cited properties pointing to the monoclausality of apparently biclausal  $V_{\text{finite}} + V_{\text{non-finite}}$  configurations is the availability of clitic climbing (also known as scrambling) – a phenomenon in which «a clitic associated with an embedded verb is actually pronounced on the higher verb» (Krapova & Cinque 2018: 151). It has been observed that in some Romance languages object clitics can climb out of the non-finite clause only in some environments. In Italian for example object clitics are placed to the left of a finite verb or to the right of a non-finite verb and they can climb from an embedded non-finite verb to the matrix verb only in case it is a modal, aspectual, or motion verb, as noticed by Rizzi (1976: 4):

- (1) a. *Gianni deve/può/comincia a/viene a leggerlo.*  
 b. *Gianni lo deve/può/comincia a/viene a leggere.*  
 ‘Gianni must/can/begins to/comes to read it.’

While (2a) and (2b) are both well-formed, just like the (3a), the same does not hold in the sentence with verb *detestare* ‘detest’ (3b) as it does not permit the clitic climbing (Grano 2015: 11-12).

- (2) a. *Gianni cominciava a vederlo.*  
 b. *Gianni lo cominciava a vedere.*  
 ‘Gianni was beginning it to see it.’  
 (3) a. *Gianni detestava vederlo.*  
 b. \**Gianni lo detestava vedere.*  
 ‘Gianni hated seeing it.’

At the same time, in Italian clitic climbing is restricted to non-finite complements as it is unavailable in the environment with a finite complement. In German, for example, the phenomenon known as long passivization is available only with some verbs, such as *versuchen* ‘to try’, but not with *planen* ‘to plan’ (Wurmbrand 2003: 19, 36):

- (4) *dass der Traktor zu reparieren versucht wurde*  
 that the tractor-NOM to repair tried was  
 ‘that they tried to repair the tractor’
- (5) \**dass der Traktor zu reparieren geplant wurde*  
 that the tractor-NOM to repair planned was  
 ‘that they planned to repair the tractor’

Numerous properties other than these have been associated with monoclausal configurations containing infinitives (s. above section 1). Albeit many of these properties are language-specific, and sometimes even speaker-specific (s. Wurmbrand 2003: 7; Cinque 2006: 44–45), it seems that a core class of verbs embedding infinitives in which monoclausal properties or restructuring effects are available can be identified in a cross-linguistically valid way, comprising modal, aspectual and motion verbs (Krapova & Cinque 2018), and beside the aforementioned, also causative (Wurmbrand 2003) and implicative verbs (Grano 2015). On the other hand, predicates without restructuring effects are found among factive and propositional verbs (e.g. *let*, *claim*, *forget*, *manage*, *decide*, *plan*), as Wurmbrand (2003: 6–7) notes.

I will now survey some of the contributions that have been made in the study of verbs that instantiate infinitival and restructuring configurations in Croatian and Serbian.<sup>32</sup> In the nineties, Progovac (1993, 1996) discusses the cases of transparency effects in Serbian obtained with modals such as *morati* ‘must’, which – as opposed to *kazati* ‘tell’ or *tvrditi* ‘claim’ – allows *ni*-NPI licensing and clitic climbing out of the embedded finite complement clause. She thus applies observations on restructuring on finite embedded contexts as well. Bošković (2001) considers restructuring to the extent of its interplay with clitic climbing in his work on clitics in S/C. Stjepanović (2004) investigates the phenomenon of clitic climbing both in non-finite context and in what appears to be a finite embedded *da*-complement clause triggered by verbs *morati* ‘must’ or *željeti* ‘want’.<sup>33</sup> The relation between restructuring

<sup>32</sup> Different authors use different labels for referring to Croatian, Serbian, and Bosnian. In the abbreviations used in the text – such as S/C or B/C/S – B stands for Bosnian, C for Croatian, and S for Serbian.

<sup>33</sup> Regarding the nature of restructuring, Bošković (2001) argued for the optionality of restructuring. Stjepanović (2004) on the other hand advocates the obligatoriness of restructuring with infinitives, whereas its status in finite environments depends on the choice of the matrix predicate.

and clitic climbing in Serbian was taken into consideration by other scholars as well, such as Aljović (2005) and Cinque (2006: 20–21). Recently a group of researchers investigated the phenomenon of clitic climbing in B/C/S in a series of corpus-based studies (Jurkiewicz-Rohrbacher *et al.* 2017; Hansen *et al.* 2018; Kolaković *et al.* 2019; Kolaković, Jurkiewicz-Rohrbacher *et al.* 2022). Although their focus of investigation led them to mention the raising-control dichotomy, the discussion of the two types of matrix predicates was marginal compared to their aim at identifying distributional properties of clitics and constraints on clitic climbing.<sup>34</sup> It is worth mentioning that Gračanin-Yukseš (2019) discusses the availability of clitic climbing in Croatian subject-control constructions and its unavailability in object-control constructions. She claims that object-control and subject-control constructions differ in their syntactic structure: while the embedded non-finite clause in an object-control configuration are structurally richer CPs, subject-control configurations lack the CP projection which is why clitics can climb in such environment. This leads Gračanin-Yukseš (2019) to conclude that subject-control configurations involve TPs or vPs.

Turning the attention to the properties of verbs embedding infinitives in general, it is worth mentioning other studies as well. Mihaljević (2009: 326–336) in an analysis of Croatian perception verbs such as *gledati* ‘watch’, *čuti* ‘hear’, *vidjeti* ‘see’, *slušati* ‘listen’ states that they do not introduce clausal complements for a number of reasons: they cannot be interrogative while at the same time direct questions are grammatical, they cannot be negated or modified with a temporal adverb with the temporal specification different from those expressed in a matrix clause. Gnjatović & Matasović (2013) examine control constructions offering a well-founded classification and description of Croatian control verbs, without including observations on restructuring effects. Batinić Angster (2019) considers the apparently long-distance bound possessive-reflexive anaphor *svoj* in the context of object-oriented permissive predicates that take infinitives, such as *dopustiti* and *pustiti* ‘allow, let’, *ostaviti* ‘leave’, *dati* ‘allow, let’, *dozvoliti* ‘permit’. The author shows their restructuring nature along with a list of other verbs categorized as follows: modals (*smjeti* ‘be allowed to’, *moći* ‘can’, *morati* ‘must’, *trebati* ‘need’, *imati* ‘have to, must’, *htjeti* ‘want’), some conative verbs (*uspjeti*

<sup>34</sup> It is worth mentioning, however, that they found out clitic climbing is not obligatory out of stacked infinitives, i.e. multiply embedded infinitives, or that they observed there could be differences related to the presence or absence of reflexive marker on verbs.

‘manage’, *nastojati* ‘try’, *pokušati* ‘try’, *usuditi se* ‘dare’, *izbjeci* ‘avoid’), verbs expressing different cognitive processes (*željeti* ‘want’, *voljeti* ‘love’, *mrziti* ‘hate’, *znati* ‘know’, *misliti* ‘think’, *nadati se* ‘hope’, *bojati se* ‘(be) afraid’), aspectuals (*početi* ‘start’, *nastaviti* ‘continue’, *završiti* ‘finish’, *stati* ‘stop’), some motion verbs (*ići* ‘go’, *doći* ‘come’). Recently, Croatian and Serbian infinitival and non-finite-like contexts were discussed in a number of studies on restructuring in a cross-linguistic perspective, on complementation and finiteness (Wurmbrand 2016; Stjepanović 2004; Todorović & Wurmbrand 2016; Wurmbrand *et al.* 2020) or in studies dedicated to certain classes of verbs that introduce infinitives (e.g. Veselinović 2017 on modals).

It is important to note that not all of the observations about restructuring considered at the beginning of this survey apply to Croatian. For example, the phenomenon of clitic climbing out of the finite complement clause embedded by a modal verb – therefore restructuring in finite contexts or «restructuring despite finiteness» (Wurmbrand 2016) – cannot be evaluated in Croatian when considering clausal complements to modals such as *morati* ‘must’, because in such environment a non-finite complement is the only one available (at least as far as my competence of native speaker is concerned).<sup>35</sup> In fact, Wurmbrand *et al.* (2020), Wurmbrand & Lohninger (2023) have already noted that the availability of finite and non-finite complements differs in Croatian and Serbian. They observed microparametric variation in South Slavic languages (Bosnian, Serbian, Croatian, and Slovenian) regarding the availability of infinitival complements and claimed that infinitives as complements seem to be more common and frequent in Croatian than in Serbian. Therefore, it is worth dedicating some attention to the properties of infinitival complements that characterize Croatian alone.

Before considering Croatian subject-oriented configurations with infinitives more in depth, I will introduce few important aspects of restructuring and define the theoretical approach to restructuring that influenced my discussion.

<sup>35</sup> A worthwhile remark about Croatian modals concerns their ability to introduce finite when it is in its impersonal form in the present tense and with an epistemic reading:

<i>Mora</i>	<i>da</i>	<i>su</i>	<i>uranili.</i>
must.PRS.3SG	that	be.PRS.3PL	come_early.PTCP.ACT.M.PL

‘It must be that they have come earlier.’

### 3. Theoretical background: verbal meaning and monoclausal effects

Two main strands of theoretical approaches to monoclausal  $V_{\text{finite}} + V_{\text{non-finite}}$  configurations may be identified in generative grammar, namely functional and lexical restructuring. The proponents of the so-called functional approach (e.g. Cinque 2004, 2006) consider the verb introducing the infinitive to be a functional head having the latter as the main predicate of the clause. On the contrary, the proponents of an alternative view (e.g. Wurmbrand 2003, 2004) argue for a dominantly lexical approach to restructuring, claiming that predicates embedding infinitives are mostly lexical verbs that select non-finite VP complements. In what follows I will refer mainly to works representing the functional perspective, because these contributions to the study of restructuring guided my analysis of Croatian verbs selecting infinitives.

Restructuring effects are restricted only to some verbs ( $V_{\text{finite}}$  in  $V_{\text{finite}} + V_{\text{non-finite}}$  configurations), but the reason why it is so is still a matter of discussion. Grano (2015), inspired by works by Wurmbrand (1998, 2001) and Cinque (2004), highlights two correlations that seem to be relevant for explaining that restriction: the one between restructuring and the partial/exhaustive control split (first noted by Wurmbrand 1998) and the correlation between restructuring and Cinque's (2004) IP hierarchy of inflectional heads.<sup>36</sup>

Following Wurmbrand's (2001, 2003) survey of infinitives in five languages (German, Dutch, Italian, Spanish, Japanese), Grano (2015) classifies the predicates she identified as restructuring into the semantic classes proposed by Landau (2000). Grano's (2015) integration of their works – reported in Figure 1 – aims at distinguishing between verbs that participate in exhaustive control from those that participate in partial control. The scholar realizes that the vast majority of exhaustive control predicates tends to restructure, whereas partial control predicates are

<sup>36</sup> As for the nature of restructuring, according to Cinque (2004), monoclausality is the defining property of restructuring predicates that remains stable throughout the derivation. Albeit Cinque (2004) and Grano (2015) agree in that restructuring classes of verbs realize functional heads, Grano (2015, 2017) believes their functional nature is derived and originally lexical taking as evidence that a grammaticalization process starts from lexical and develops into functional.

generally non-restructuring predicates. As a result, Grano (2015: 15) posits a generalization according to which exhaustive verbs are restructuring, while the partial control verbs are not.<sup>37</sup>

Table 1. Grano's (2015: 15) integration of Wurmbrand's (2001) classification of restructuring predicates and Landau's (2000) exhaustive control (EC) and partial control (PC) verb classes

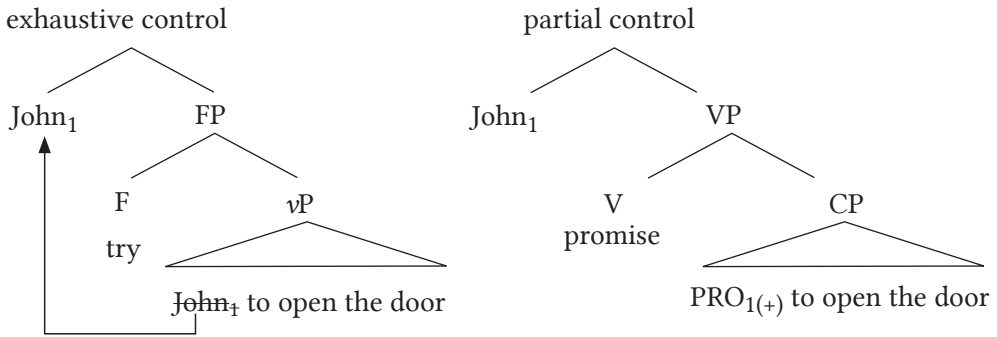
Landau's classes	Wurmbrand's predicates	Restructuring status (Wurmbrand 2001: 342)				
		German	Dutch	Italian	Spanish	Japanese
<b>EC: aspectual</b>	<i>begin</i>	+	+	+	+	+
	<i>continue</i>	N/A	N/A	+	+	+
	<i>start</i>	N/A	N/A	+	+	+
	<i>finish, stop</i>	N/A	N/A	+	+	+
	<i>be about to</i>	N/A	N/A	+	+	+
	<i>(be) used to</i>	+	+	+	+	N/A
<b>EC: modal</b>	<i>can, may</i>	+	+	+	+	+
	<i>must, need</i>	+	+	+	+	+
	<i>know how</i>	N/A	+	+	+	N/A
	other (semi-)modals	+	+	+	+	+
<b>EC: implicative</b>	<i>manage/succeed</i>	+	+	±	±	+
	<i>forget</i>	+	+	±	±	+
	<i>fail</i>	+	N/A	N/A	N/A	+
	<i>dare</i>	+	+	±	-	N/A
<b>EC: (other)</b>	<i>try</i>	+	+	±	±	±
<b>PC: desiderative</b>	<i>want</i>	+	+	+	+	+
	<i>intend (≈ want, mean)</i>	+	+	±	±	N/A
	<i>refuse, reject</i>	-	+	-	-	-
	<i>prefer</i>	-	-	-	-	-
	<i>decide, choose</i>	-	-	-	-	-
	<i>plan</i>	-	-	-	-	-
	<i>offer</i>	-	-	-	-	-
	<i>wish</i>	-	-	-	-	-
	<i>propose</i>	-	-	-	-	-
<b>PC: factive</b>	factive	-	-	-	-	-
<b>PC: propositional</b>	propositional	-	-	-	-	-

<sup>37</sup> Grano (2015) at the same time, however, admits that some partial control predicates tend to restructure, namely desiderative verbs, and that some exhaustive verbs, such as weak implicatives, may fail to restructure.

In Table 1 the abbreviation N/A following some predicates means that it was not possible to apply the classification to them (‘not applicable’). The sign ± indicates that the verb in question is subject to inter-speaker variation or that its behaviour is not uniform.

The exhaustive control constructions, as Grano (2015) claims, have monoclausal syntax whereby a predicate takes a *vP* complement with subject raising – which explains why restructuring effects are available only with those predicates. On the other hand, the structure of partial control constructions is biclausal in that partial control verbs have true control status instantiating a CP complement with PRO as a subject. The difference in the two structures is illustrated below in Figure 1.

Figure 1. The structures instantiated by a restructuring verb *try* and a non-restructuring *promise* (Grano 2015: 17)



Cinque’s (2004, 2006) account of restructuring predicates is outlined in his cartography research programme. He observes that the ordering of restructuring predicates is fixed just like the ordering of inflectional morphemes and adverbs (proposed in Cinque 1999) and claims that this is a consequence of them being realization of functional heads in the inflectional layer of the clause. Cinque (2004, 2006: 12) proposes an IP hierarchy that constitutes a cross-linguistically valid functional portion of the clause, an excerpt of which is reported in Krapova & Cinque (2018: 11), and in (6).<sup>38</sup>

<sup>38</sup> Combining the hierarchy proposed by Cinque (1999: 106) and a portion of it laid out in Cinque (2006: 12), the order of all the relevant functional heads is as follows: MoodP<sup>act</sup> > MoodP<sup>evaluative</sup> > MoodP<sup>evidential</sup> > ModP<sup>epistemic</sup> > TP(Past) > TP(Future) > MoodP<sup>speech</sup> > ModP<sup>necessity</sup> > ModP<sup>possibility</sup> > ModP<sup>alethic</sup> > AspP<sup>habitual</sup> > Asp<sup>repetitive</sup> > AspP<sup>frequentative</sup> > ModP<sup>volitional</sup> > AspP<sup>celerative</sup> > TP(Anterior) > AspP<sup>terminative</sup> > AspP<sup>continuative</sup> > AspP<sup>retrospective</sup> > AspP<sup>proximative</sup> > AspP<sup>durative</sup> > AspP<sup>generic/progressive</sup> > AspP<sup>prospective</sup> > ModP<sup>obligation</sup> > ModP<sup>permission/ability</sup> > AspP<sup>completive</sup> > VoiceP > AspP<sup>celerative(II)</sup> > AspP<sup>repetitive(II)</sup> > AspP<sup>frequentative(II)</sup>.



Cinque head	Corresponding predicates	Restructuring status
Asp <sub>inceptive</sub>	<i>to start</i>	+
Mod <sub>obligation</sub>	<i>have, must</i>	+
Mod <sub>ability</sub>	<i>can, be able</i>	+
Asp <sub>frustrative</sub>	<i>forget</i>	%
Asp <sub>success</sub>	<i>manage</i>	%
Mod <sub>permission</sub>	<i>can, may</i>	+
Asp <sub>conative</sub>	<i>try</i>	%
Asp <sub>completive</sub>	<i>finish</i>	+

The table shows that the meanings of predicates that participate in monoclausal configurations correspond to heads below Tense on the functional spine of the clause proposed by Cinque (2004, 2006). This leads Grano (2015: 62) to the hypothesis in (7):

- (7) **Restructuring hypothesis (final):** A verb *V* restructures just in case the meaning of *V* is subsumed by the meaning of an inflectional-layer functional head *F*, and either *F* is below Tense or *V* is non-subject-oriented.

Above Tense, predicates that lack restructuring status are listed, and here we can find predicates with the following meanings: speech act, evaluative, evidential, epistemic. In conclusion, we can observe that the functions matrix verbs realize, i.e. their meanings, correspond to specific syntactic properties of the configurations in which they participate. In other words, the presence or absence of monoclausal effects in a configuration depends on the choice of the matrix predicate, as cross-linguistic evidence suggests.

#### 4. Monoclausal $V_{\text{finite}} + V_{\text{non-finite}}$ configurations in Croatian

The monoclausal character of  $V_{\text{finite}} + V_{\text{non-finite}}$  configurations, signalled by the presence of transparency effects, is related to a particular functional notion expressed by a finite verb that introduces the non-finite one. It should be noted that Croatian infinitives are bare infinitives (e.g. *moram ići* ‘I must go, I have to go’), with few exceptions (s. below interrogative verbs). Before considering arguments for a monoclausal approach to certain configurations with infinitives in Croatian, I will classify Croatian subject-oriented verbs.

## 4.1. Croatian subject-oriented predicates with infinitives: classification

Croatian predicates embedding infinitive can be classified into six out of seven semantic classes proposed by Landau (2000: 38). See in Table 3 a tentative classification of Croatian subject-oriented predicates selecting infinitives (with the exception of propositionals, see below).<sup>40</sup>

Table 3. Croatian subject-oriented predicates embedding infinitives classified into Landau's (2000) classes

Landau's (2000) classes	Corresponding verbs in Croatian
<b>modals</b>	<i>moći</i> 'can', <i>morati</i> 'must', <i>trebati</i> 'need', <i>smjeti</i> 'be allowed to', <i>imati</i> 'must, have to'
<b>aspectuals</b>	<i>početi</i> 'start, begin', <i>nastaviti</i> 'continue', <i>završiti</i> 'finish', <i>prestati</i> 'stop'
<b>implicatives</b> <sup>41</sup>	<i>uspjeti</i> 'manage', <i>izbjeći</i> 'avoid', <i>usuditi se</i> 'dare', <i>sjetiti se</i> 'remember', <i>zaboraviti</i> 'forget'
<b>desideratives</b>	<i>odlučiti</i> 'decide', <i>planirati</i> 'plan', <i>obećati</i> 'promise', <i>odbiti</i> 'refuse, reject', <i>željeti</i> 'want', <i>nadati se</i> 'hope', <i>ponuditi se</i> 'offer', <i>bojati se</i> '(be) afraid', <i>namjeravati</i> 'intend', <i>misliti</i> 'mean, intend', <i>odabrati</i> 'choose', <i>težiti</i> 'strive', <i>(biti) spreman</i> '(be) ready'
<b>factives</b>	<i>mrziti</i> 'hate', <i>voljeti</i> 'love, like', <i>(biti) žao</i> '(be) sorry', <i>(biti) drago</i> '(be) glad'
<b>interrogatives</b>	<i>pitati se</i> 'wonder', ask', <i>znati</i> 'know'
<b>propositionals</b> <sup>42</sup>	<i>tvrditi</i> 'claim', <i>izjaviti</i> 'assert, declare', <i>vjerovati</i> 'believe', <i>pretpostavljati</i> 'suppose'

<sup>40</sup> The list of Croatian predicates is not exhaustive, it comprises a selection of members of each class.

<sup>41</sup> Among implicatives Landau (2000: 38) lists *decline*, and *refuse* among desideratives. Since in Croatian a single verb (*odbiti*) corresponds to both of them, I decided not to include it in my survey of Croatian verbs.

<sup>42</sup> The Croatian verb *misliti* in Table 3 is included in the class of desideratives. However, its main meaning corresponds to that of English *think*, a propositional verb in Landau's (2000: 38) classification, embedding a finite complement. That is the reason why I do not list it among propositionals with infinitives. When *misliti* expresses the meaning 'intend' or 'mean', thus belonging to desideratives, it selects a non-finite complement verb.

Modal and aspectual verbs express meanings pertaining to the notions of modality and aspect. More specifically, considering functions in Cinque's (2004) hierarchy of functional heads, modals can express epistemic, deontic or dynamic modality.<sup>43</sup> Epistemic predicates are higher than Tense on the functional spine of the clause, whereas obligation (e.g. *morati* 'must', *trebati* 'need', *imati* 'must, have to'), permission (e.g. *smjeti* 'be allowed to'), and ability (e.g. *moći* 'can') predicates are comprised below Tense in Cinque's hierarchy.

Aspectuals can realize inceptive (e.g. *početi* 'start, begin'), continuative (e.g. *nastaviti* 'continue'), terminative (e.g. *prestati* 'stop'), or completive (e.g. *završiti* 'finish') meanings. All these meanings are listed under Tense.

As for the class of implicative verbs, the truth of their complement is entailed if they are positive (e.g. *uspjeti* 'manage', *izbjeći* 'manage not to do'), whereas it is denied in case of negative or negated implicatives (e.g. *ne uspjeti* 'fail, not manage') (cf. Karttunen 1971; Landau 2000: 37; Grano 2015: 9). To put it differently – as Landau (2000: 37) clarifies it following Karttunen (1971) – implicatives «presuppose that the truth of the entire sentence is a necessary and sufficient condition for the truth (or falsity) of the complement». Taking into consideration the functions on Cinque's (2004) list, Croatian implicatives such as *uspjeti* and *izbjeći* express success, while some of them express the frustrative function (e.g. *zaboraviti* 'forget'), both listed below Tense. Clearly, verbs such as *sjetiti se* 'remember' and *zaboraviti* 'forget', despite belonging to Landau's (2000) class of implicatives, do not correspond to Cinque's (2004) function success. The class of implicatives in Croatian deserves a more thorough discussion which however goes beyond the scope of this paper. For this reason the following discussion will focus only on verbs *uspjeti* 'manage', *izbjeći* 'avoid' and *usuditi se* 'dare'.

The class of desiderative predicates in Landau's (2000) sense is a heterogeneous class comprising verbs expressing desire such as *željeti* 'want', intention (e.g. *namjeravati* 'intend', *planirati* 'plan', *odlučiti* 'decide').<sup>44</sup> Among these, verbs expressing volition (*željeti* 'want'), occupy a

<sup>43</sup> According to Landau (2000: 37, 60) dynamic modals expressing ability in English are unquestionably control predicates, whereas the others might be raising.

<sup>44</sup> Landau (2000) also includes verbs expressing command within the class of desiderative predicates, but I failed to find Croatian verbs expressing that meaning which are able to embed a non-finite complement and in which the implicit subject is coreferential with the matrix subject.

position below Tense in Cinque's (2004) terms, whereas those with speech act function (e.g. *obećati* 'promise') are listed above Tense.

Both positive and negated factives embed complements whose truth is presupposed (Kiparsky & Kiparsky 1970; Landau 2000: 37–38; Grano 2015: 9). Some of them are emotive factives (e.g. *(biti) žao* '(be) sorry') expressing the speaker's «subjective value of a proposition rather than knowledge about it or its truth value» (Kiparsky & Kiparsky 1970: 169). According to Cinque's (2004) IP hierarchy emotive factives are evaluative and as such precede Tense.

Croatian interrogatives embedding non-finite complements such as *pitati se* 'wonder' introduce interrogative complements mostly via the complementizer *kako* 'how'. They can also have other forms of interrogative complement clause characterized by the presence of the interrogative particle *li* introduced by *da* 'that'. In Cinque's (2004) IP hierarchy these realize the epistemic function.

As epistemic or declarative in nature (Landau 2000: 37), propositional predicates introduce complements whose «truth or falsity can be predicated», as Grano (2015: 9) states, exemplifying it with the following sentence: *John claimed to have solved the problem, which was true*. They introduce an infinitive in English. However, Croatian propositional verbs embed almost exclusively finite complements. Therefore, the list of Croatian propositional matrix predicates is highlighted in grey in Table 3 and they will not be considered any further.

#### 4.2. Croatian subject-oriented predicates with infinitives: monoclausal-type vs. biclausal-type

Assuming that the availability of restructuring effects is related to the meaning of the verb that selects an infinitive and taking Grano's (2015) observations presented in section 3 as valid, the following hypothesis can be made:

- (8) In Croatian monoclausal properties will be obtained with infinitives selected by modals, aspectuals and implicatives.

These three classes of predicates encode meanings below Tense in Cinque's (2004) hierarchy and, according to Landau (2000), in English they belong to the category of exhaustive control verbs. Desiderative verbs expressing volition, such as *željeti* 'want', although occupying a lower position with

respect to Tense, will be left aside since they are not at the same time exhaustive control (EC) predicates. As Landau (2000: 5) put it, the «*want*-type complements fall under PC [partial control]». <sup>45</sup>

Therefore, in Table 4, in which Landau's classes are crossed with Cinque's (2004, 2006) functions expressed by the members of particular classes, I distinguish between monoclausal-type and biclausal-type predicates in Croatian. On the one hand, modals, aspectuals, and implicatives should share monoclausal properties, while on the other hand, the remaining subject-oriented predicates should lack them. This latter group should include both desideratives and factives listed in Table 3. Propositionals and interrogatives are no longer included because, as seen in 4.1 above, very few of them can embed non-finite complements in Croatian.

Table 4. Croatian subject-oriented predicates with infinitives: monoclausal-type vs. biclausal-type

Landau's (2000) classes	Cinque's (2004, 2006) functions	Predicates with infinitives in Croatian	Configuration
modals	obligation, ability, permission	<i>moći</i> 'can', <i>morati</i> 'must', <i>trebati</i> 'need', <i>smjeti</i> 'be allowed to', <i>imati</i> 'must, have to'	monoclausal-type
aspectuals	inceptive, continuative, terminative, completive	<i>početi</i> 'start', begin', <i>nastaviti</i> 'continue', <i>završiti</i> 'finish', <i>prestati</i> 'stop'	
implicatives	success	<i>uspjeti</i> 'manage', <i>izbjeci</i> 'avoid', <i>usuditi se</i> 'dare'	
other subject-oriented predicates: desideratives, factives	volitional, evaluative, speech act	<i>odlučiti</i> 'decide', <i>planirati</i> 'plan', <i>obećati</i> 'promise', <i>željeti</i> 'want', <i>nadati se</i> 'hope', <i>ponuditi se</i> 'offer', <i>bojati se</i> 'afraid', <i>namjeravati</i> 'intend', <i>misлити</i> 'mean, intend', <i>odabrati</i> 'choose', <i>mrziti</i> 'hate', <i>voljeti</i> 'love, like', etc.	biclausal-type

In 4.3. I will continue to pursue the distinction between these two types of predicates: modals, aspectuals, and implicatives represent the monoclausal-

<sup>45</sup> The distinction between partial and exhaustive control will be considered in 4.3.1.

type, whereas the biclausal-type is represented by a rather heterogeneous group of verbs that can be classified as desideratives and factives (Landau 2000) or as verbs denoting volitional, evaluative and speech act functions in Cinque's (2015) sense.<sup>46</sup> It can also be observed that the biclausal-type predicates express a range of cognitive processes. In fact, some of them are often classified as attitude or attitude-ascribing predicates, e.g. mental attitude predicates, such as *željeti* 'want', *obećati* 'promise' (s. Pearson 2012, 2016). Sometimes I will refer to the biclausal type of predicates as *decide*-type predicates following Wurmbrand (2003): *odlučiti* 'decide' is an exemplary member of the class and a rather stable one in its restructuring nature.

### 4.3. Properties of monoclausal configurations with infinitives in Croatian

As stated in 4.2, I expect that monoclausal properties of configurations with infinitives in Croatian will be obtained with three classes of Croatian verbs: modals, aspectuals, and implicatives. In what follows I aim at verifying this expectation, i.e. at testing the hypothesis given in (8). I will analyse the members of the three classes given in Table 4 by applying several tests for monoclausality to examples built and judged based on my intuitions. I will compare non-finite configurations taken by the – presumably restructuring – classes of verbs with the ones embedded by other – presumably non-restructuring – subject-oriented verbs. I will first consider the split between partial and exhaustive control, showing that the monoclausal-type verbs do not admit semantic mismatches between the matrix subject and the understood subject of the embedded predicate, thus exhibiting a property of exhaustive control verbs. I will then show that these configurations display other relevant properties such as preserved synonymy under passivization, matrix licensing of negative polarity items, unavailability of finite complementation and, finally, temporal uniformity. The latter

---

<sup>46</sup> Beside the one I follow in this paper, there is another influential approach recently proposed, i.e. the three-way classification in attitude, *irrealis* and tenseless or in proposition, event and situation complements (see e.g. Wurmbrand *et al.* 2020; Kovač *et al.* 2022; Wurmbrand & Lohninger, 2023). One of the aspects in which this approach significantly diverges from the position taken here is reflected in conceiving of the distinction between monoclausal and biclausal properties of complex structures as a scale rather than a binary opposition.

property regards the unavailability of tense-mismatches between matrix and embedded event descriptions and the availability of simultaneously interpreted event descriptions.

#### 4.3.1. Exhaustive control: unavailability of semantic mismatches

The bipartition of subject control in exhaustive (EC) and partial control (PC) is acknowledged among researchers of non-finite complementation. One of the aspects in which EC and PC differ is related to the identity relation between the controller and the controlee. In case of PC, the embedded non-finite predicate can be semantically plural notwithstanding the presence of syntactically singular subject in the matrix clause. In other words, in PC structures mismatches in semantic plurality are allowed since PRO can appear with a collective reading being thus incompatible with a syntactically singular matrix subject (cf. Grano 2015: 8–9; Pitteroff *et al.* 2017: 141–143).

Consider the difference between (9) and (10) (Pitteroff *et al.* 2017: 141–143).

- (9) *John<sub>i</sub> wanted/preferred PRO<sub>i+</sub> to gather at six.* PC  
 (10) \**John<sub>i</sub> tried/managed PRO<sub>i+</sub> to gather at six.* EC

When the matrix subject is singular, lexically collective predicates such as *gather* are disallowed with *try* and *manage* as matrix predicates, while perfectly acceptable if introduced by *want* and *prefer*. The same can be observed when the collectivizer *together* is placed in a complement clause (Landau 2000: 44–45): whereas (11) and (12) are acceptable, the same cannot be claimed for (13) (Grano 2015: 9).

- (11) *John claimed/wanted to (have) solved the problem (together).* PC  
 (12) *John enjoyed solving the problem (together).* PC  
 (13) *John began/had to/managed to solve the problem (\*together).* EC

This split is tied to the choice of the matrix predicate, which was noted by Landau (2000). Propositional, desiderative, factive, and interrogative predicates are predicate classes that tend to allow partial control. Predicates that do not allow it are modals, aspectuals, and implicatives.<sup>47</sup>

<sup>47</sup> See Landau (2000) and Grano (2015). However, as reported by various linguists and experimental evidence (e.g. White & Grano 2014; Pitteroff *et al.* 2017) the availability of PC or EC readings is subject both to cross-linguistic and intra-linguistic variation.

The correlation between the choice of the predicate and the cited effects is confirmed by Croatian predicates as well. Modals, aspectuals and implicatives disallow mismatches in semantic plurality, that is they disallow partial control readings, which can be observed in the contrast between (14a) and (14b).

- (14) a. \**On se morao<sub>mod</sub>/uspio<sub>impl</sub> PRO<sub>i+</sub> naći | počeo<sub>asp</sub> PRO<sub>i+</sub> nalaziti u šest.*  
 ‘He had/managed to meet<sub>pfv</sub> | started to meet<sub>ipfv</sub> at six.’  
 b. *On je morao<sub>mod</sub>/uspio<sub>impl</sub> riješiti | počeo<sub>asp</sub> rješavati probleme (\*zajedno).*  
 ‘He had/managed to resolve<sub>pfv</sub> | started to resolve<sub>ipfv</sub> the problems (together).’

Collective readings notwithstanding the presence of a singular matrix subject are allowed only with what, following Wurmbrand (2003), we can call *decide*-type predicates:

- (15) a. *On se odlučio/planirao/želio PRO<sub>i+</sub> naći | mrzio PRO<sub>i+</sub> nalaziti u šest.*  
 ‘He decided/planned/wanted to meet<sub>pfv</sub> | hated to meet<sub>ipfv</sub> at six.’  
 b. *On je odlučio/planirao/želio riješiti | mrzio rješavati probleme zajedno.*  
 ‘He decided/planned/wanted to resolve<sub>pfv</sub> | hated to resolve<sub>ipfv</sub> the problems (together).’

#### 4.3.2. Availability of synonymy under passivization

Availability of synonymy under passivization is known to be a diagnostic for distinguishing raising from control constructions (Davies & Dubinsky 2004: 6). If applied to Croatian, it can be observed that the passivized sentence remains synonymous to its active counterexample only if the infinitive is selected by modal, aspectual, and implicative verbs (16), while with control predicates such as *odlučiti* ‘decide’ the synonymy is not preserved (17).

- (16) *Profesor je trebao<sub>mod</sub>/uspio<sub>impl</sub> ispitati | počeo<sub>asp</sub> ispitivati<sub>i</sub> učenika.*  
 = *Učenik je trebao<sub>mod</sub>/uspio<sub>impl</sub> biti ispitan | počeo<sub>asp</sub> biti ispitivan.*  
 ‘Professor had/managed to interrogate<sub>pfv</sub> | started to interrogate<sub>ipfv</sub> a student.’  
 = ‘The student had/managed to be interrogated<sub>pfv</sub> | started to be interrogated<sub>ipfv</sub>.’
- (17) *Profesor je odlučio/planirao/želio ispitati | mrzio ispitivati učenika.*  
 ≠ *Učenik je odlučio/planirao/želio biti ispitan | mrzio biti ispitivan.*

‘Professor decided/planned/wanted to interrogate<sub>pfv</sub> | hated to interrogate<sub>ipfv</sub> the student.’  
 ≠ ‘The student decided/planned/wanted to be interrogated<sub>pfv</sub> | hated to be interrogated<sub>ipfv</sub>.’

Therefore, in Croatian decide-type predicates fail this test, whereas the three verbal classes – modals, aspectuals, and implicative – display a characteristic of raising predicates.

#### 4.3.3. Matrix licensing of negative polarity items

Considering that under restructuring «apparently biclausal structures act as transparent domains for ordinarily clause-bound process» (Grano 2015: 11), it is important to investigate the distribution of negative polarity items (e.g. *nitko* ‘nobody’, *nikamo* ‘nowhere’, *ništa* ‘nothing’) in infinitival configurations.

Croatian is a negative concord language in which negative indefinites are licensed clause-internally. The so-called *ni*-NPIs (e.g. *nitko* ‘nobody’, *nikamo* ‘nowhere’, *ništa* ‘nothing’, etc.) are under the local scope of negation (cf. Laka 1990; Zovko Dinković 2013, 2021; Zovko Dinković & Ilc 2017; Zanuttini 1991; Grano 2015):<sup>48</sup>

- (18) a. *Ivan ne radi ništa.*<sup>49</sup>  
           Ivan not do.3SG.PRES nothing  
       b. \**Ivan radi ništa.*  
           Ivan do.3SG.PRES nothing  
           ‘Ivan does not do anything.’

While matrix licensing of NPIs out of the local and apparently embedded non-finite clause is possible when the infinitive is taken by a modal, aspectual, or implicative verb (19), it is degraded, if not unacceptable, with *decide*-type predicates (20).

<sup>48</sup> The examples are taken from Batinić Angster (2019).

<sup>49</sup> In this paper glosses will be provided for examples whose structure deviates from the structure of the English translation and these deviations are relevant for the argumentation. Glosses follow the Leipzig Glossing Rules (Bickel *et al.* 2015).

- (19) *Profesor nije uspio ispitati*  
 professor[NOM.SG] not\_be.3SG.PRES manage.PTCP.3SG.M examine.INF  
*nikoga.*  
 nobody.ACC.SG  
 ‘Professor didn’t manage to orally examine anybody.’
- (20) *Profesor nije odlučio ispitati*  
 professor[NOM.SG] not\_be.3SG.PRES decide.PTCP.3SG.M examine.INF  
*nikoga.*  
 nobody.ACC.SG  
 ‘Professor didn’t decide to orally examine anybody.’

Considering that Croatian NPIs require local licensors, the domains in which they can be licensed out of the non-finite clause are local. This is certainly the case when infinitives are introduced by modals, aspectuals, and implicatives. However, the judgements on that matter are subject to inter-speaker variation, some speakers might not find them ill-formed, despite being somewhat hard to process.

#### 4.3.4. Unavailability of finite complementation

The classes of Croatian predicates I assume to display restructuring effects – monoclausal-type predicates – cannot project finite clausal complements. Modals, aspectuals, and few implicatives require a non-finite complementation. Consider (21):

- (21) *\*Profesor je trebao<sub>mod</sub> / uspio<sub>impl</sub> da ispita | počeo<sub>asp</sub> da ispituje sve studente.*  
 ‘The professor had/managed to interrogate<sub>pfv</sub> | started to interrogate<sub>ipfv</sub> all students.’

As for the other subject-oriented predicates, grouped so far under the name of *decide*-type verbs or biclausal-type predicates, they do not display a homogeneous behaviour regarding the possibility of introducing a finite clausal complement.<sup>1</sup> The subject-oriented predicates in (22) can embed both non-finite and finite complements via the complementizer *da* ‘that’:

- (22) *Profesor je odlučio/obećao | se nadao/bojao da će ispitati sve studente.*  
 ‘The professor decided/promised | hoped/was afraid that he will interrogate all students.’

On the other hand, not all verbs in the heterogenous class of desideratives in Landau’s (2000) sense share the availability of both types of complementation.

For example, *željeti* ‘want’, the core member of the desiderative class, requires an infinitive, as well as *planirati* ‘plan’, *namjeravati* ‘intend’, *misliti* ‘mean, intend’,<sup>50</sup> *odabrati* ‘choose’:

- (23) \**Profesor je želio/planirao/namjeravao/mislio da će ispitati sve studente.*  
lit. ‘The professor wanted/planned/intended that he will interrogate all students.’

Concerning (emotive) factives, it is worth noting that *da*-complementation is completely unacceptable with *mrziti* ‘hate’ and *voljeti* ‘love’ (s. (24a)), while it only tends to be unavailable with (*biti*) *žao* ‘(be) sorry’, (*biti*) *drago* ‘(be) glad’. Confront the ill-formed (24b) to the well-formed (24c).

- (24) a. \**Profesor je mrzio da će ispitivati sve studente.*  
professor.NOM.SG be.PRS.3SG hate.PTCP.ACT.M.SG that will.PRS.3SG  
interrogate.INF all students
- b. \**Profesoru je bilo drago/žao da će ispitati sve studente.*  
professor.DAT.SG be.PRS.3SG be.PTCP.ACT.N.SG glad/sorry that will.PRS.3SG  
interrogate.INF all students
- c. *Profesoru će biti drago da ispita | žao da ne ispita sve studente.*  
professor.DAT.SG will.PRS.3SG be.INF glad that interrogate.PRS.3SG |  
sorry that NEG interrogate.PRS.3SG all students
- ‘The professor will be glad to interrogate | sorry that he will not interrogate all students.’

However, it seems that all factives can introduce a finite complement clause via the complementizer *što* ‘that’:<sup>51</sup>

<sup>50</sup> See note 42 regarding the properties of the verb *misliti* as far as its different meanings are concerned.

<sup>51</sup> The alternatives to subordinate *što*-clause of the sentence (24a) are: (*mrzio*) *to što je ispitao sve studente* and (*mrzio*) *činjenicu da je ispitao sve studente*, both translated in English as ‘(hated) the fact that he interrogated all students’. The behaviour of factives would need further discussion which is beyond the scope of this paper, but see e.g. Krapova & Soćanac (2022).

- (25) a. *Profesor je mrzio što je ispitao sve studente.*  
 ‘The professor hated that he interrogated all students.’
- b. *Profesoru je bilo drago/žao što je ispitao sve studente.*  
 professor.DAT.SG be.PRS.3SG be.PTCP.ACT.N.SG glad/sorry that be.PRS.3SG  
 interrogate.PTCP.ACT.M.SG all students  
 ‘The professor was glad/sorry he interrogated all students.’

For the purpose of this study, it is important to note that the members of the three classes I claim to instantiate monoclausal configurations systematically disallow finite complementation in Croatian.

#### 4.3.5. Temporal uniformity

Another evidence for treating configurations instantiated by modal, aspectual and implicative verbs as monoclausal – thus as functional heads forming complex predicates with their non-finite complements – comes from the unavailability of distinct temporal modification. In this respect, monoclausal structures are characterized by temporal uniformity (s. McCawley 1979; Marušić & Žaucer 2006; Grano 2015). Consider the unacceptability of the following example:

- (26) \**Sutra će profesor ispitati studente prekosutra.*  
 ‘Tomorrow professor will examine the students the day after tomorrow.’

Although both *sutra* ‘tomorrow’ and *prekosutra* ‘the day after tomorrow’ in (26) are temporal adverbs of position referring to the future with respect to the speech event, they place the event of examining the students on two distinct, non-overlapping positions in time which is impossible in single event descriptions. That is the reason why we consider a sentence with two contrasting temporal adverbs of position such as (26) unacceptable.

On the contrary, (27) – despite the presence of two temporal adverbials – constitutes a single event description where the adverbial *rano ujutro* ‘early in the morning’ functions as a further specification of the adverb *sutra* ‘tomorrow’.

- (27) *Sutra će profesor ispitati studente rano ujutro.*  
 ‘Tomorrow professor will examine the students early in the morning.’

In a sentence with just one event description it is possible to have only one temporal adverbial of position, that is one temporal specification intended as placement of the event on the timeline (s. Marušič & Žaucer 2006: 1097). In other words, following Cinque (2004) and Krapova & Cinque (2018: 156), monoclausal structures have «single deictic Tense». In the context of  $V_{\text{finite}} + V_{\text{non-finite}}$  configurations, it will be possible to insert a temporal adverbial of position that conflicts with a matrix tense only in an embedded complement clause. We can therefore expect non-agreeing temporal adverbials of position only in configurations with biclausal-type predicates.

Considering the acceptability of the following two examples (28a) and (28b) with *odlučiti* ‘decide’ as the matrix verb – in which *jutros* ‘this morning’ modifies the deciding and *sutra* ‘tomorrow’ places the event of going to the cinema on the timeline – we can conclude that both of these examples have biclausal syntax and that the non-finite complement of the predicate *odlučiti* is a full-fledged CP just like the finite CP introduced by a complementizer *da* ‘that’.

(28) a. *On je jutros odlučio da će sutra ići u kino.*

‘This morning he decided that tomorrow he will go to the cinema.’

b. *On je jutros odlučio ići sutra u kino.*

‘This morning he decided to go tomorrow to the cinema.’

The infinitive taken by *odlučiti* in (28b) is said to be temporally independent (s. Pitteroff *et al.* 2018) just like its finite counterpart in (28a)

If we now turn our attention to the same type of temporal modification in a configuration with a modal, aspectual, or implicative predicate (29), we clearly see that non-agreeing adverbs that signal tense-mismatches are disallowed:

(29) \**On je jutros trebao<sub>mod</sub> / počeo<sub>asp</sub> / uspio<sub>impl</sub> ići sutra u kino.*

‘This morning he managed to go tomorrow to the cinema.’

In conclusion, we can state that infinitives to modal, aspectual, and implicative predicates in Croatian are temporally dependent or anaphoric (s. Pitteroff *et al.* 2018) and that these three classes of predicates instantiate monoclausal configurations.

Grano (2015) notes the correlation between the exhaustive/partial control split and the temporal reading of infinitives. According to him, partial control predicates – i.e. those that cannot instantiate monoclausal configurations – disallow simultaneity readings of embedded infinitival complements with eventive (non-stative) semantics. English admits three possibilities of temporal relation expressed by the embedded infinitive: simultaneity, posteriority, or anteriority relation.

In the present discussion regarding the temporal interpretation of Croatian embedded infinitives, it is important to stress that, unlike English infinitives (e.g. *He managed to have opened the door*), they cannot be tensed (e.g. *\*On je uspio biti otvorio vrata*). It is, nonetheless, relevant to consider the availability of their simultaneous and posterior interpretation with respect to the matrix tense and to identify whether it correlates or not with the type of matrix verb. Considering Grano's account (2015) we can expect that Croatian modal, aspectual, and implicative verbs impose simultaneity reading on their infinitives. On the other hand, the *decide*-type verbs should admit future or posteriority reading on infinitives they select.

Let's consider the simultaneity interpretation of infinitives embedded by *odlučiti* 'decide' using an adverbial that imposes such reading, such as *istog časa* 'in that same moment' or *upravo tada* 'right then':

- (30) *?Ivan je odlučio otvoriti vrata istog časa/upravo tada.*  
 'Ivan decided to open the door in that same moment/ right then.'
- (31) *?Ivan je odlučio otići istog časa/upravo tada.*  
 'Ivan decided to go away in that same moment/ right then.'

In the examples (30) and (31) the act of deciding (the same holds for the acts of promising, hoping, etc.) seems to precede the event described by non-finite complements – being it the event of opening the door (30) or going away (31). That is the reason why these sentences sound infelicitous with a temporal adverbial imposing the simultaneity reading. The simultaneity interpretation should be acceptable only in case of the embedded statives, such as 'be happy' (Grano 2015: 109):

- (32) *Ivan je odlučio biti sretan istog časa/upravo tada.*  
 'Ivan decided to be happy in that same moment/ right then.'

At the same time, modals, aspectuals, and implicatives allow or, better yet, force overlapping or simultaneity interpretation of the eventive infinitives they introduce:

- (33) *Ivan je trebao<sub>mod</sub>/uspio<sub>impl</sub> otvoriti | počeo<sub>asp</sub> otvarati vrata istog časa/upravo tada.*  
 ‘Ivan had/ managed to open<sub>pfv</sub> | started opening<sub>ipfv</sub> the door in that same moment/ right then.’
- (34) *Ivan je trebao<sub>mod</sub>/uspio<sub>impl</sub> otići | počeo<sub>asp</sub> odlaziti istog časa/upravo tada.*  
 ‘Ivan had/ managed to leave<sub>pfv</sub> | started leaving<sub>ipfv</sub> in that same moment/ right then.’

The time of opening the door in (33) or going away in (34) overlaps with the time of needing to do so, starting, or managing in doing so.<sup>52</sup>

Table 5 offers an overview regarding the availability of different temporal readings of infinitives selected by the Croatian subject-oriented verbs taken in consideration in this paper.

Table 5. Temporal relations expressed by infinitives to Croatian verbs

Verb classes	Croatian verbs	Simultaneity relation	Posteriority relation	Anteriority relation	Configuration
modals	<i>moći</i> ‘can’, <i>morati</i> ‘must’, <i>trebati</i> ‘need’, <i>smjeti</i> ‘be allowed to’, <i>imati</i> ‘must, have to’				monoclausal-type
aspectuals	<i>početi</i> ‘start, begin’, <i>nastaviti</i> ‘continue’, <i>završiti</i> ‘finish’, <i>prestati</i> ‘stop’	+	–	N/A	
implicatives	<i>uspjeti</i> ‘manage’, <i>izbjeći</i> ‘avoid’, <i>usuditi se</i> ‘dare’				
other subject-oriented predicates	<i>odlučiti</i> ‘decide’, <i>planirati</i> ‘plan’, <i>obećati</i> ‘promise’, <i>željeti</i> ‘want’, <i>nadati se</i> ‘hope’, <i>ponuditi se</i> ‘offer’, <i>namjeravati</i> ‘intend’, etc.	– (eventive complements) + (stative complements)	+	N/A	biclausal-type
	<i>mrziti</i> ‘hate’, <i>voljeti</i> ‘love’, ( <i>biti</i> ) <i>žao</i> ‘(be) sorry’, ( <i>biti</i> ) <i>drago</i> ‘(be) glad’ <sup>53</sup>	+	–	N/A	

<sup>52</sup> The fact that the aspectual verbs force a simultaneity reading of their complement verb explains why the respondents in Marini (*this volume*, section 7.2, example (12)) opted for the use of an imperfective form of the non-finite verb when it was preceded by the verb *početi* ‘begin’, a perfective favoured by the presence of *čim* ‘as soon as’, conjunction that usually introduces a perfective environment.

<sup>53</sup> Interestingly, (emotive) factives do not display biclausal-like properties with respect to temporal readings of embedded infinitives despite having an otherwise biclausal behaviour. The reason for that may lie in their habitual interpretation which rules out a temporal mismatch. This calls for a more fine-grained analysis, which I nonetheless will not tackle any further in this paper.

Investigating two groups of predicates that select infinitives, we observed the asymmetry regarding the availability of the split of temporal deixis between finite and non-finite domain and the availability of simultaneity readings if the selected infinitive is eventive. These findings support the idea that in Croatian infinitives embedded by modals, aspectuals, and implicatives contain no embedded Tense, while the non-finite complements to *decide*-type predicates do, with the exception of emotive factives. To sum up and following Grano's (2015) line of reasoning,<sup>54</sup> this amounts to stating that the  $V_{\text{finite}} + V_{\text{non-finite}}$  structure in the latter case – with the Tense being projected separately – is biclausal, whereas in the former case a monoclausal  $V_{[\text{mod/asp/impl}]_{\text{finite}}} + V_{\text{non-finite}}$  configuration is projected.

## 5. Concluding remarks

This study is based on the cross-linguistically well-confirmed observation that monoclausal properties of configurations with infinitives are tied to the choice of the matrix predicate, i.e. the meanings it encodes. Following that, I classified subject-oriented Croatian verbs embedding infinitives into semantic classes proposed by Landau (2000) for English. Considering Cinque's (2004, 2006) and Grano's (2015) theoretical accounts, I identified three Croatian verbal classes expected to give rise to monoclausal configuration: modals, aspectuals, and implicatives. Unlike other subject-oriented predicates taking infinitives, such as desideratives and factives, they belong to exhaustive control verbs and display properties such as preserved synonymy under passivization, matrix licensing of negative polarity items, unavailability of finite complementation and, furthermore, they constitute single event descriptions with their embedded context, i.e. their configurations exhibit temporal uniformity. Using this set of diagnostics, I showed that in Croatian  $V_{[\text{mod/asp/impl}]_{\text{finite}}} + V_{\text{non-finite}}$  configurations behave like complex predicates or monoclausal structures, thus confirming the hypothesis laid out in 4.2. In other words, according to the accounts we followed, a restructuring verb participates in monoclausal raising structures, which seems to be confirmed by Croatian data. However, we do not exclude the possibility that there

---

<sup>54</sup> It is worth noting that already Stowell (1982) suggested that temporal deixis involves the presence of a CP domain.

are other restructuring verbs, e.g. motion verbs, that were not taken into consideration in this paper.

In the course of this exploratory study of monoclausal configurations, other aspects worth of further investigation emerged. A more fine-grained semantic classification of both predicates that select infinitives and their complements – non-finite and finite – is needed. Studies dedicated to particular classes of Croatian predicates with infinitives (e.g. on implicatives, factives, etc.) could also result in valuable insights. Furthermore, a detailed morphosyntactic analysis of these configurations – considering tense, aspect, and mood (TAM) dependencies – might shed further light on the phenomenon. In addition, gathering data on this phenomenon via a corpus-based analysis or an experimental methodology would be beneficial to overcome the possible shortcomings of an intuition-based study such as this one. Finally, cross-linguistic studies focused on observing the phenomenon in a wider context of subordination, such as Wurmbrand *et al.* (2020), Kovač *et al.* (2022), or Wurmbrand & Lohninger (2023), offer a promising line of research.

## Appendix

### Acknowledgments

This work has been fully supported by the Croatian Science Foundation under the projects *Syntactic and Semantic Analysis of Arguments and Adjuncts in Croatian – SARGADA* (IP-2019-04-7896) and *Semantic-Syntactic Classification of Croatian Verbs – SEMTACTIC* (IP-2022-10-8074).

I am grateful both to two anonymous reviewers of the abstract for the conference and to the reviewers of an earlier version of this paper for their valuable comments.

### Abbreviations

3	third person
ACC	accusative
ACT	active
asp	aspectual
DAT	dative
EC	exhaustive control
F	functional head

impl	implicative
INF	infinitive
ipfv	imperfective
M	masculine
mod	modal
N	neuter
NEG	negation
NOM	nominative
PC	partial control
pfv	perfective
PRO	null subject pronoun in non-finite clauses
PRS	present
PTCP	participle
SG	singular
V	verb

## References

- Aissen, J. & Perlmutter, D. M. (1976). Clause Reduction in Spanish. In: Thompson, H., Whistler, K., Edge, V., Jaeger, J. J., Javkin, R., Petruck, M., Smeall, C. & Van Valin, R. D. (eds) *Proceedings of the Second Annual Meeting of the Berkeley Linguistics Society*, 1–30. Berkeley Linguistics Society.
- Aljović, N. (2005). On clitic climbing in Bosnian/Croatian/Serbian. *Forum Bosnae*, 34, 58–84.
- Batinić Angster, M. (2019). Može li hrvatska anafora svoj biti dalekometna? Odgovor na temelju analize infinitivnih dopuna. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje*, 45(2), 279–297.
- Bickel, B., Comrie, B. & Haspelmath, M. (2015). *Leipzig Glossing Rules*. Available online: <https://www.eva.mpg.de/lingua/resources/glossing-rules.php>, Accessed on 2023-12-06).
- Bošković, Ž. (2001). *On the nature of the syntax-phonology interface: Cliticization and related phenomena*. Brill.
- Cardinaletti, A. & Shlonsky, U. (2004). Clitic positions and restructuring in Italian. *Linguistic Inquiry*, 35(4), 519–557.
- Cinque, G. (1999). *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.
- Cinque, G. (2004). Restructuring and functional structure. In: Belletti, A. (ed) *Structures and Beyond: The Cartography of Syntactic Structures*, 132–191. Oxford: Oxford University Press.
- Cinque, G. (2006). *Restructuring and functional heads: the cartography of syntactic structures. Volume 4*. Oxford: Oxford University Press.
- Davies, W. D. & Dubinsky, S. (2008). *The grammar of raising and control: A course in syntactic argumentation*. Malden/Oxford: Blackwell Publishing.

- Gnjatović, T. & Matasović, R. (2013). Some observations on verbs with obligatory control in Croatian. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje*, 39(2), 405–422.
- Gračanin-Yukseš, M. (2019). Clitics that don't climb. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje*, 45(2), 401–416.
- Grano, T. (2015). *Control and restructuring*. Oxford: Oxford University Press.
- Hansen, B., Kolaković, Z. & Jurkiewicz-Rohrbacher, E. (2018). Clitic climbing and stacked infinitives in Bosnian, Croatian and Serbian – A corpus-driven study. In: Fuß, E., Konopka, M., Trawiński, B. & Waßner, U. H. (eds), *Grammar and Corpora*, 259–268. Heidelberg: Heidelberg University Publishing.
- Jurkiewicz-Rohrbacher, E., Hansen, B. & Kolaković, Z. (2017). Clitic climbing, finiteness and the raising-control distinction: A Corpus-based study. *Journal of Linguistics/Jazykovedný casopis*, 68(2), 179–190.
- Karttunen, L. (1971). Implicative verbs. *Language*, 47(2), 340–358.
- Kiparsky, P. & Kiparsky, C. (1970). Fact. In: Bierwisch, M. & Heidolph, K. E. (eds) *Progress in linguistics*, 143–173. Mouton: The Hague.
- Kolaković, Z., Jurkiewicz-Rohrbacher, E. & Hansen, B. (2019). Clitic climbing, the raising-control dichotomy and diaphasic variation in Croatian. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje*, 45(2), 505–522.
- Kolaković, Z., Jurkiewicz-Rohrbacher, E., Hansen, B., Filipović Đurđević, D. & Fritz, N. (2022). *Clitics in the wild: Empirical studies on the microvariation of the pronominal, reflexive and verbal clitics in Bosnian, Croatian and Serbian*. Language Science Press.
- Kovač, I., Lohninger, M. & Wurmbrand, S. (2022). The many facets of restructuring. Hand-out of a talk given at the 5th Virtual NYI, July 2022.
- Krapova, I. & Cinque, G. (2018). Universal Constraints on Balkanisms. A Case Study: The absence of Clitic Climbing. In: Krapova, I. & Joseph, B. (eds) *Balkan Syntax and (Universal) Principles of Grammar*, 151–191. Berlin/Boston: De Gruyter Mouton.
- Krapova, I. & Sočanac, T. (2021). Factivity in South Slavic Languages: Complement and Relative Clauses. In: Franks, S. L., Timberlake, A. H. & Wietecha, A. W. (eds) *Selected Proceedings of the 14<sup>th</sup> Meeting of the Slavic Linguistics Society*, 89–111. Berlin: Peter Lang.
- Laka Mugarza, M. I. (1990). *Negation in syntax – on the nature of functional categories and projections*. PhD thesis, Massachusetts Institute of Technology.
- Landau, I. (2000). *Elements of control: Structure and meaning in infinitival constructions*. Dordrecht: Kluwer.
- Marušič, F. & Žaucer, R. (2006). On the intensional FEEL-LIKE construction in Slovenian: A case of a phonologically null verb. *Natural Language & Linguistic Theory*, 24(4), 1093–1159.
- McCawley, J. D. (1979). *Adverbs, vowels, and other objects of wonder*. Chicago: University of Chicago Press.
- Mihaljević, M. (2009). The Structure of Complements of Verbs of Perception in Croatian. In: Franks, S., Chidambaram, V. & Joseph, B. (eds) *A Linguist's Linguist: Studies in South Slavic Linguistics in Honor of E. Wayles Browne*, 317–353. Bloomington: Slavica Publishers.
- Pearson, H. A. (2012). *The Sense of Self: Topics in the Semantics of De Se Expressions*. PhD thesis, Harvard University.
- Pearson, H. A. (2016). The semantics of partial control. *Natural Language & Linguistic Theory*, 34(2), 691–738.

- Progovac, L. (1993). Locality and subjunctive-like complements in Serbo-Croatian. *Journal of Slavic linguistics*, 1(1), 116–144.
- Progovac, L. (1996). Clitics in Serbian/Croatian: Comp as the second position. In: Halpern, A. L. & Zwicky, A. M. (eds), *Approaching second: Second position clitics and related phenomena*, 411–428. Stanford: CSLI Publications
- Pitteroff, M., Alexiadou, A., Darby, J. & Fischer, S. (2017). On partial control in German. *The Journal of Comparative Germanic Linguistics*, 20, 139–185.
- Rizzi, L. (1976). Ristrutturazione. *Rivista di grammatica generativa*, 1, 1–54.
- Rizzi, L. (1978). A restructuring rule in Italian syntax. In: Keyser, S. J. (ed) *Recent Transformational Studies in European Languages*, 113–158. Cambridge (MA): MIT Press.
- Roussou, A. (2009). In the mood for control. *Lingua*, 119, 1811–1836.
- Stjepanović, S. (2004). Clitic Climbing and Restructuring with “Finite Clause” and Infinitive Complements. *Journal of Slavic Linguistics*, 12, 177–212.
- Stowell, T. (1982). The tense of infinitives. *Linguistic Inquiry*, 13, 561–70
- Todorović, N. & Wurmbrand, S. (2020). Finiteness across domains. In: Radeva-Bork, T. & Kosta, P. (eds) *Current Developments in Slavic Linguistics. Twenty Years After (based on selected papers from FDSL 11)*, 47–66. Berlin: Peter Lang.
- Veselinović, D. (2017). Structural Differences Between Epistemic and Root Modality: Evidence from BCS. In: Kantarovich, J., Tuong, T. & Xherija, O. (eds) *Proceedings of CLS52 (2016)*, 501–515. Chicago: Chicago Linguistics Society.
- Wurmbrand, S. (1998). *Infinitives*. PhD thesis, Massachusetts Institute of Technology.
- Wurmbrand, S. (2003). *Infinitives: Restructuring and Clause Structure*. Berlin/New York: Mouton de Gruyter.
- Wurmbrand, S. (2004). Two types of restructuring – Lexical vs. functional. *Lingua*, 114(8), 991–1014.
- Wurmbrand, S. (2016). Restructuring as the regulator of clause size. Talk given at the Workshop on Shrinking Trees, 10 October 2016, University of Leipzig.
- Wurmbrand, S. & Lohninger, M. (2023). An implicational universal in complementation: Theoretical insights and empirical progress. In: Hartmann, J. M. & Wöllstein, A. (eds) *Propositionale Argumente im Sprachvergleich: Theorie und Empirie./Propositional Arguments in Cross-Linguistic Research: Theoretical and Empirical Issues*, 183–231. Tübingen: Gunter Narr Verlag.
- Wurmbrand, S., Kovač, I., Lohninger, M., Pajančić, C. & Todorović, N. (2020). Finiteness in South Slavic complement clauses: Evidence for an implicational finiteness universal. *Linguistica*, 60(1), 119–137.
- Zanutini, R. (1991). *Syntactic properties of sentential negation. A comparative study of Romance languages*. PhD thesis, University of Pennsylvania.
- Zovko Dinković, I. (2013). *Negacija u jeziku: kontrastivna analiza negacije u engleskome i hrvatskome jeziku*. Zagreb: Hrvatska sveučilišna naklada.
- Zovko Dinković, I. (2021). *O dvostrukoj negaciji i niječnome slaganju*. Zagreb: Hrvatska sveučilišna naklada.
- Zovko Dinković, I. & Ilc, G. (2017). Pleonastic negation from a cross-linguistic perspective. *Jezikoslovlje*, 18(1), 159–180.

## 5. Croatian middle *se*-constructions

BRANIMIR BELAJ

Keywords: *middles; reflexives; passives; impersonals; transitivity; intransitivity*

This paper addresses Croatian middle *se*-constructions, viz. middle verbs, *se*-passives, and impersonal *se*-constructions from the joint perspective of Cognitive Grammar and some work of typological-functional provenance. The Croatian grammatical literature virtually never discusses these constructions under the label “middles”. In this paper we argue that all Croatian *se*-constructions, be they morphological (middle verbs) or syntactic (middles, passives, impersonals) have the basic semantic and pragmatic properties of middles. These properties can be subsumed under a single schematic meaning, viz. *low degree of event structure elaboration*, which results from detopicalizing event participants and topicalizing the event itself. From this follows our strong conviction that the term *povratni glagoli* ‘reflexive verbs’ must be replaced with the term *medijalni glagoli* ‘middle verbs’, and that the notion and label “reflexivity” should only remain applicable to reflexive transitive constructions, which differ profoundly from their middle counterparts semantically and syntactically.

*Original scientific paper*

### 1. Introduction\*

Upon reviewing Croatian reference grammars and other scholarly publications, one can conclude, without a hint of reservation, that middles, or the middle voice, are among the most undervalued and overlooked morphological and syntactic categories in the description of the Croatian language. Middles do not figure at all in most of our descriptive and prescriptive grammars (e.g. Pavešić & Vince 1971; Babić *et al.* 1991; Katičić 1991; Raguž 1997; Silić & Pranjković 2005). Even where they do make an appearance (e.g. Barić *et al.* 1995: 231), middles are handled with insufficient breadth and depth. At best, sections devoted to morphology mention middle verbs as one subtype of so-called pseudo-reflexive verbs. They are said to designate various natural occurrences (e.g. *mračiti se* ‘to grow dark’,

\* A more concise Croatian version of this paper is published in Belaj (2023).

*naoblačiti se* ‘to turn cloudy’, *bjelasati se* ‘to shimmer/glimmer white’) or processes taking place within humans (e.g. *sramiti se* ‘to feel ashamed, shy, bashful’, *čuditi se* ‘to wonder, to feel puzzled’, *nadati se* ‘to hope’). However, as a syntactic category, middles are nowhere to be found in Croatian grammars (although Belaj & Tanacković Faletar 2017: 292 mention the category). As far as other types of publications are concerned, middle constructions<sup>55</sup> are addressed in e.g. Oraić Rabušić (2018: 175–177) and Batinić Angster (2020: 64–65). However, they are limited to only generic *se*-structures exemplified by *Riža se kuha petnaest minuta* ‘Rice is cooked for fifteen minutes’, *Vrtovi se okopavaju motikama* ‘Gardens are tilled with hoes’, *Trava se zalijeva rano ujutro i uvečer* ‘Grass is watered in the early morning and evening’, etc. These generic structures are described as qualifying the designated process by expressing how it is typically performed. The neglect of the middle voice evidenced in Croatian linguistics and beyond should probably be attributed to the absence of a clear and unequivocal syntactic definition of the category, comparable to that available for the passive. And the reasons why such a definition is still missing may be due to the existence, cross-linguistically, of a range of functionally similar but far less controversial constructions. The same meanings that could be carried by middle-voice constructions find their expression, cross-linguistically, in a range of semantically more coherent and transparent morphological and syntactic constructions known under the much more current and universally well-established labels like the passive, impersonal constructions, reflexive verbs, etc.

«Since the term applies to diverse constructions, and in some cases to elaborate families of constructions, any single characterization is bound to be oversimplified». (Langacker 2008: 385)

One of the first attempts to define middles speaks to its cross-categorial and schematic character. Lyons (1969: 373) describes middles broadly as the voice where «the ‘action’ or ‘state’ affects the subject of the verb or his interests», which allows other morphosyntactic categories, foremost passives and other intransitive constructions, to be classed as middles.<sup>56</sup>

<sup>55</sup> In line with cognitive constructional approaches to grammar we use the term *construction* in reference to any linguistic unit above the level of the morpheme.

<sup>56</sup> For instance, in their grammar of Bosnian, Jahić, Halilović & Palić (2000: 262) list middles as the third voice joining the usual active and passive. The middle is argued to express processes taking place without the subject’s volition. There is no mention of *se*-constructions as middles; the category is exemplified with constructions like *Ujesen lišće vene* ‘Leaves wither in the fall’, *Doći će ljepše stoljeće* ‘A better century will come’ and *Dedo je ostario* ‘Grandpa has grown old’.

Middles are made no less slippery by the label they carry. The designation ‘middle’ or ‘middle voice’ pegs them morphosyntactically as a very divergent, in-between category, belonging neither among transitive two-participant constructions nor among prototypical intransitive one-participant constructions. Yet, semantically and by extension, syntactically, middles are intimately connected to various reflexive, passive and impersonal constructions. In other words, they are extremely polysemous. Leaning on the relevant insights from the cognitive linguistic scholarship of the last few decades, this paper will argue for a somewhat broad conception of middles. From this perspective, Croatian will turn out to be a language with one of the most articulated systems of middle constructions, morphological and syntactic. All Croatian *se*-constructions will be shown to fall out from the schematic meaning of middles, viz. *low degree of event structure elaboration*. Their schematic counterpart for two-participant transitive constructions and constructions with the reflexive pronoun *sebe* ‘self’ is *high degree of event structure elaboration*. Both obviously derive from the semantic parameter of *relative elaboration of events* (Kemmer 1993: 3).<sup>57</sup> Based on this parameter, we may distinguish between middle and reflexive constructions. At the same time, as far as their position on the transitivity scale, both middles and reflexives land somewhere between its two ends marked by prototypical transitive two-participant (two-argument) constructions and prototypical intransitive one-participant (one-argument) constructions. The schematic meaning of low degree of event elaboration, common to all situation types coded by middles, follows from the *low degree of distinguishability of participants*; the lower the participant distinguishability, the lower the degree of event elaboration, and vice versa. In the case of passive and, especially impersonal constructions, their low degree of event elaboration is due to reduced event structure – the latter being a consequence of detopicalizing one or all participants and topicalizing the process itself.<sup>58</sup> Therefore, the degree of participant reduction is proportional to the degree of process topicalization,

<sup>57</sup> Lichtenberk’s (1991: 176) term *the degree of elaboration of situations* assumes that every situation can be linguistically coded in alternative ways, with more or less detail. It corresponds to the Cognitive Grammar notion of *construal*. See also Buttinelli (*this volume*, Chapter 11, section 2) for the distinction between subjective and objective construal in the context of the continuum between arguments and adjuncts.

<sup>58</sup> *Participant distinguishability* corresponds to what Langacker & Munro (1975) and Langacker (1976) call *non-distinctness of arguments*, Timberlake (1975, 1977) and Hopper & Thompson (1980) refer to as *individuation of participants*, and Lichtenberk (1991) terms *the degree of distinctness of participants*.

viz. processualization, while the low degree of event structure elaboration is ultimately a consequence of participant reduction. The remainder of this study elaborates on these semantic parameters, explaining how they bear on the analysis of middle verbs (2.1.) as well as middle, passive, and impersonal syntactic constructions (2.2.).

## 2. Discussion

This section examines in more detail single-participant<sup>59</sup> and multi-participant middle verbs (2.1.),<sup>60</sup> and syntactic middle *se*-constructions, viz. the prototypical middle, passive, and impersonal constructions (2.2.).

<sup>59</sup> The term *single-participant* is used in reference to single-participant middle verbs as opposed to multi-participant reciprocal middle verbs. The term *one-participant* is used to refer to intransitive verbs and absolute constructions/thematic processes featuring a single participant.

<sup>60</sup> A finer classification and semantic elaboration of middle verbs is outside the scope of this paper. A few remarks are in order though. Middle verbs fork into two main classes: (i) *single-participant* and (ii) *multi-participant*. Single-participant middle verbs designate processes directed at or taking place within one subject – agent or experiencer (singular or plural). They include (a) prototypical middle verbs, which have their counterparts in direct or indirect transitive reflexive constructions (e.g. *kupati se / kupati sebe* ‘bathe / bathe oneself’, *brijati se / brijati sebe* ‘shave / shave oneself’, *gledati se / gledati sebe* ‘look at oneself’, *sagraditi si / sagraditi sebi* ‘to build for oneself’, *pokloniti si / pokloniti sebi* ‘to gift oneself’, etc.); (b) middle verbs with no counterparts in reflexive constructions but with counterparts in the form of prototypical transitive verbs (e.g. *buditi se / buditi* ‘to wake up / to wake up Obj’, *pomicati se / pomicati* ‘to move / to move Obj’, *micati se / micati* ‘to move / to move Obj’ *osjećati se / osjećati* ‘to feel Adj/ to feel Obj’, *spuštati se / spuštati* ‘to lower oneself / to lower Obj’, *dizati se / dizati* ‘to lift oneself / to lift Obj’, *zapustiti se / zapustiti* ‘to let oneself go, to neglect oneself / to neglect Obj’, etc.); (c) redundant middle verbs, which have similar meanings with or without the particle *se*, i.e. in their middle and prototypically intransitive shape (e.g. *klizati se / klizati* ‘to skate’, *skijati se / skijati* ‘to ski’, *blistati se / blistati* ‘to glisten’, *šetati se / šetati* ‘to stroll’, etc.) and (d) middle verbs found only with the particle *se*, so-called *reflexiva tantum* verbs (e.g. *kajati se* ‘to regret’, *smijati se* ‘to laugh’, *bojati se* ‘to fear’, *svadati se* ‘to quarrel’, etc.). Multi-participant middle verbs designate processes which two or more subjects (agents or experiencers) direct at each other. These are reciprocal middle verbs, with counterparts in explicit constructions featuring the phrase *jedan drugog* ‘each other, one another’ in various gender, case and number forms (e.g. *grudali su se / grudali su jedni druge* ‘they were in a snowball fight / they were in a snowball fight with each other’, *svadali su se / svadali su se jedni s drugima* ‘they argued / they argued with each other’, *poljubili su se / poljubili su jedno drugo* ‘they kissed / they kissed each other’, etc.). Except for Silić & Pranjković (2005), Croatian reference grammars call single-participant middle verbs from class (a) and multi-participant reciprocal verbs *real reciprocal verbs*, while single-participant verb classes (b), (c) and (d) are tagged as *pseudo-reflexive verbs*.

## 2.1. Middle verbs

### 2.1.1. Single-participant middle verbs

Three features are common to all accounts of *se*-constructions in the Croatian grammatical literature: (i) regardless of theoretical and methodological bias, no use is ever made of the term “middle verb”; without exception these verbs are called reflexive verbs; (ii) the prevailing view is that, with all reflexive verbs, *se* is the enclitic form of the reflexive pronoun *sebe* ‘self’ (Barić *et al.* 1995: 231–232; Raguž 1997: 163; Pavešić & Vince 1971: 374–375), even in passive and impersonal constructions (Katičić 1991: 143–146; Barić *et al.* 1995: 452); and (iii) in verbs where *se* cannot be replaced by *sebe*, *se* is categorized as a particle; where the substitution is an option, *se* is categorized as an accusative-marked object clitic (Ivić 1962, 1967; Browne 1974; Belaj 2001; Silić & Pranjković 2005: 40–41).<sup>61</sup> To the best of our knowledge, only Babić *et al.* (1991: 675–676) and Oraić Rabušić (2018) treat *se* uniformly as a particle, whereby Babić *et al.* (1991: 675) claim that with so-called real reflexive verbs *se* is substitutable by *sebe*. Examples (1a–d) suffice to show that *se* cannot always be replaced with *sebe* and that it is not an enclitic accusative form of the reflexive pronoun, therefore not an object. Instead, it should be treated as a particle clitic even with verbs where it formally alternates with the reflexive pronoun.<sup>62</sup>

- (1) a. *Marija se<sub>MM</sub> počešljala.*  
 ‘Mary combed her hair’  
 b. \**Marija je počešljala sebe<sub>RM</sub>.*<sup>63</sup>  
 ‘Mary combed her hair.’  
 c. *Marija je prvo počešljala sebe<sub>RM</sub>, a onda dijete.*  
 ‘Mary first combed her hair, and then the child’s.’  
 d. \**Marija se<sub>MM</sub> počešljala, a onda dijete.*  
 ‘Mary combed her hair, and then the child’s.’

<sup>61</sup> In the remainder of the paper the enclitic form *se* will be referred to as the *light form*, while the full form *sebe* will be referred to as the *heavy form*.

<sup>62</sup> The idiomatic English translations of Croatian examples often do not reflect the distinctions between Croatian middle and reflexive constructions. For this reason, in section 2.1. we will provide minimum glossing, using the gloss MM for the middle marker *se* and RM for the reflexive marker *sebe*. We will also gloss any case markings that prove instrumental to the middle-reflexive contrast.

<sup>63</sup> Putting semantic-pragmatic criteria aside, this sentence is acceptable; otherwise, it is not. Without a supporting co-text and context no native speaker of Croatian would ever use such a construction.

When it comes to proving that the mentioned constructions feature pronouns, we might be inclined to rely on a comparison with similar constructions featuring personal pronouns – on the assumption that the distributional restrictions on the heavy (HF) and light (LF) form of personal pronouns were identical. Compare (2a–d):

- (2) a. *Marija ga<sub>LF</sub> je počesljala.*  
 ‘Mary combed his hair.’  
 b. \**Marija je počesljala njega<sub>HF</sub>.*  
 ‘Mary combed his hair.’  
 c. *Marija je prvo počesljala njega<sub>HF</sub>, a onda dijete.*  
 ‘Mary first combed his hair, and then the child’s.’  
 d. \**Marija ga<sub>LF</sub> je počesljala, a onda dijete.*  
 ‘Mary combed his hair, and then the child’s.’

However, this reasoning is flawed – primarily because the use of light vs. heavy forms of personal pronouns has nothing to do with event structure reduction caused by low participant distinguishability. The use of heavy forms is motivated only by the pragmatic context, i.e. the need to emphasize the participant, which then requires that a relationship or contrast with another entity also be established. Therefore, the light and heavy forms of personal pronouns do not differ in kind – categorically, they are both pronouns. On the other hand, reaching for a light form of the reflexive pronoun *sebe* is iconically motivated; the light form results from a reduced distinguishability of event participants. This makes the light form *se* an integral part of the verb, and thus also a particle, not a pronoun.

Several syntactic tests well-known from the generative literature support this view. Essentially the *sebe/se* alternation is compared to that between the heavy (*njega, njih, mene*) and light (*ga, ih, me*) accusative form of personal pronouns, showing that the forms of personal pronouns are in free distribution, whereas those of the reflexive pronoun are not. These are the so-called secondary predication test, shown in (3) (Oliva 2001: 203; Oraić Rabušić 2018: 66), the test of modification with the *kao*-phrase in (4) (Moskovičević 1997: 111; Oraić Rabušić 2018: 67–68) and the comparative construction test, shown in (5) (Zec 1985: 365; Oraić Rabušić 2018: 71). These tests show that in the case of personal pronouns both forms are coreferential with the accusative of the secondary predicate and of the nominal in the *kao*-phrase and the comparative construction. As for the reflexive pronoun, only the heavy form is coreferential with the accusative while the light form binds the nominative, cf. (3–5):

- (3) a. *Petar je njega<sub>ACC</sub> cijeloga<sub>ACC</sub> uredio.*  
 ‘Peter cleaned all of him up.’  
 b. *Petar ga<sub>ACC</sub> je cijeloga<sub>ACC</sub> uredio.*  
 ‘Peter cleaned all of him up.’  
 c. *Petar je sebe<sub>RM-ACC</sub> cijeloga<sub>ACC</sub> uredio.*  
 ‘Peter cleaned himself all up.’  
 d. *Petar se<sub>MM</sub> cijeli<sub>NOM</sub> / \*cijeloga<sub>ACC</sub> uredio.*  
 ‘Peter cleaned himself all up.’
- (4) a. *Njih<sub>ACC</sub> su predstavili kao dobre studente<sub>ACC</sub>.*  
 ‘They presented them as good students.’  
 b. *Predstavili su ih<sub>ACC</sub> kao dobre studente<sub>ACC</sub>.*  
 ‘They presented them as good students.’  
 c. *Sebe<sub>RM-ACC</sub> su predstavili kao dobre studente<sub>ACC</sub>.*  
 ‘They presented themselves as good students.’  
 d. *Predstavili su se<sub>MM</sub> kao dobri studenti<sub>NOM</sub> / \*dobre studente<sub>ACC</sub>.*  
 ‘They presented themselves as good students.’
- (5) a. *Petar je branio mene<sub>ACC</sub> uspješnije nego Anu<sub>ACC</sub>.*  
 ‘Peter defended me more successfully than (he defended) Anne.’  
 b. *Petar me<sub>ACC</sub> je branio uspješnije nego Anu<sub>ACC</sub>.*  
 ‘Peter defended me more successfully than (he defended) Anne.’  
 c. *Petar je branio sebe<sub>RM-ACC</sub> uspješnije nego Anu<sub>ACC</sub>.*  
 ‘Peter defended himself more successfully than (he defended) Anne.’  
 d. *Petar se<sub>MM</sub> branio uspješnije nego Ana<sub>NOM</sub> / \*Anu<sub>ACC</sub>.*  
 ‘Peter defended himself better than Anne (defended herself) / better than (he defended) Anne.’

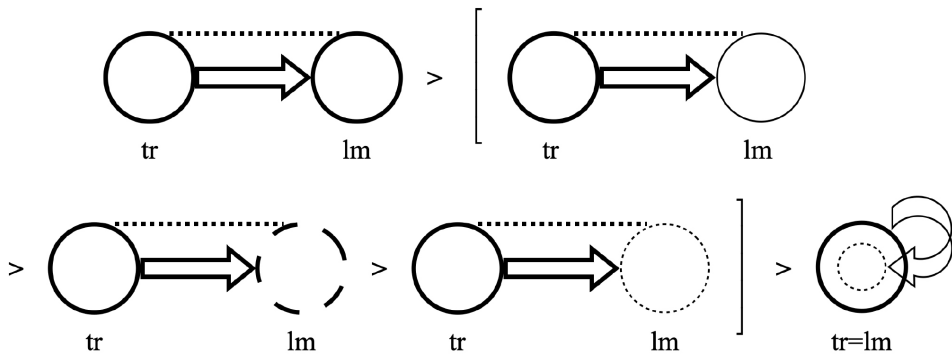
The examples above prove that the heavy and the light form of the reflexive pronoun are not syntactically equivalent. *Se*-verb constructions show a markedly different syntactic behaviour in comparison with constructions featuring reflexive pronouns, suggesting that they also symbolize different conceptual/semantic configurations. Syntactically, the light form can by no means or criteria be considered a reflexive pronoun. Ergo, the label *reflexive verb* is a poor fit for what should in fact be termed *middle verb*. To be sure, the label *reflexive verb* is not inaccurate. It is even more common in the pertinent literature (e.g. Hopper & Thompson 1980; Haiman 1983; König 2001; König & Siemund 2000; Smith 2004, etc.). Its downside, however, is that it is motivated exclusively by morphological facts. This is inconsistent with the continuous view of grammar as spanning morphology,

syntax, and semantics – the only view, to our mind, which guarantees a proper explanation of how structure and function meet in the grammatical constructions of any language. The light form *se* originated in the reflexive pronoun, but their diachronic morphological relationship has been weakened through grammaticalization. Syntactically, the link between the two has all but disappeared as the pronominal function of the light form changed to that of a particle via the following grammaticalization chain [<sub>MORPH.</sub> *se*<sub>PRON</sub> > *se*<sub>PRON</sub>] > [<sub>SYNT.</sub> *se*<sub>PART.</sub>]. The particle function of the Croatian middle marker *se* is semantically consistent with comparable grammaticalization processes taking place in other languages where an etymological link exists between the primary reflexive marker and the middle marker.<sup>64</sup> In other words *se* was subject to *desemanticization*, semantic bleaching, depletion (Heine & Reh 1984; Lehmann 1985) or loss (Traugott & König 1991) of the primary form's semantic function – a grammaticalization tendency usually accompanied by the loss of phonological, morphological, and syntactic autonomy. The Croatian middle marker *se* is thus a component of every middle verb, i.e. all middle verbs are special lexical units. This grammaticalization path also explains why it is not possible for a language to have middle markers without also having reflexive markers (Kemmer 1993: 241), while the opposite holds, viz. languages can have reflexive markers without having middle markers – as evidenced by English. Croatian light form *se* is a particle because, like all other particles, its function is modificational. It modifies the verb-designated

<sup>64</sup> Middle markers are attested in relatively few languages. Among languages that have them, there are so-called two-form languages, where the reflexive-middle contrast is coded by two forms (Haiman 1983 and Smith 2004 use the terms *heavy* and *light form*), like Russian, Croatian, Icelandic, Romanian, Hungarian and Turkish. There are also languages with one only form, the so-called one-form languages like French, Spanish and German. Languages also differ in the morphological means available for coding middle markers. Clitics or unbound morphemes are used in Croatian, French, Spanish (Maldonado 1992, 1999, 2000, 2008) and Romanian (Calude 2017), whereas other languages use verb affixes – for instance *-mai* in Classical Greek, *-st* in Icelandic, *-sja* in Russian, *-kod-* and *-köz-* in Hungarian. Finally, in some languages, the reflexive and the middle marker are morphologically related. For instance, in Croatian, Czech and Russian the middle developed from the primary reflexive. In others, like Romanian, Turkish, Latin, modern Greek there is not even so much as a semantic connection (Manney 2001: 22). Other languages missing a morphological connection between middles and reflexives include some South-American and Mexican languages (Maldonado 2009). Maldonado (2009) also argues that in Spanish there are many cases suggesting that the middle was not derived from the reflexive.

action by limiting the extent of event structure elaboration via reducing the degree of conceptual distinguishability of its participants.<sup>65</sup> This semantic function makes *se* different from the reflexive pronoun *sebe*, whose meaning has the opposite effect. The relationship between the reflexive pronoun *sebe* and the particle clitic *se* may be compared to what happens in photography when we zoom in or out by adjusting focal length in order to change the angle of view or magnification. Compare Figure 1, where dotted lines mark correspondences between the same entity functioning as the trajector (agent/subject) and the landmark (theme<sup>66</sup> / object).

Figure 1. Metaphorical portrayal of the reduction of participant distinguishability in transitive reflexive constructions and middle verbs



Various intermediate stages are normally not accessible to processing, which is why in Figure 1 they are enclosed in brackets. Rarely, they may gain some cognitive salience and thus become accessible, cf. the middle verb *okupati se* ‘take a bath/bathe oneself; take a swim’ in (6a–b):

- (6) a. *Ivan se<sub>MM</sub> ujutro okupao i otišao na posao.*  
 ‘Ivan took a bath in the morning and went to work.’

<sup>65</sup> This is why generativists have traditionally treated such particles and other grammaticalized reflexive forms as *detransitivizers* (e.g. Rosen 1989; Grimshaw 1982; Zec 1985; Moskovljević 1997). Interestingly, Maldonado (2008) argues the opposite for Spanish, viz. a middle marker is found to increase transitivity.

<sup>66</sup> In accordance with Cognitive Grammar, the label *theme* is used in reference to a macrorole, whose prototypical microrole is patient. It corresponds to the notion of *undergoer* in Role and Reference Grammar (Foley & Van Valin 1984; Van Valin & LaPolla 1997).

- b. *Na putu smo zastali u jednoj uvali, **okupali se**<sub>MM</sub> i nastavili dalje.*  
 ‘We stopped along the way in a lagoon, took a dip and moved on.’

*Okupati se* in (6a) is a verb denoting a body care routine. This type of verbs assumes that particular parts of the body are affected by the process which the agent initiates and directs at himself. The same verb in (6b) does not imply involvement of particular body parts. It simply denotes the process of immersing the body in the water and moving the limbs to keep oneself afloat. This example suggests that it may be profitable to think of some middle verbs as scalar in terms of the degree to which they elaborate the structure of the designated event.

Since in Croatian the reflexive-middle contrast between *sebe/se* is very pronounced, we will next turn to exploring its syntactic and semantic ramifications.

The basic difference between reflexive *sebe-* and middle *se-* constructions is that the former are only valid under specific semantic and pragmatic conditions. They depend on the existence of a relationship or, in the words of Haiman (1983: 804), of a contrast with some other event, as shown above in (1b–c). More examples are given below:

- (7) a. *Pogledaj **se**<sub>MM</sub> kako izgledaš!*  
 ‘Take a look at yourself.’  
 b. \**Pogledaj **sebe**<sub>RM</sub> kako izgledaš!*  
 ‘Take a look at yourself.’  
 c. *Pogledaj prvo **sebe**<sub>RM</sub> a onda sudi drugima!*  
 ‘First take a look at yourself, and then judge others.’  
 d. \**Pogledaj **se**<sub>MM</sub> a onda sudi drugima!*  
 ‘First take a look at yourself, and then judge others.’
- (8) a. *Zašto **se**<sub>MM</sub> stalno gledaš u ogledalu?*  
 ‘Why do you keep looking at yourself in the mirror?’  
 b. \**Zašto **sebe**<sub>RM</sub> stalno gledaš u ogledalu?*  
 ‘Why do you keep looking at yourself in the mirror?’  
 c. *Gledaj **sebe**<sub>RM</sub> baš te briga za druge.*  
 ‘Mind yourself, don’t concern yourself with others.’  
 d. \**Gledaj **se**<sub>MM</sub> baš te briga za druge.*  
 ‘Mind yourself, don’t concern yourself with others.’

Several factors explain why the acceptability of reflexive transitive constructions with an emphatic reflexive pronoun, as they are commonly dubbed, depends on the existence of a relationship with other entities or events (e.g. Ivić 1983: 118; Haiman 1983; Calude 2017). The first is the contrast between expected and unexpected information, i.e. between informational predictability and unpredictability. Haiman (1983: 802) associates informational predictability/unpredictability with language economy and shows that with reflexive constructions, among others, the predictable information is preferably coded by shorter, phonologically lighter forms. Longer forms are associated with unpredictable information.<sup>67</sup> Using a frequentist argument, where higher frequency is seen as proportional with predictability, Haspelmath (2008) argues that more frequent scenarios will be coded with less linguistic material and vice versa. In his analysis of reflexive constructions, Haiman sets up another distinction, viz. between so-called *introverted* and *extroverted* verbs.<sup>68</sup> Introverted verbs are self-directed, i.e. the activity designated by such verbs is directed at the subject (for instance body care verbs, like *kupati se<sub>MM</sub>* ‘bathe (oneself)’, *brijati se<sub>MM</sub>* ‘shave (oneself)’, *češljati se<sub>MM</sub>* ‘comb (oneself)’, etc.). Extroverted verbs are other-directed, e.g. *voljeti* ‘love’, *mrziti* ‘hate’, *udarati* ‘hit’, *gledati* ‘watch’, etc. In view of these two factors, verbs of body care clearly must take the light form in neutral contexts. Since they are prototypically introverted verbs, the information that the action is spontaneous and subject-directed is expected and the verbs receive lighter coding. If, in turn, these verbs are used with the heavy form, viz. the reflexive pronoun *sebe*, the notion of unexpectedness follows automatically. The situation assumes the air of unexpectedness as it now communicates a non-spontaneous, voluntary action directed at the agent’s body but begs for “additional explanation”. In examples (1a–d), for instance, this explanation involves reference to a relationship with another animate entity. *Pogledati* ‘take a look’ and *gledati* ‘look at, mind’ (7–8) in turn are extroverted verbs – where the other-directedness of action is the expected information. When accompanied by the heavy form *sebe*, these verbs

<sup>67</sup> Haiman (1983: 802) draws an interesting analogy between the relationship of predictable/unpredictable contrast to lighter/heavier linguistic coding, and not spelling out the country of destination in national mail traffic.

<sup>68</sup> Haiman’s (1983) introverted/extroverted contrast was embraced by e.g. Everaert (1986, 1991), Reuland (1998), Huang (2000), König & Siemund (2000), König (2001) – who labels introverted verbs *self-directed* and extroverted verbs *other-directed* – and Smith (2004).

communicate an unexpected information and involve an obligatory contrast with “the other” (7c). Such a contrast is excluded in the case of the light form *se* (7–8d). If, however, these verbs migrate semantically from the domain of extrovertedness toward that of introvertedness, as they do in (7a) and (8a), the situation is interpreted as expected and the light form *se* is sanctioned. In (7a) the imperative form *pogledaj se<sub>MM</sub>* ‘take a look at yourself’ is used as a pragmatic strategy to chastise the interlocutor and make him aware of the consequences of his actions (this is often the case with reprimanding the child for dirtying their clothes during play). Example (8a) is something of an idiomatic expression communicating narcissistic demeanour. It is not so much about directing the gaze to the entity as a whole, as it is about the gaze being directed at specific body parts and details of appearance. Haiman (1983: 803) adds the unexpected/expected quality of the object as another factor in the choice between heavy or light form. This factor contributes to the neutrality/acceptability of (8a) as the expected object of gaze in the mirror is the subject, not someone else. With the extroverted verbs *voljeti* ‘love’ and *mrziti* ‘hate’ (9a–d), other-directedness explains the distribution of the heavy form *sebe* vs. the light form *se*. The light middle form *se* is not acceptable as it directs the action at the subject.<sup>69</sup> The heavy form *sebe*, as usual, implies a contrast, as in (9d):

- (9) a. ??*Mrzio / Volio sam se<sub>MM</sub> cijeli život.*  
 ‘I’ve hated/loved myself all my life.’  
 b. ??*Mrzio / Volio sam sebe<sub>RM</sub> cijeli život.*  
 ‘I’ve hated/loved myself all my life.’  
 c. \**Mrzio / Volio sam se<sub>MM</sub> više nego tebe.*  
 ‘I’ve hated/loved myself more than (I’ve hated/loved) you.’  
 d. *Mrzio / Volio sam više sebe<sub>RM</sub> nego tebe.*  
 ‘I’ve hated/loved myself more than (I’ve hated/loved) you.’

The bulk of so-called *reflexiva tantum* verbs are introverted verbs.<sup>70</sup> In fact, they are highly introverted verbs featuring subjects in the semantic

<sup>69</sup> Some Croatian native speakers may find *se* acceptable with the verbs *mrziti* and *voljeti*. In our native speaker judgement, it is not. In fact, an informal search of the *Riznica* corpus did not return a single such token. Perhaps some examples could be found in other corpora, but what matters more is that with these verbs *se* forms are highly marked, and sit on the very brink of acceptability.

<sup>70</sup> Milković (*this volume*, Chapter 6) offers a diachronic survey of *reflexiva tantum* psychological verbs in the history of Croatian.

role of experiencer. The light form *se* is non-optional because the process takes place subject-internally. In other words, the conceptual configuration underlying these verbs reflects iconically onto the form. Given the extremely low degree of event structure elaboration, these verbs are among the most prototypical middle constructions. Among prototypical middle verbs we also find those that designate natural states and processes. Recall that in some grammars these are the only types of verbs tagged as middles. They include deadjectival verbs like *zelenjeti se<sub>MM</sub>* ‘to be/glow green’, *crvenjeti se<sub>MM</sub>* ‘to be/glow red’, *žutjeti se<sub>MM</sub>* ‘to be/glow yellow’, *bjelasati se<sub>MM</sub>* ‘to be/gleam white’, *crnjeti se<sub>MM</sub>* ‘to be/gleam black’ etc., and designate processes resulting in states and properties coded by the adjectives.

Redundant middle verbs are usually described as having the same meaning with or without the *se* particle, cf. *skijati se<sub>MM</sub>* / *skijati* ‘to ski’, *klizati se<sub>MM</sub>* / *klizati* ‘to skate’, *šetati se<sub>MM</sub>* / *šetati* ‘to walk’, *blistati se<sub>MM</sub>* / *blistati* ‘to glisten’, etc. However, these variants only share a semantic base. Their differences come to the fore especially with verbs denoting various sporting activities (Belaj 2001). The middle form will be more common in non-competitive contexts, while the particle-less form is the only acceptable form in references to competitive, especially professional sports. Compare:

- (10) a. *Na zimovanju je bilo jako lijepo, bilo je puno snijega, pa smo se<sub>MM</sub> skijali / ?skijali od jutra do mraka.*  
 ‘We had such a good time during winter holidays, there was lots of snow, so we skied from morning till night.’
- b. *Janica Kostelić jučer je izvrsno skijala u Wengenu / \*Janica Kostelić jučer se<sub>MM</sub> izvrsno skijala u Wengenu.*  
 ‘Janica Kostelić skied amazingly yesterday in Wengen.’

What is more, in the context of professional sports, forms without the *se* particle may even be transitive:

- (11) a. *Jučer je skijala svoju posljednju utrku.*  
 ‘Yesterday, she skied her last race.’
- b. *Osim slaloma skijat će i veleslalom.*  
 ‘Besides slalom, she will also ski the giant-slalom.’

Shifts between middle and non-middle forms are mainly due to the fact that the middle form blocks transitivity, allowing the action to be construed as an exercise undertaken for the subject’s sheer amusement.

In the domain of professional sports, however, reference is made to other animate participants or to inanimates metaphorically construed as event participants. A similar contrast is found in measure expressions, which can be coded as adverbials or as arguments (participants) with prototypical intransitive verbs like *trčati* ‘run’, *hodati* ‘walk’, *plivati* ‘swim’, etc. (Belaj & Tanacković Faletar 2017: 168), cf. 12:

- (12) a. *Trčao je osamsto metara do kuće.*  
 ‘He ran **eight hundred meters** to the house.’  
 b. *Na Olimpijskim igrama trčao je osamsto metara.*  
 ‘He ran **800-meters** (i.e. the 800-meter race) in the Olympics.’

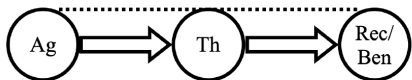
Apart from verbs with the particle *se*, which represents the highest grade of grammaticalization of the light accusative form of the reflexive pronoun, we should also consider indirect middle constructions with *si*, the light dative form of the reflexive pronoun marking a beneficiary/recipient. The *si* form alternates with the heavy form *sebi* in reflexive constructions. Cf. (13):

- (13) a. *Kupio si<sub>MM-DAT</sub> je kuću.*  
 ‘He bought himself a house.’  
 b. \**Kupio je kuću sebi<sub>RM-DAT</sub>*  
 ‘He bought a house for himself.’  
 c. *Kupio je kuću sebi<sub>RM-DAT</sub> a ne djeci.*  
 ‘He bought a house for himself, not for the children.’  
 d. \**Kupio si<sub>MM-DAT</sub> je kuću, a ne djeci.*  
 ‘He bought himself a house, and not for the children.’

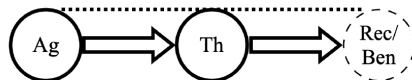
Given the distribution of the light and the heavy form shown in (13a–d) above, we conclude that indirect middle and reflexive pairs behave exactly like direct middles and reflexives in (1). In other words, the contrasts can also be explained by appealing to the conceptual-semantic parameter, viz. a lower or higher degree of participant distinguishability and event elaboration. The only difference with indirect pairs is that *si* is not a particle but a reflexive pronoun, which means that unlike *se*, it functions syntactically as an indirect object – just like its heavy form *sebi*. Compare Figures 2a–b.

Figure 2. Conceptual-semantic difference between indirect reflexive and middle constructions

a. indirect reflexive



b. indirect middle



Further evidence that the light form *si* is as much a pronoun and an indirect object as the heavy form *sebi* comes from the following distributional fact. Namely, unlike the distribution of *se/sebe* in (5), the distribution of *si/sebi* with respect to the case (nominative/dative) of the comparative construction in (14c–d) is identical to the distribution of the light and heavy forms of personal pronouns in (14a–b), allowing of course for the difference in meaning:

- (14) a. *Petar meni<sub>DAT</sub> kupuje odjeću češće nego Ana<sub>NOM</sub> / Ani<sub>DAT</sub>*  
 ‘Petar buys clothes for me more often than Ana does.’  
 ‘Petar buys clothes for me more often than for Ana.’
- b. *Petar mi<sub>DAT</sub> kupuje odjeću češće nego Ana<sub>NOM</sub> / Ani<sub>DAT</sub>*  
 ‘Petar buys clothes for me more often than Ana does.’  
 ‘Petar buys clothes for me more often than for Ana.’
- c. *Petar sebi<sub>RM-DAT</sub> kupuje odjeću češće nego Ana<sub>NOM</sub> / Ani<sub>DAT</sub>*  
 ‘Petar buys clothes for himself more often than Ana does for herself.’  
 ‘Petar buys clothes for himself more often than for Ana.’
- d. *Petar si<sub>MM-DAT</sub> kupuje odjeću češće nego Ana<sub>NOM</sub> / Ani<sub>DAT</sub>*  
 ‘Petar buys clothes for himself more often than Ana does for herself.’  
 ‘Petar buys clothes for himself more often than for Ana.’

### 2.1.2. Multi-participant reciprocal middle verbs

Reciprocal middle verbs designate activities which two or more participants direct at each other/one another, which means that subjects/agents/experiencers also qualify as themes. Typologically, reciprocal middle verbs are secondary relative to single-participant middle verbs, viz. they constitute a functional extension of single-participant middle verbs and not the other way around (Lichtenberk 1991: 171). As with single-participant middle verbs, an iconic meaning-form relationship exists with multi-participant middles too. Middle constructions with the *se* particle involve a low degree

of participant distinguishability and event elaboration; a higher degree of their distinguishability/elaboration is found with explicit constructions (EC) featuring the phrase *jedan drugoga* ‘each other, one another’, cf. (15a–d):

- (15) a. *Grudali su se*<sub>MM</sub>  
 ‘They were in a snowball fight.’  
 b. *Grudali su jedni druge*<sub>EC-ACC</sub>  
 ‘They threw snowballs at each other.’  
 c. *Poljubili su se*<sub>MM</sub>  
 ‘They kissed.’  
 d. *Poljubili su jedno drugo*<sub>EC-ACC</sub>  
 ‘They kissed each other.’

Similarly to single-participant middle verbs discussed in 2.1.1., multi-participant middles also allow both direct reciprocal middle verbs and indirect reciprocal middles with their explicit counterparts. Indirect reciprocal middles are exemplified in (16a–d). In their object slots we no longer see the *se* particle, but the light dative form *si* with the semantic role of beneficiary/recipient:

- (16) a. *Dali su si*<sub>MM-DAT</sub> *poklone za Božić.*  
 ‘They gave **themselves** Christmas presents.’  
 b. *Dali su jedni drugima*<sub>EC-DAT</sub> *poklone za Božić.*  
 ‘They gave **each other** / **one another** Christmas presents.’  
 c. *Čestitali su si*<sub>MM-DAT</sub>  
 ‘They congratulated **each other/one another.**’  
 d. *Čestitali su jedni drugima*<sub>EC-DAT</sub>  
 ‘They congratulated **each other/one another.**’

In both direct and indirect constructions, the light and heavy form contrast mainly along the parameter of simultaneity/sequentiality (Lichtenberk 1991, Kemmer 1993: 109).

The lighter, middle form communicates a simultaneous event, the form with *jedan drugoga/jedni drugima* symbolizes a sequential event. Put differently, the light form involves holistic event condensation, the heavy form communicates a dispersive, analytical event construal. Example (15c) thus reads as one simultaneous kiss and (15d) as two kisses. As with constructions featuring the heavy reflexive form *sebe* (cf. 1c and 7c), this difference can be accentuated by adding a contrasting adversative clause (cf. *Poljubili su jedno drugo* ‘They kissed each other’ > *Prvo je x poljubio*

*y*, **a potom je y poljubio x** ‘First x kissed y, and then y kissed x’). This is also possible in the case of indirect reciprocal middles (cf. *Čestitali su jedni drugima* ‘They congratulated each other/one another’ > *Prvo je x čestitao y, a potom je y čestitao x* ‘First x congratulated y, and then y congratulated x’).<sup>71</sup> Kemmer (1993: 110) remarks quite accurately that the simultaneity inherent to the light form often carries implications of social interaction (absent from the heavy form). This is especially prominent in verbs of perception, cf. (17a–d):

- (17) a. *Vidjeli su se<sub>MM</sub> jučer u gradu.*  
 ‘They saw each other in town yesterday.’ i.e. ‘They met.’
- b. *Vidjeli su jedan drugoga<sub>EC-ACC</sub> jučer u gradu.*  
 ‘They saw each other in town yesterday.’ i.e. ‘They sighted each other but did not meet.’
- c. *Čuli su se<sub>MM</sub> jučer telefonom.*  
 ‘They heard from each other by phone yesterday.’
- d. *Čuli su jučer jedan drugoga<sub>EC-ACC</sub> na zabavi.*  
 ‘They heard each other yesterday at the party.’

Whereas (17a and c) imply social interaction – (17c) explicitly and (17a) implicitly – in (17b and d) there are no such implicatures, there is not even the idea of simultaneity of perception. It is just as likely that two people saw/heard each other at different moments in time. However, social interaction may be the only implicature triggered by reciprocal middle constructions, in which case explicit reciprocal constructions are ruled out, cf. (18a–d):

- (18) a. *Vidjet ćemo se<sub>MM</sub> uskoro.*  
 ‘We’ll see each other soon.’
- b. *\*Vidjet ćemo uskoro jedan drugoga<sub>EC-ACC</sub>.*  
 ‘We’ll see each other soon.’
- c. *Čut ćemo se<sub>MM</sub> uskoro.*  
 ‘We’ll hear from each other soon.’
- d. *\*Čut ćemo uskoro jedan drugoga<sub>EC-ACC</sub>.*  
 ‘We’ll hear from each other soon.’

<sup>71</sup> Lichtenberk (1991: 172) presents an interesting case from the Austronesian language To’aba’ita in the Solomon Islands. The language has an explicit reciprocal form *kwailiu*, which also means, among others, *front-back*. This form proves that the more basic sense of spatial reciprocity may metaphorically map onto sequentiality in explicit reciprocal constructions.

The reason why explicit reciprocal constructions in (18b and d) are objectionable is that they do not allow the visual or auditory event to be construed as a matter of pure perception – rather the purpose/goal of the visual/auditory event is highly salient and implies some form of planned social interaction (cf. e.g. *Svakako ćemo se<sub>MM</sub> uskoro vidjeti / čuti* ‘We will definitely soon hear from / see each other’ > *Svakako ćemo se<sub>MM</sub> uskoro vidjeti / čuti kako bismo porazgovarali o tome* ‘We will definitely soon see each other / hear from each other so we can talk about this’). In both (17) and (18), the metonymy VISUAL OR AUDITORY PERCEPTION FOR SOCIAL INTERACTION allows the notion of social interaction to be communicated via verbs of perception. However, (18) arguably involves a higher degree of pragmatic inferencing. An explicit reciprocal construction is also unsuitable for so-called collective reciprocal meanings (Lichtenberk 1985; Kemmer 1993: 98–102) as found with verbs like *okupiti se<sub>MM</sub>* ‘gather’, *ujediniti se<sub>MM</sub>* ‘unite’, *naći se<sub>MM</sub>* ‘come together’, etc. The middle construction is the only coding option since collective scenarios do not profile interactions between event participants to any notable extent/at all. Collectives communicate a summary construal of individuals as a compact unity without any emphasis on the interaction between them, cf. (19a–d):

- (19) a. *Okupili su se<sub>MM</sub> kako bi proslavili godišnjicu mature.*  
 ‘They gathered to celebrate their graduation anniversary.’  
 b. \**Okupili su jedni druge<sub>EC-ACC</sub> kako bi proslavili godišnjicu mature.*  
 ‘\*They gathered one another to celebrate their graduation anniversary.’  
 c. *Ujedinili su se<sub>MM</sub> kako bi se obranili od neprijatelja.*  
 ‘They united to defend themselves against the enemy.’  
 d. \**Ujedinili su jedni druge<sub>EC-ACC</sub> kako bi se obranili od neprijatelja.*  
 ‘\*They united one another to defend themselves against the enemy.’

However, since there are verbs whose middle forms have single-participant meanings, the explicit reciprocal construction is the only coding option for multi-participant reciprocity, as in (20a–d):<sup>72</sup>

- (20) a. \**Hoće li doći? – pitali su se<sub>MM</sub>*  
 ‘Are they coming? – they asked each other.’

<sup>72</sup> *Hoće li doći? – pitali su se* in (20a) is grammatical in the sense ‘they were wondering if they would come’.

- b. *Hoće li doći? – pitali su jedan drugoga*<sub>EC-ACC</sub>  
 ‘Are they coming? – they asked each other.’
- c. \**Ubili su se*<sub>MM</sub>  
 ‘They killed each other.’
- d. *Ubili su jedan drugoga*<sub>EC-ACC</sub>  
 ‘They killed each other.’

A reciprocal middle may occasionally take the explicit form. Instead of (15b), for instance, one may opt for the tautologous *Grudali su se*<sub>MM</sub> *jedni s drugima*<sub>EC-INSTR</sub> ‘They were in a snowball fight with each other’. Because of tautology, this construction is rare (even colloquially) but it is not unacceptable. It is a comitative construction featuring an obligatory instrumental case. Since its function is to increase the level of explicitness (cf. *Grudali su se*<sub>MM</sub> *jedni s drugima*<sub>EC-INSTR</sub> ‘They were in a snowball fight with each other’/\**Grudali su se*<sub>MM</sub> *jedni druge*<sub>EC-ACC</sub> ‘They threw snowballs at each other’) we may argue that it is an example of an explicit-explicative construction. Recall that it was argued in section 2.1.1. that even among middle verbs events may be elaborated to various degrees. The claim was illustrated with *Na putu smo zastali u jednoj uvali, okupali se*<sub>MM</sub> *i nastavili dalje* ‘We stopped on our way in a lagoon, took a dip and moved on’, which elaborates the event to a lower degree than *Ivan se*<sub>MM</sub> *ujutro okupao i otišao na posao* ‘Ivan took a bath in the morning and went to work’. It is similar with reciprocal middle verbs, except that differences in degree are here brought out by contrasting different verbs. Thus, e.g. *Njih se*<sub>MM</sub> *dvoje dobro slažu* ‘The two of them are getting along well’ or *Ova se*<sub>MM</sub> *dva dijela police savršeno poklapaju* ‘These two sections of the shelf align perfectly’ elaborate the events to a lower degree than *Grudali su se*<sub>MM</sub> ‘They were in a snowball fight’ or *Svađali su se*<sub>MM</sub> ‘They were arguing’. Generally, higher degrees of event elaboration will tend to be associated with animate participants rather than inanimate, because not all participant types are equally conceptually accessible on the animacy hierarchy.

### 2.1.3. Reciprocal-middle contrast and the scalar nature of transitive constructions

Let us wrap up this section with a cursory look at the reflexive-middle contrast from the perspective of the gradience of transitive constructions. The transitive prototype or the *canonical event model* (Langacker 1991: 286; Belaj & Tanacković Faletar 2017: 157–158) and the associated *source – path*

– *goal* image schema involve a transfer of energy along the event chain<sup>73</sup> featuring two prototypical participants – the agent (subject) as the “head” or source of the event chain and the patient (direct object) as its “tail” or target/goal – causing a change of state in the target participant. Against these, there are prototypical one-participant intransitive constructions. In the transitivity cline discussed in Hopper & Thompson (1980: 277), reflexive constructions were argued to sit between prototypical transitive and prototypical intransitive scenarios:<sup>74</sup>

«Reflexives in many languages have properties which can be explained by appealing to their intermediate status between one-argument and two-argument clauses: compared with one-argument clauses, they may be more Transitive...; compared with two-argument clauses, they typically display features associated with lower Transitivity». (Hopper & Thompson 1980: 277)

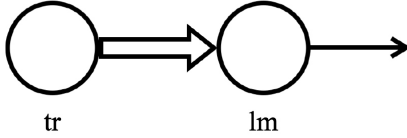
In the typological study cited above, Hopper & Thompson do not discuss middle constructions per se. Nevertheless, as there are important conceptual-semantic and consequently syntactic differences between prototypical reflexive constructions (referred to by Faltz (1977: 3) as the archetypal reflexive context), and middle constructions, their implicit transitivity scale *two-participant event* > *reflexive* > *one-participant event* should also accommodate middles. Because of their low degree of participant distinguishability and low degree of event elaboration, middle constructions sit squarely within the domain of intransitivity, i.e. they are closer to prototypical intransitive constructions. Therefore, we may propose the following transitivity scale for languages which, like Croatian, differentiate between transitive reflexive and intransitive middle constructions – [ <sub>TRAN.</sub> *two-participant event* > *reflexive constructions (reflexive)* ] > [ <sub>INTR.</sub> *middle verbs* > *one-participant event* ] – as shown in Figure 3.

<sup>73</sup> Cognitive linguistics is indebted to Talmy’s (1985, 2000) model of *force dynamics* for the idea of an event chain, which proved instrumental to the analysis of various grammatical constructions.

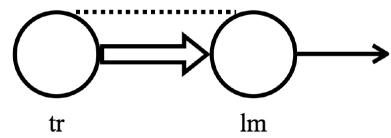
<sup>74</sup> On the gradience of transitive constructions in Croatian, cf. Belaj & Tanacković Faletar (2017), and for English, cf. Rice (1987).

Figure 3. A scale showing progressive reduction in participant distinguishability and event structure elaboration

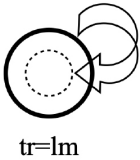
a. prototypical two-participant transitive event



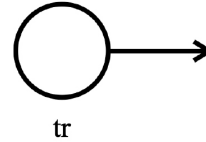
b. reflexive event



c. middle



d. prototypical one-participant intransitive event



The conceptual proximity between middle verbs and prototypical intransitive verbs is corroborated in Croatian dialects by various intransitive verbs that regularly appear in the middle form. For instance, *plakati se<sub>MM</sub>* ‘to cry’ vs. *plakati* ‘to cry’, *sjesti se<sub>MM</sub>* ‘to sit oneself down’ vs. *sjesti* ‘to sit’. Cf. also the reciprocal middle *razgovarati se<sub>MM</sub>* ‘to talk’ vs. *razgovarati* ‘to talk’. Further evidence of their proximity comes from the already mentioned redundant middles like *ustati se<sub>MM</sub>* / *ustati* ‘to get up’, *skijati se<sub>MM</sub>* / *skijati* ‘to ski’, *klizati se<sub>MM</sub>* / *klizati* ‘to skate’, *šetati se<sub>MM</sub>* / *šetati* ‘to walk’, etc.<sup>75</sup>

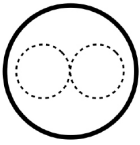
While reflexive constructions are characterized by the separation of ego, or as put by Langacker & Munro (1975), by the mind/body distinction, in middle verbs the ego is whole; there is not so much as an inkling of distinct participants being involved. One could say that, as far as the degree of participant distinguishability and event elaboration, middles represent more schematic construals than reflexive constructions, and reciprocal middle verbs compare to explicit reciprocal constructions in the same way. Middle verbs thus designate a holistic entity where a reduced phonological pole iconically represents a reduction on the conceptual-semantic pole.

<sup>75</sup> The form *sjesti se<sub>MM</sub>* ‘to sit oneself down’ is also motivated by analogy with *ustati se<sub>MM</sub>* ‘to get (oneself) up’.

Langacker (1991: 371) draws a remarkable parallel between middle single-participant verbs and *pluralia tantum* nouns, on the one hand, and reflexive constructions and regular plurals on the other. This is illustrated in Figure 4a–d.

Figure 4a–d. Parallel between *pluralia tantum* nouns/middle verbs and regular plurals/reflexive constructions

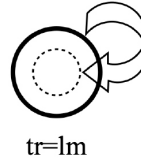
a. *pluralia tantum* nouns



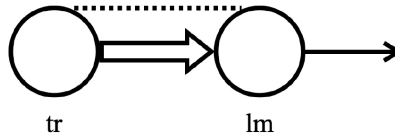
c. regular plural



b. middle verbs



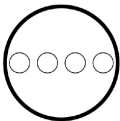
d. reflexive constructions



In addition to the parallelism between *pluralia tantum* nouns and single-participant middle verbs, we may also propose an analogy between collective nouns and reciprocal middle verbs, and between regular plurals and explicit reciprocal constructions (Fig. 5a–d).

Figure 5a–d. Parallel between collective nouns/reciprocal middle verbs and regular plurals/explicit reciprocal constructions

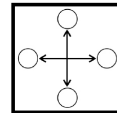
a. collective nouns



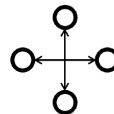
c. regular plurals



b. reciprocal middle verbs



d. explicit reciprocal constructions



Examples (21a–b), cited by Moskovljević (1997: 110) and Oraić Rabušić (2018: 69) demonstrate perfectly the difference between a holistic and separate conception of the ego and how this underlies the distinction between middle and reflexive constructions. In this example *se* is seemingly substitutable by *sebe*:

- (21) a. *Ivan se<sub>MM</sub> brani bolje od Marka.*  
 ‘Ivan defends himself better than Marco.’  
 b. *Ivan sebe<sub>RM</sub> brani bolje od Marka.*  
 ‘Ivan defends himself better than Marco.’

As duly noted by the mentioned authors, (21b) can be interpreted in three ways: *Ivan brani sebe<sub>RM</sub> bolje nego što brani Marka* ‘Ivan defends himself better than he defends Marco’/ *Ivan sebe<sub>RM</sub> brani bolje nego što ga Marko brani* ‘Ivan defends himself better than Marco defends him’ > *Ivan brani sebe<sub>RM</sub> bolje nego što Marko brani sebe<sub>RM</sub>* ‘Ivan defends himself better than Marco defends himself’. Yet, only the latter matches the meaning of the *se*-construction in (21a). Besides *se*- and *sebe*-constructions differing in the number of permissible interpretations (one vs. many), a more fundamental difference is that not even the construction *Ivan brani sebe<sub>RM</sub> bolje nego što Marko brani sebe<sub>RM</sub>* ‘Ivan defends himself than Marco defends himself’, stands as a true semantic equivalent to the *se*-construction in (21a). Their semantic base is indeed equivalent, however, they each involve a different construal or type of event elaboration – a far more important contrast to our mind. In the middle *se*-construction characterized by low event elaboration, the profile only includes the fact that Ivan is more successful defending himself than Marco. In the reflexive *sebe*-construction, in turn, because the ego is separated, Ivan and Marco come to be construed as defence attorneys each defending himself as the defendant.

## 2.2. Middle, passive, and impersonal constructions

This section is devoted to syntactic constructions with the particle *se*. This includes prototypical middle constructions where middle is the only profiled meaning, and the functionally adjacent passive and impersonal *se*-constructions, which can be construed as functional extensions of general middle semantics. The latter constructions hold on to their primary passive and impersonal semantics but also harbour the meaning of middles in their semantic base, which is the contribution of the basic semantics of the

*se* particle. We first discuss the relationship between middle and passive constructions (2.2.1.), and then move on to examine the relationship between middles and impersonals (2.2.2.)

### 2.2.1. Middle and passive constructions

As stated in the introduction to this paper, middle constructions are hardly ever mentioned in the Croatian literature. Even when they are (Oraić Rabušić 2018: 175–177; Batinić Angster 2020: 64–65), only the most prototypical middles denoting static generic/habitual situations – as exemplified in (22a–c) are considered:<sup>76</sup>

- (22) a. *Krumpir se kuha pola sata.*  
       ‘Potato is boiled for half an hour.’  
       b. *Vrtovi se okopavaju motikama.*  
       ‘Gardens are tilled with hoes.’  
       c. *Kruh se reže nožem.*  
       ‘Bread is sliced with a knife.’

Such middles occasionally surface in discussions of the passive construction, where they are usually singled out for their generic nature. Because they are perceived as lacking dynamicity they suffer one of two fates: (a) they are either excluded from the passive category (Vukojević 1992) or (b) they are included among passives, but only as a highly marginal member (Belaj 2004; Belaj & Tanacković Faletar 2017). Their generic nature is indicated by calling attention to their incompatibility with adverbials signalling a passive interpretation, like *sada* ‘now’, *upravo* ‘at this very moment’, *danas* ‘today’, *jučer* ‘yesterday’, etc.; and conversely, their compatibility with those indicating habituality, like *obično* ‘usually’, *u pravilu* ‘as a rule’, *redovito* ‘regularly’, etc. Indeed, given that middles topicalize the process at the expense of its participants (Manney 2001: 40; Calude 2017), these generic constructions come out as prototypical middles with their extremely low degree of event elaboration and consequently a

<sup>76</sup> In her typological study couched in the generative Theta System theory (Reinhart 2002), Marelj (2004) also advocates limiting the middle to the prototypical generic constructions shown in (22a–c).

strongly profiled process, stripped of its participants.<sup>77</sup> They have a highly schematic, generic agent, which is cognitively hardly accessible. In that respect, these middles are similar to passives since both are strategies of detopicalization of the agent. But unlike passives, these middles take this to the extreme. Passives generally, including the Croatian *se*-passive, are primarily strategies of topicalization of the theme participant. This has strong cognitive ramifications since, despite agent detopicalization, the topicalization of the theme in the target domain of the event chain implies that the agent is still cognitively accessible and quite specific. The dynamic nature of the passive construction implies the existence of a relation between the entity at which the action is directed and the entity responsible for its instigation. Let us compare the middle construction *Vrtovi se okopavaju u proljeće* ‘Gardens are tilled in the spring’ and the passive *Vrtovi se okopavaju već tjedan dana* ‘The gardens are being tilled for a week now’. The agent cannot be expressed in either the middle or the passive, but in the passive, it is cognitively more accessible because it is more specific (cf. the active counterpart *Moj susjed vrt okopava već tjedan dana* ‘My neighbour has been tilling the garden for a week now’). In the middle, the agent is a very schematic type and could be glossed as *ljudi* ‘people’ or at its most specific *vrtlari* ‘gardeners’. Nonetheless, we still contend that Croatian *se*-passives are akin to the middle because they feature the *se* particle. To put it simply, passive *se*-constructions are not middles if compared with the prototypical generic middle. But they are middles considering how they relate to the Croatian periphrastic passive, which is formed with the passive participle and the auxiliary *biti* ‘be’. Compare *Kuća je izgrađena u dva mjeseca* ‘The house was built in two months’ and *Kuća se gradila dva mjeseca* ‘The house was built for two months’. The periphrastic passive clearly profiles a state, the *se*-passive profiles the process.<sup>78</sup> Since *se*-passives involve a lower degree

<sup>77</sup> Their generic nature often results in idiomatization, which may be thought of as “habitualization” of the highest order. Cf. *Loš glas daleko se čuje* – similar in meaning to ‘Bad news travels fast’ – but here indicating the fast spread of bad reputation. Or *Prvi se mačići bacaju u vodu* lit. ‘First kittens end up down the drain’, a saying which suggests that the first failed attempt may be excused, etc.

<sup>78</sup> Because they profile a state, periphrastic passive constructions are much more frequent with participial forms of perfective verbs. But even when they feature imperfective verb participles, e.g. *Kuća je gradjena dva mjeseca* ‘The house was being built for two months’, the profile includes not only the duration of the construction process, but also various stages of construction completion over the two-month span. This is the contribution of the passive participle, which is an atemporal relational predication profiling the end state of the process in the conceptual base.

of event elaboration due to lower participant distinguishability, they do not admit agents, only non-volitional agents or effectors.<sup>79</sup> Periphrastic passives are not subject to that restriction (cf. *Kuća je izgrađena u dva mjeseca od (strane) jedne dobre građevinske tvrtke*/\**Kuća se izgradila u dva mjeseca od (strane) jedne dobre građevinske tvrtke* ‘The house was built in two months by a good construction company’). The reason why agents cannot be coded in *se*-passives has much to do with event structure reduction. As a prototypical instigating participant, an agent would increase the degree of event elaboration. Also, since agents are cognitively more accessible than effectors or non-volitional agents, their presence in prototypical agentive events also increases the degree of participant distinguishability. All in all, we may conclude that Croatian aligns itself with languages where passive meanings can be coded by middle markers because passives and middle verbs are both characterized by low participant distinguishability or non-distinctness of arguments (Langacker & Munro 1975: 801; Langacker 1976).<sup>80</sup> Of course, passives and middle verbs differ in the nature of non-distinctness of their arguments. In passives, the arguments are not the same, but their non-distinctness comes from detopicalizing the agent and consequently, reducing and levelling out the argument structure.

There is a gradience to the generic quality of middles. Apart from fully habitual constructions like those in (22a–c), there are also limited generic constructions, where the degree of event elaboration is somewhat higher. Compare, for instance, *Ove se godine jabuke dobro prodaju* ‘This year apples sell well’ or *Danas se knjige slabo kupuju* ‘Today books sell poorly’. Since there is a time limit to the temporal extent of these actions, we will refer to these constructions as *semi-generic* or *semi-habitual middles*. Besides generic constructions with a theme-marked subject, middle semantics are also found in agentive generic constructions like *On se stalno tuče* ‘He’s always picking a fight’, *Pčele se bodu* ‘Bees sting’, *Mačke se grebu* ‘Cats scratch’, etc., where the theme in the target domain of the event chain is detopicalized. Lichtenberk (1991) calls such constructions *depatientive* and Geniušienė (1987) labels them *absolute reflexive verbs*. We shall refer to them as *agentive*

<sup>79</sup> For instance, *Stadion se puni gledateljima* ‘The stadium is being filled with spectators’ features a non-volitional agent; the spectators do not have the intention of filling the stadium, their intention only extends as far as simply watching the game. In other words, the agent’s volitional action has nothing to do with the process designated by the verb.

<sup>80</sup> Langacker & Munro (1975) and Langacker (1976) do not discuss middles, only the similarities between reflexive and passive constructions.

*generic middles*. In the agentive generic middle construction, participant distinguishability is also low because the theme is highly schematic, generalized, and undifferentiated – which makes it harder to cognitively access. Both generic middles with the theme subject and those with the agent subject are iconic because – if a participant fails to qualify for profiling and only figures as part of the conceptual background on the construction’s semantic pole – it will not qualify for coding on the phonological pole either. At this juncture it may be useful to recall section 2.1.3., where it was argued that middle verbs sit before prototypical one-participant intransitive verbs and immediately after reflexive constructions on the transitivity scale, which confirms middles as the “intermediate” or “middle” voice. The same can be observed with syntactic constructions, where middles also surface as an in-between category. In the hierarchy of event elaboration, the generic, the agentive generic and the semi-generic *se*-constructions fall between prototypical one-participant absolute constructions/thematic processes and passives:<sup>81</sup>

[[<sub>INTR.</sub> *absolute/thematic process* > *generic/habitual middle* > *agentive generic/habitual middle* > *semi-generic/semi-habitual middle* > *passive*]  
> [<sub>TRAN.</sub> *active*]]

The more defocused a participant, the more salient the middle or absolute meaning and vice versa. Put differently, a higher degree of agent defocusing is proportional to a higher degree of profiling of the middle or absolute meaning. This conceptual hierarchy is shown in Figures (6a–f). Thickness of the dotted lines symbolizes the degree of defocusing of one of the participants or the degree of cognitive salience of the actions leading to the generic state (6b–d).<sup>82</sup> The construction in (6a) designates a pure

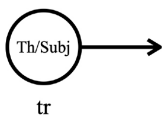
<sup>81</sup> In Cognitive Grammar, absolute constructions or absolute construals involve independent profiling of thematic processes, viz. a pure conception of thematic relations without co-activation of any event chain. For more, see Langacker (1991: 291, 389–393); Belaj & Tanacković Faletar (2017: 284).

<sup>82</sup> In order for a statement like that in (6b) – viz. that iron should be melted at 1000 degrees – to become an accepted truth, the process must have been construed dynamically in the past. In other words, someone must have tried experimenting with different melting degrees to establish the optimal one. This is why Belaj (2004) argued that generic constructions must not be banished from the passive category. With a little bit of cognitive unpacking of the associated event structure, they too can be associated with dynamic construals. In other words, their static generic nature has emerged from past dynamic processes, and their profiled trajectors correspond to the landmarks of those dynamic processes.

thematic process, neither the agent nor the event chain is evoked. The example simply means that the chemical composition of iron is such that it melts at 1000, rather than, say, 2000 degrees. Next in hierarchy are generic and agentive generic constructions in (6b–c), where one of the participants is highly schematic but, unlike in (6a, remains within conscious awareness. In semi-generic middles (6d) the degree of event elaboration is even higher because constructions of this type activate agents in more specific domains like e.g. apple growers and sellers in the domain of fruit growing and trade, or the majority of today’s population, especially youth, in the domain of digitalized publishing in *Danas se knjige slabo kupuju* ‘Books sell poorly today’. In passive constructions, due to the dynamic construal of the process, its agent is even more specific. This then results in a higher degree of event elaboration and participant distinguishability – e.g. of employees in an ironworks in example (6e). Finally, prototypical two-participant active constructions, where event chain participants from both source and target domain are profiled, enjoy the highest degree of event elaboration.

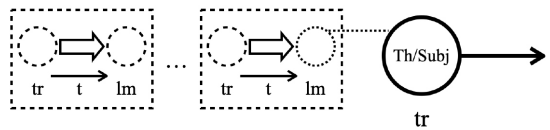
Figure 6a–f. Hierarchy of degrees of event structure elaboration in Croatian *se*-constructions

- a. Željezo se tali na 1000 stupnjeva. >  
(Talište je željeza na 1000 stupnjeva.)  
‘Iron melts at 1000 degrees.’  
(‘Iron’s melting point is at 1000 degrees.’)



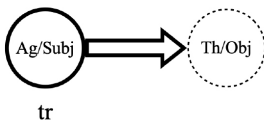
(absolute/thematic process)

- b. Željezo se tali na 1000 stupnjeva. >  
(Željezo se obično tali na 1000 stupnjeva.)  
‘Iron melts at 1000 degrees.’  
(‘Iron usually melts at 1000 degrees.’)



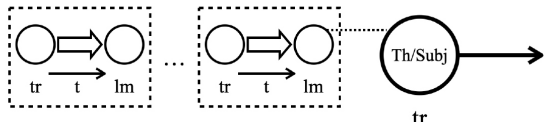
(generic/habitual middle)

- c. On se stalno tuče.  
‘He’s always picking a fight.’



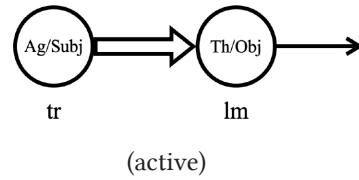
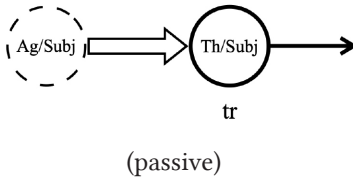
(agentive generic/habitual middle)

- d. Ove godine jabuke se dobro prodaju.  
‘This year apples sell well.’



(semi-generic/semi-habitual middle)

- e. Željezo se jučer talilo na 1000 stupnjeva. 'Yesterday iron was melted at 1000 degrees.'  
 f. Jučer smo talili željezo na 1000 stupnjeva. 'Yesterday we melted iron at 1000 degrees.'



The boundary between passives, middles, and absolute constructions is not always clear, to be sure. Compare (23a–c):

- (23) a. *Prozor se brzo popravio.*  
 'The window got fixed quickly.'  
 b. *Prozor se brzo otvorio.*  
 'The window opened quickly.'  
 c. *Prozor se otvorio.*  
 'The window opened.'

Example (23a) is unquestionably passive because the agent, though unexpressed, is conceptually very much present and distinguishable. A window cannot fix itself; a human agent must be part of the conceptual background.<sup>83</sup> Example (23c) in turn is a case of absolute construal since the situation is conceptualized autonomously, i.e. without reference to the causal chain. If construed dependently, the causal chain is made part of active conceptualization and we can appreciate the link between this construction and the passive. Namely, a window cannot open itself, there is always some kind of causal force, an effector of sorts, like the wind or drought. This makes these constructions different from the construction instantiated by *Željezo se tali na 1000 stupnjeva* 'Iron melts at 1000 degrees' (Fig 6a), as in the latter

<sup>83</sup> Passive or absolute interpretations are very often sanctioned constructionally, i.e. they depend on the verb at the hub of the construction. For instance, *Kuća se gradi* 'The house is being built' is unquestionably passive, just like (23a). The example *Kuća se ruši* 'The house is being demolished/falling apart' can be interpreted as a passive if the event involves planned demolition by a contractor, but it can also be interpreted in the absolute sense because a house can collapse all on its own due to age or disrepair. Finally, *Kuća se urušava* 'The house is falling in on itself' can only be interpreted as an absolute because the process does not presuppose a causal chain.

no causal force exists. Regardless of A/D alignment, this is an instance of a purely thematic process. However, if we exclude from the analysis dependent event conceptualization and treat (23c) as an absolute construction, the status of (23b) remains unresolved. On the one hand, it is close to the passive, because the adverbial *brzo* ‘fast’ increases the likelihood of agent activation. But it is not barred from the absolute interpretation, because, ultimately, a window can open fast without the agent’s interference. We seem to be looking at a meaning that sits somewhere between prototypical passives and prototypical absolutes. It does involve a higher degree of participant distinguishability and event elaboration than that typically found in absolute constructions, but this degree is still lower than with passives. Langacker (2008: 385) classes constructions like (23b) as middles, because their degree of agentivity is higher compared to absolute constructions, cf. (23c), and is lower in comparison to prototypical two-participant actives like *Otvorio sam prozor* ‘I opened the window’. However, in English, the passive only corresponds to Croatian periphrastic constructions, cf. *Prozor je (bio) otvoren* (‘The window was opened (by me)'). Thus, unlike the middle, which profiles only what happens to the theme, the passive profiles the full theme-agent relation. That is why English middles do not contrast so much with the passive as they do with absolute constructions. Because the situation is different in Croatian (where the passive is also coded by a middle marker), and because the meaning is somewhat closer to the passive due in large part to the adverbial modifier, which always affects event structure elaboration,<sup>84</sup> we feel it is only fair to credit these constructions with the label *middle-passive*. These relationships are shown in Figures (7a–d). The extended hierarchy of event structure elaboration now looks as follows:

---

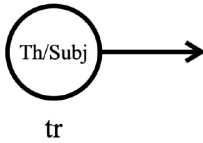
<sup>84</sup> Various adverbial (in the source domain of event chain) and adjectival (in the target domain of event chain) modifiers affect the factuality of events, and with it also the degree of prototypicality of the transitive scenario (even where it is only part of the background of intransitive constructions). Verbal aspect and telicity are also important. They have a considerable impact on event dynamicity, which is proportional to the degree of prototypicality of the transitive event. If, for instance, we were to replace the perfective aspect in (23) with a non-perfective, the constructions would become “static” and would transition to prototypical generic middles (cf. *Prozor se brzo popravlja* ‘The window repairs fast’ / *Prozor se brzo otvara* ‘The window opens fast’/ *Prozor se otvara* ‘The window opens’). For more on various factors behind the prototypicality of transitivity, see Hopper & Thompson (1980), for English see Rice (1987), for Croatian, see Belaj & Tanacković Faletar (2017).

[[<sub>INTR.</sub> *absolute / thematic process > generic / habitual middle > agentive generic / habitual middle > semi-generic / semi-habitual middle > middle-passive > passive*] > [<sub>TRAN.</sub> *active*]]

Figure 7a–d. Hierarchy of increasing event elaboration in middle-passive and passive constructions

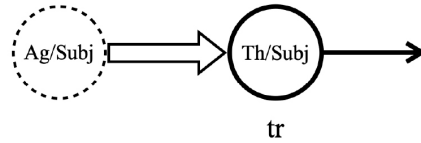
a. Prozor se otvorio.

‘The window opened.’



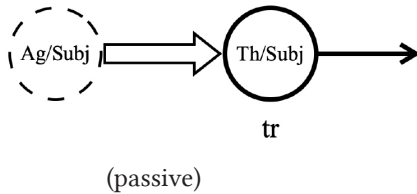
b. Prozor se brzo otvorio.

‘The window opened quickly.’



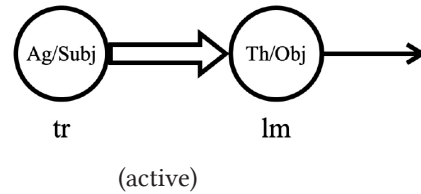
c. Prozor se brzo popravio.

‘The window was fixed quickly.’



d. Brzo je popravio prozor.

‘He quickly fixed the window.’



### 2.2.2. Middle and impersonal constructions

A few words remain about the relationship between the semantics of middles and impersonal *se*-constructions. Besides characterizing prototypical generic middle constructions in (22), one of the central properties of “middles”, viz. the topicalization of the process, also underlies the impersonals illustrated in (24a–e):

(24) a. *Priča se o tome.*

‘There is talk of that.’

b. *Gradi se kuću.*

‘A house is being built.’

c. *Spava mi se.*

‘I’m sleepy.’

d. *Stari se.*

‘We are getting old/People get old.’<sup>85</sup>

e. *Naoblačilo se.*

‘It turned cloudy.’

Like every other impersonal construction, Croatian impersonal *se*-constructions detopicalize the agent, making this participant ineligible for subject coding. As such, they join a cohort of other constructions with the basic semantics of middles, viz. a low degree of participant distinguishability and event elaboration. Impersonal *se*-constructions differ from this set, though, in how the theme argument is coded, if it is coded at all. The result is an exclusive or partial profiling of the process. A gradience can be established between impersonal *se*-constructions on the basis of this parameter. Example (24a) features a construction most commonly occupied by verbs of speaking and cognition (e.g. *Priča se* ‘There’s talk/gossip’/ *Govori se* ‘There are stories/rumors’ / *Raspravlja se* ‘There’s discussion about x’ / *Razmišlja se o tome* ‘X is being considered’). The agent is detopicalized, schematic, and plural; the theme object is coded in the locative case. Constructions like (24b) are alternatives to passive *se*-constructions, from which they only differ in having a direct object in the accusative instead of a passive subject. As a result, the passive interpretation becomes more salient but so does the process itself since the object’s affectedness is also part of the profile. Impersonal constructions of the type shown in (24c) are especially interesting, both semantically and syntactically. They feature verbs of psychological and physiological events with dative-marked arguments (*Kiše mi se* ‘I feel like sneezing’, *Spava mi se* ‘I feel like sleeping’, *Vrti mu se* ‘He’s feeling dizzy’, *Kašlje joj se* ‘She feels like coughing’, etc.). The latter are occasionally dubbed *logical subjects* (e.g. Barić *et al.* 1995: 427) because they are recognized as subjects on semantic and pragmatic, but not on morphosyntactic grounds (cf. e.g. *Spava mi se* ‘I feel like sleeping’ > *Osjećam potrebu za spavanjem* ‘I feel the need to sleep’). However, looking deeper into their cognitive background, it becomes clear that these constructions also display semantic-pragmatic features of objects

<sup>85</sup> English lacks an impersonal construction as versatile as the Croatian sentence in (24d). Depending on the context the sentence can be translated having any subject (e.g. *Uvijek zaboravljam/zaboravljaš gdje sam/si parkirao. Stari se* ‘I/you always forget where I/you park the car. I am/You are getting old’).

(Belaj 2015). By detopicalizing the effector (e.g. in 24c *umor* ‘tiredness’), agentivity is passed on to the process itself. The effector is absorbed and becomes equivalent to the process it causes. This process is stripped of its participants in the source domain of the event chain and is directed at the dative-marked referent in the target domain. Additional justification for seeing in these dative complements semantic features of objects comes from Cognitive Grammar, where the dative-marked indirect object prototype is defined as the *active experiencer in the target domain* (Langacker 1991: 327) and as the *target person* (Dąbrowska 1997; Palić 2006; Belaj & Tanacković Faletar 2012). Without the grammatical subject and direct object, the dative-marked indirect object is the next one in line to become the pragmatic topic (cf. Shibatani 1985: 832).

The semantic and pragmatic features of objects also become evident in these dative arguments in a type of paraphrase where the dative referent assumes the prototypical position of the direct object, not of the subject (cf. *Spava mi se* ‘I feel like sleeping’ > *Obuzela me potreba za spavanjem* ‘The urge to sleep came over me’). Constructions like (24b–c) are especially interesting because they appear to depart from the universal feature of middles, viz. focus on the subject’s dominion (Maldonado 1992, 1999). Still, although they do feature objects, these objects are each in its own special way connected to the subject. In (24b) the object is an alternative to the more prototypical passive subjects and is thus associatively linked with them.<sup>86</sup> As argued above, in (24c) this link is even clearer via the notion of the logical subject. The highest degree of process topicalization is found in constructions exemplified in (24d–e), where the theme argument falls short of being coded at all – thus, they are all only about the process. (24d) exemplifies generic<sup>87</sup> constructions of the type *Živi se* ‘I’m getting by’, *Životari se* ‘I’m scraping by’, *Stari se* ‘I’m getting old’, *Gura se* ‘I’m getting

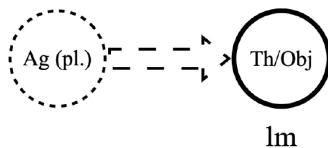
<sup>86</sup> Older Croatian prescriptive literature (Jonke 1952, Vince 1955) pegged such constructions as irregular due to the presence of objects in passive sentences – an alleged influence from other languages. Due to their ubiquity, especially in colloquial language, these normative assessments are no longer accepted. As a rule, descriptive grammars describe these constructions as subtypes of impersonals.

<sup>87</sup> Note that all types of impersonal constructions in (24a–e) receive a habitual construal if accompanied by a generic adverbial (cf. *Kod njih se uvijek puno priča, a malo radi* ‘Over there, there’s always a lot of talk, but little work gets done’/*Kuću se gradi u proljeće* ‘Houses are built in the spring’/*Ujutro mi se uvijek spava* ‘I always feel sleepy in the morning’/*Zimi se kasno razdanjuje* ‘In winter it dawns late’, etc.).

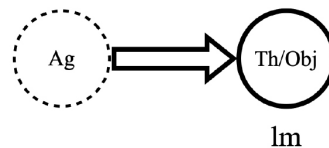
by', etc., where the singular or plural theme is present in the conceptual background, or, as in (24e), where it is not present anywhere at all. In (24e) the impersonals feature *verba meteorologica*, of which some are in the middle form with the particle *se* (e.g. *Razvedrilo se* 'The clouds broke', *Naoblačilo se* 'It turned cloudy', *Razdanilo se* 'It dawned', *Smrknulo se* 'It grew dark', etc.), and others are without it (e.g. *Kiši* 'It's raining', *Sijeva* 'There's a lightning storm', *Pljušti* 'It's pouring with rain', *Pada* 'It's raining/snowing', *Sniježi* 'It's snowing', *Grmi* 'There's a thunder storm', etc.).<sup>88</sup> In these constructions participant distinguishability is at its lowest, as is the degree of event elaboration; conceptually, they code events that are all-encompassing, completely levelled out and homogenized so much so that no significant conceptual partitioning can be detected between the process, participants, place or time of the event.<sup>89</sup> In other words, a processual relation devoid of any participant construal comes into profile. These middle impersonal *se*-constructions are shown in Figure (8a–e), with a focus on the gradience of process topicalization.

Figure 8a–e. Gradient increase in process topicalization among middle impersonal *se*-constructions

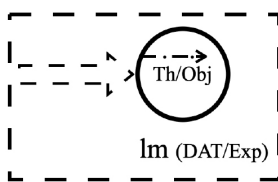
a. *Priča se o tome.* 'There's talk of that.'



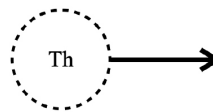
b. *Gradi se kuću.* 'A house is being built.'



c. *Spava mi se.* 'I feel sleepy.'



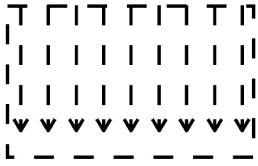
d. *Stari se.* 'We are getting old.'



<sup>88</sup> Both types of verbs appear either in 3rd person singular form (e.g. *Kiši* 'It's raining' / *Razvedrava se* 'It's clearing up') or in the singular, neuter gender form of the active past participle (*Kišilo je* 'It rained' / *Razvedrilo se* 'It cleared up'). For more on these constructions see Ruwet (1986), Langacker (1991: 365–366), Smith (2002), and for Croatian, Belaj (2007).

<sup>89</sup> This conceptual configuration is best illustrated with Givón's (2001: 119) example from spoken Palestinian Arabic, where the corresponding verbs take the noun meaning 'world' as their subject (id-*dunya* ti-shti, translated loosely as *Svijet kiši* 'The world rains').

e. Naoblačilo se. 'It's become cloudy.'



The syntactic and semantic diversity of impersonal *se*-constructions hinders their unique assignment to a specific position in the hierarchy of increasing event elaboration. For instance, constructions in (24e) would fall more naturally ahead of absolute constructions since they are marked by the lowest degree of event elaboration. Those in (24a–b) in turn would sit closer to passive constructions. If, however, one took into account only their common denominator, viz. a strongly profiled process with more or less reduced participant distinguishability, middle constructions would then cluster in that same hierarchy together with the three types of generic middles sharing that particular property.

### 3. Conclusion

This study examined Croatian middle *se*-constructions, including both single-participant and reciprocal multi-participant middle verbs (2.1.) and syntactic *se*-constructions (2.2.). The main argument was that the term *reflexive verb*, current within and beyond Croatian linguistics, should be replaced with the term *middle verb*. This is in part due to the former's purely morphological and diachronic motivation. However, a range of additional arguments of both semantic-pragmatic and syntactic nature were advanced in favour of this proposal. The bottom line is, if two construction types differ so markedly as to merit inclusion into separate categories, this should also be mirrored in a clear terminological separation. Croatian middle verbs feature particle *se*, the light form of the reflexive pronoun in a highly grammaticalized particle function. This function symbolizes the schematic meaning of middles, glossed as *low degree of participant distinguishability* and *low degree of event elaboration* and situates middle verbs in the transitivity hierarchy somewhere between transitive reflexive constructions and prototypical one-participant intransitive constructions. It was also argued that the concept of reflexivity should in turn only be associated with transitive reflexive

constructions featuring the heavy form of the reflexive pronoun (*sebe* ‘self’). These constructions elaborate the said features to a higher degree, and therefore behave differently syntactically. This property also situates them closer to the prototypical two-participant transitives. Besides middle verbs, we also examined how the particle *se* contributes the semantics of middles to higher-order syntactic constructions. These include prototypical generic middle constructions and, somewhat closer to the category fringe, absolute, passive, and impersonal constructions. The latter also exhibit, though in smaller measure, universal semantic properties of middles. What they all have in common is some degree of process topicalization, which is itself proportional to the degree to which the event is construed without distinguishable participants.

## Appendix

### Acknowledgments

This work has been fully supported by the Croatian Science Foundation under the projects *Syntactic and Semantic Analysis of Arguments and Adjuncts in Croatian – SARGADA* (IP-2019-04-7896) and *Semantic-Syntactic Classification of Croatian Verbs – SEMTACTIC* (IP-2022-10-8074).

I wish to thank anonymous reviewers for their valuable comments.

### Abbreviations

A/D	autonomous/dependent
ACC	accusative
Ag	agent
Ben	benefactive
DAT	dative
EC	explicit construction
Exp	experiencer
HF	heavy form of personal pronoun
INTR.	intransitive
LF	light form of personal pronoun
lm	landmark
MM	middle marker <i>se</i>
MORPH.	morphological system

NOM	nominative
Obj	object
PART	particle
Pl.	plural
PRON	pronoun
Rec	recipient
RM	reflexive marker <i>sebe</i>
Subj	subject
SYNT.	syntactic system
Th	theme
tr	trajector
TRAN.	transitive

## References

- Babić, S., Brozović, D., Moguš, M., Pavešić, S., Škarić, I. & Težak, S. (1991). *Povijesni pregled, glasovi i oblici hrvatskoga književnog jezika*. Zagreb: Globus.
- Barić, E., Lončarić, M., Malić, D., Pavešić, S., Peti, M., Zečević, V. & Znika, M. (1995). *Hrvatska gramatika*. Zagreb: Školska knjiga.
- Batinić Angster, M. (2020). *Rečenična anafora: Morfologija i sintaksa povratnih imenskih izraza u hrvatskom i talijanskom jeziku*. Zagreb: Institut za hrvatski jezik i jezikoslovlje.
- Belaj, B. (2001). Prototipno-kontekstualna analiza povratnih glagola u hrvatskom jeziku. *Suvremena lingvistika*, 51/52(1/2), 1–11.
- Belaj, B. (2004). *Pasivna rečenica*. Osijek: Filozofski fakultet.
- Belaj, B. (2007). Konceptualnosemantički aspekti prototipnih struktura s bezličnim i obezličnim glagolima i njihove sintaktičke implikacije. In: Kuna, B. (ed) *Sintaktičke kategorije*, 21–51. Osijek/Zagreb: Filozofski fakultet/Institut za hrvatski jezik i jezikoslovlje.
- Belaj, B. (2015). Što kognitivna gramatika može reći o tzv. dativnim subjektima?. *Croatica: časopis za hrvatski jezik, književnost i kulturu*, 59, 3–17.
- Belaj, B. (2023). *Kognitivna lingvistika i hrvatski jezik*. Osijek: Filozofski fakultet.
- Belaj, B. & Tanacković Faletar, G. (2012). Space, Conceptualization and Case Meaning: A Cognitive Account of the dative in Croatian. In: Brdar, M. & Raffaelli, I. & Žic Fuchs, M. (eds) *Cognitive Linguistics between Universality and Variation*, 53–93. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Belaj, B. & Tanacković Faletar, G. (2017). *Kognitivna gramatika hrvatskoga jezika, Knjiga druga, Sintaksa jednostavne rečenice*. Zagreb: Disput.
- Browne, W. (1974). On the problem of enclitic placement in Serbo-Croatian. In: Brecht, R. D. & Chvany, C. V. (eds) *Slavic Transformational Syntax*, 36–52. Ann Arbor, Michigan: Michigan Slavic Materials.
- Calude, A. S. (2017). Testing the boundaries of the middle voice: Observations from English and Romanian. *Cognitive Linguistics*, 28(4), 599–629.

- Dąbrowska, E. (1997). *Cognitive Semantics and the Polish Dative*. Berlin/New York: Mouton de Gruyter.
- Everaert, M. (1986). *The Syntax of reflexivization*. Dordrecht: Foris.
- Everaert, M. (1991). Contextual determination of the anaphor/pronominal distinction. In: Koster, J. & Reuland, E. (eds) *Long-distance Anaphora*, 77–118. Cambridge: Cambridge University Press.
- Faltz, L. (1977). *Reflexivization. A study in Universal Syntax*. PhD thesis, University of California.
- Foley, W. A. & Van Valin, R. D. Jr. (1984). *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Geniušienė, E. (1987). *The typology of reflexives*. Berlin/New York/Amsterdam: Mouton de Gruyter.
- Givón, T. (2001). *Syntax: An Introduction. Volume I*. Amsterdam/Philadelphia: John Benjamins.
- Grimshaw, J. (1982). On the lexical representation of Romance reflexive clitics. In Bresnan, J. (ed) *The mental representation of grammatical relations*, 87–148. Cambridge (MA): MIT Press.
- Haiman, J. (1983). Iconic and economic motivation. *Language*, 59(4), 781–819.
- Haspelmath, M. (2008). A frequentist explanation of some universals of reflexive marking. *Linguistics Discovery*, 6(1), 40–63.
- Heine, B. & Reh, M. (1984). *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske.
- Hopper, P. J. & Thompson, S. A. (1980). Transitivity in Grammar and Discourse. *Language*, 56(2), 251–299.
- Huang, Y. (2000). *Anaphora. A Cross-Linguistic Study*. Oxford: Oxford University Press.
- Ivić, M. (1962). Jedan problem slovenske sintagmatike osvetljen transformacionom metodom. *Južnoslovenski filolog*, 25, 137–151.
- Ivić, M. (1967). Types of direct object in Serbocroatian. In *To Honor Roman Jakobson. Essays on the Occasion of his Seventieth Birthday. Volume 2*, 989–994. Mouton: The Hague.
- Ivić, M. (1983). *Lingvistički ogledi*. Beograd: Prosveta.
- Jahić, Dž., Halilović, S. & Palić, I. (2000). *Gramatika bosanskoga jezika*. Zenica: Dom štampe.
- Jonke, Lj. (1952). Čita se Šenou ili čita se Šenoa?. *Jezik*, 1(3), 90–91.
- Katičić, R. (1991). *Sintaksa hrvatskoga književnog jezika, drugo, ponovljeno izdanje*. Zagreb: Globus.
- Kemmer, S. (1993). *The Middle Voice*. Amsterdam: John Benjamins.
- Kemmer, S. (1994). Middle voice, transitivity and events. In: Fox, B. A. & Hopper, P. J. (eds) *Voice: Form and Function*, 179–230. Amsterdam/Philadelphia: John Benjamins.
- König, E. (2001). Intensifiers and reflexives. In: Haspelmath, M. (ed) *Sprachtypologie und sprachliche Universalien: ein internationales Handbuch/Language typology and language universals: an international handbook*, 747–760. Berlin/New York: Mouton de Gruyter.
- König, E. & Siemund, P. (2000). Intensifiers and reflexives: a typological perspective. In: Frajzyngier, Z. & Curl, T. S. (eds) *Reflexives: Forms and Functions*, 41–74. Amsterdam: John Benjamins.
- Langacker, R. W. (1976). *Non-distinct arguments in Uto-aztecan*. Berkeley: University of California Press.

- Langacker, R. W. (1991). *Foundations of Cognitive Grammar. Volume 2: Descriptive Application*. Stanford, California: Stanford University Press.
- Langacker, R. W. (2008). *Cognitive Grammar. A Basic Introduction*. Oxford: Oxford University Press.
- Langacker, R. W. & Munro, P. (1975). Passives and their meaning. *Language*, 51(4), 789–830.
- Lehmann, C. (1985). Grammaticalization: Synchronic variation and diachronic change. *Lingua e stile*, 20(3), 303–318.
- Lichtenberk, F. (1985). Multiple uses of reciprocal constructions. *Australian Journal of Linguistics*, 5, 19–41.
- Lichtenberk, F. (1991). Reciprocals and depatientives in To'aba'ita. In: Blust, R. (ed) *Currents in Pacific linguistics: Papers on Austronesian languages and ethnolinguistics in honor of George W. Grace*, 171–183. Canberra: Australian National University.
- Lyons, J. (1969). *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.
- Maldonado, R. (1992). *Middle voice: The case of Spanish se*. San Diego (CA): University of California.
- Maldonado, R. (1999). A media voz: problemas conceptuales del clítico SE en español. México: Instituto de Investigaciones Filológicas.
- Maldonado, R. (2000). Spanish reflexives. In: Frajzyngier, Z. & Walker, T. (eds) *Reflexives: Forms and functions*, 153–185. Amsterdam/Philadelphia: John Benjamins.
- Maldonado, R. (2008). Spanish middle Syntax: A usage based proposal for grammar teaching. In: De Knop, S. & De Rycker, T. (eds) *Cognitive approaches to pedagogical grammar*, 155–196. Berlin: Mouton de Gruyter.
- Maldonado, R. (2009). Middle as a basic voice system. In: Guerrero, L. & Ibanez, S. & Belloro, V. (eds) *Studies in role and reference grammar*, 69–112. México: Instituto de Investigaciones Filológicas.
- Manney, L. (2001). *Middle voice in Modern Greek: Meaning and function of a morphosyntactic category*. Amsterdam/Philadelphia: John Benjamins.
- Marelj, M. (2004). *Middles and Argument Structure across Languages*. Utrecht: LOT.
- Moskovljević, J. (1997). Leksička detranzitivizacija i analiza pravih povratnih glagola u srpskom jeziku. *Južnoslovenski filolog*, 53, 107–114.
- Oliva, K. (2001). Reflexe reflexivity reflexiv (A reflexion on a reflexivity of reflexives). *Slovo a slovesnost*, 62(3), 200–207.
- Oraić Rabušić, I. (2018). *Struktura povratnih glagola i konstrukcije sa se u hrvatskome jeziku*. Zagreb: Institut za hrvatski jezik i jezikoslovlje.
- Palić, I. (2006). *Dativ u savremenom bosanskom jeziku (sintaksičko-semantički opis)*. PhD thesis, University of Sarajevo.
- Pavešić, S. & Vince, Z. (1971). Gramatika. In: Pavešić, S. (ed.) *Jezični savjetnik s gramatikom*, 329–446. Zagreb: Matica hrvatska.
- Raguž, D. (1997). *Praktična hrvatska gramatika*. Zagreb: Medicinska naklada.
- Reinhart, T. (2002). The Theta System – An Overview. *Theoretical linguistics*, 28(3), 229–290.
- Reuland, E. (1998). The fine structure of grammar: anaphoric relations. In: Frajzyngier, Z. (ed) *Reflexives: Forms and Functions*, 1–40. Amsterdam: John Benjamins.
- Rice, S. A. (1987). *Towards a cognitive model of transitivity*. PhD thesis, University of California.

- Rosen, C. (1989). *The relational structure of reflexive clauses: Evidence from Italian*. New York: Garland.
- Ruwet, N. (1986). On Weather Verbs. *Papers from the regional Meetings, Chicago Linguistic Society*, 22(1), 195–215.
- Shibatani, M. (1985). Passives and related constructions: a prototype analysis. *Language*, 61(4), 821–848.
- Silić, J. & Pranjković I. (2005). *Gramatika hrvatskoga jezika za gimnazije i visoka učilišta*. Zagreb: Školska knjiga.
- Smith, M. B. (2002). The polysemy of German es, iconicity, and the notion of conceptual distance. *Cognitive Linguistics*, 13(1), 67–112.
- Smith, M. (2004). Light and heavy reflexives. *Linguistics*, 42(3), 573–615.
- Talmy, L. (1985). Force Dynamics in Language and Thought. *Cognitive Science*, 12(1), 49–100.
- Talmy, L. (2000). *Toward a Cognitive Semantics. Volume I*. Cambridge (MA): MIT Press.
- Timberlake, A. (1975). Hierarchies of the genitive of negation. *Slavic and East European Journal*, 19(2), 123–138.
- Timberlake, A. (1977). Reanalysis and actualization in syntactic change. In: Li, C. (ed) *Subject and topic*, 141–177. New York: Academic Press.
- Traugott, E. C. & König, E. (1991). The semantics-pragmatics of grammaticalization revisited. In: Traugott, E. C. & Heine, B. (eds) *Approaches to grammaticalization. Volume 1*, 189–218. Amsterdam: John Benjamins.
- Van Valin, R. D. Jr. & LaPolla, R. J. (1997). *Syntax, structure, meaning and function*. Cambridge: Cambridge University Press.
- Vince, Z. (1955). O nepravilnoj upotrebi objekta u akuzativu u pasivnim rečenicama. *Jezik*, 4(3), 93–96.
- Vukojević, L. (1992). Sintaksa pasiva. *Rasprave Zavoda za hrvatski jezik*, 18(1), 235–257.
- Zec, D. (1985). Objects in Serbo-Croatian. In: Niepokuj, M., Van Clay, M., Nikiforidou, V. & Feder, D. (eds) *Proceedings of the Eleventh Annual Meeting of the Berkley Linguistics Society*, 358–371. Berkeley, California: University of California.

## 6. A diachronic change of the *reflexiva tantum* psychological verbs in Croatian

LIDIJA MILKOVIĆ

Keywords: *diachronic change; argument structure; Croatian language; psychological verbs; reflexivity*

In this paper, the changes in the valency of the *reflexiva tantum* psychological verbs in Croatian are analysed within the framework of the Slavic languages' history. Mainly, lexicographic sources were used to study the diachronic change of verb valency. The valency change was observed through five parameters. These parameters form a hierarchy of diachronic stability, starting from the least stable: the morphosyntactic properties of arguments > reflexive marking > transitivity > causativity > semantic class. In general, according to the results shown in this paper, most *reflexiva tantum* psychological verbs in the history of the Croatian language have a relatively recent origin. According to the typological parameter of the basic valency orientation, Croatian was, to a lesser degree, a detransitivizing language in the earlier periods of its history.

*Original scientific paper*

### 1. Introduction

Psychological verbs are a group of verbs with a stable syntactic and semantic structure which is visible in their potential for regular diathesis alternations and in their semantic role composition (Brač & Bošnjak Botica 2015; Padučeva 2004: 287; Van Gelderen 2019; Levin & Rappaport Hovav 2005). The *reflexiva tantum* verbs differ from the majority of verbs by appearing exclusively with a morphological reflexive marking (Oraić Rabušić 2018: 76; Belaj 2001: 4, *this volume*, section 2.1.1). A greater or smaller presence of the *reflexiva tantum* verbs in various periods within a language history can point to a diachronic change of the typological parameter of the basic valency orientation (Nichols, Peterson & Barnes 2004). The change of the possibility of entering reflexive pairs affects the other syntactic and semantic characteristics of verbs. Along with

the possibility of verbs taking a reflexive and a non-reflexive form, four additional parameters will be studied in this work: a) the change of the arguments, b) transitivity, c) causativity, and d) the verb meaning change. Based on the stated parameters, verbs will be classified in groups according to the changes which have modified the verb through time. Taking into consideration that for Croatian there are currently no dictionaries or corpora available online specialized in the earlier periods of the language and that the sources from one Slavic language do not make it possible to establish the diachronic primary state of the (non) reflexive form, the data on the stated parameters will be excerpted from the earlier periods of other Slavic languages (Old East Slavic dictionaries, Old Czech Dictionary (henceforth OCZ) and Old Polish Dictionary (henceforth SS)).<sup>90</sup> The comparison of the Slavic languages from the earlier periods with the contemporary Croatian language makes it possible to answer the question about the diachronic primary status of the *reflexiva tantum* psychological verbs and about the special characteristics of the change of the studied parameters in the history of Croatian.

## 2. *Reflexiva tantum* psychological verbs

The term *Reflexiva tantum* denotes a group of reflexive verbs which occur in language exclusively with a reflexive marker, i.e. which do not have a transitive verb pair (Oraić Rabušić 2018: 76; Belaj 2001: 4). The verbs within this morphologically defined category can be subjected to various classification principles. For the *reflexiva tantum* verbs, Belaj (2001: 4), also uses the Croatian term *primarnopovratni* (primary reflexive verbs) and separates these verbs into two categories based on the criterion of the relationship of the actants involved in an action. Verbs marked by an action that includes one or more individual subjects belong to the first group and verbs with reciprocal-reflexive meaning belong to the second

---

<sup>90</sup> There are open-access transliterations of the Beram breviaries available on the website of the *Old Church Slavonic Institute*. The Beram breviaries are written in Croatian Church Slavonic, which differs from Old Croatian. In addition, there is also the *Regensburg Croatian diachronic* corpus, consisting of texts written between the 16th and 19th century (Hansack *et al.* 2016: 6). Unfortunately, the corpus is not currently available online, so it is not included in this research.

group. Oraić Rabušić (2018: 77-79) divides the group of the *reflexiva tantum* verbs to subcategories according to the following criteria: 1) whether the participants of an action marked by the verb are animate or inanimate and 2) the semantic roles of the predicates in the position of the main parts of the sentence. The semantic and syntactic principles of classification are intertwined in the mentioned categorization. Verbs researched in this work have been selected based on the semantic verb classification criterion i.e. further classification of the *reflexiva tantum* verbs is based on the semantic class<sup>91</sup> to which a psychological verb belongs. The psychological verbs selection criterion and the names for other semantic classes rely on the verb classification in the *Croatian Verb Valency database*. According to the classification in the *Croatian Verb Valency database* (Brač & Bošnjak Botica 2015), psychological verbs are the fifth most frequent semantic group of verbs with 34 groups in total. Along with the *Croatian Verb Valency database*, there is another comprehensive source of data on the verb valency, the *Croatian Valency Lexicon of Verbs* or *CROVALLEX* (Mikelić Preradović 2020). The verb classification in that lexicon is based on syntactic-semantic criteria and contains 180 semantic classes. Due to such a difference in the number of thematic classes, the same verb can sometimes belong to different thematic classes in the two classifications. The verbs that belong to the psychological verbs' class both in the *Croatian Verb Valency database* and the *CROVALLEX* are also found in other classes, such as the verb *zaljubiti se* 'to fall in love', which is classified as a *marriage verb* (Mikelić Preradović 2020: 59). The shared characteristic that makes psychological verbs a separate class is the description of the emotional state of a human being (Brač & Bošnjak Botica 2015). The psychological verbs' thematic class is specific in that it is mostly unique both by semantic and syntactic criteria i.e. most members of this class share the same thematic roles. All psychological verbs have one actant in common, which marks an entity, a human being, in the thematic role of an Experiencer, and another actant that marks the cause, which can belong to various categories from

---

<sup>91</sup> Semantic classes are groups of words with a common semantic constituent, which has a central position in the semantic structure, described using dictionary entries (Padučeva 2004: 43). The classes are formed according to the principle of consolidation of the common semantic characteristics and not by finding differences to form separate classes.

the non-linguistic reality (events, sounds, natural forces).<sup>92</sup> Psychological verbs are also often characterized by the thematic role of patient, which most usually describes the face to which the experiencer's feelings are transmitted (Padučeva 2004).

Psychological verbs can form regular polysemy relations<sup>93</sup> with other thematic classes of verbs. The verbs of perception are in regular polysemy relation with verbs of thinking, because, for example, the verbs *vidjeti* 'to see', *razmatrati* 'to consider', *sagledati* 'to perceive', can mark in different uses both the visual perception and the process of thinking about what has been perceived. The similarity between these meanings is established using a metaphor as a cognitive mechanism. According to Padučeva (2004), psychological verbs enter the regular polysemy relationships with 1) communication verbs and 2) mental-state verbs. Padučeva (2004: 537) holds the view that the mechanism of the regular polysemy works on both the synchronic and the diachronic level, i.e. that throughout the history of a language, the same changes of meaning which occur in the synchrony, have occurred on the diachronic level as well. This idea fits the perspective that synchrony and diachrony are not completely separate from each other, and that polysemic lexemes often determine the diachronic processes' intensity, as these processes can be more or less schematic (Raffaelli 2009: 141–143).

---

<sup>92</sup> In different theoretical approaches the semantic role of psychological verbs is termed differently and attitudes on the equivalence of the semantic roles of psychological verbs with different valency patterns are not compliant. The hierarchical polystratal depictions of verb arguments, typical for generative grammar, are more prone to make use of a greater number of semantic roles, with the aim of preserving the position equivalence in the syntactic structure, and the assigning of the semantic roles (Grimshaw 1990; Pesetsky 1995; Van Gelderen 2019). In functional approaches, the tendency to consolidate semantic roles and/or accepting the possibility that one argument can have more than one semantic role is predominant (Van Valin & LaPolla 1997; Padučeva 2004; Levin & Rappaport Hovav 2005).

<sup>93</sup> Apresjan (1974: 16) defines the regular polysemy of words as the following: «Polysemy of the word A with the meanings  $a_i$  i  $a_j$  is called regular if, in a given language, there exists at least one other word B with the meaning  $b_i$  and  $b_j$ , which are semantically distinguished from each other in exactly the same way as  $a_i$  and  $a_j$  and if  $a_i$  and  $b_i$ ,  $a_j$  and  $b_j$  are non-synonymous.» According to Padučeva (2004), the regular polysemy is a semantic relationship which does not only relate to words, but also to semantic classes.

### 3. Verb valency of reflexive verbs

The change caused by adding a reflexive marking to a verb can be accompanied by the reorganization of verb actants, which can lead either to a minimum change of meaning or to a complete change of verb meaning. The formal change by which the verb is necessarily used with a reflexive marking can highlight, in diachronic perspective, how verb arguments have changed (increase or the decrease in their number, change of the case of the argument, replacement of the bare case argument with a prepositional argument) or it can depict the derivation of the actant. Actant derivation is a model which describes the minimal change of the number and the composition of semantic roles, caused by adding morphological marking to the verb base or the root (Plungân 2011). According to Plungân (2011), it is sometimes difficult to distinguish between actant derivation and verbal aspect, because the morphological indicators of these two grammatical phenomena are often homonymic. Nevertheless, the core difference lies in the fact that a change in verb aspect also changes the pragmatic interpretation of the situation, and actant derivation changes the number and/or the composition of semantic roles. There are three different types of derivation: 1) valency-increasing derivation; 2) valency-decreasing derivation; and 3) interpretive derivation (Plungân 2001). A morphologically-marked valency change in Slavic languages can include all three types of derivation. Verb valency increase is expressed by a causative which is considerably less frequent than the formal decausative in Slavic languages. The causative meaning of the verb is determined in contrast to the non-causative meaning, i.e. the causative form enters a minimal pair with the non-causative.

Z. Gołąb (1968) highlighted four types of causative pairs' formation in the history of Slavic languages. The opposition of two verbs differing from each other by their suffixes, or by the apophony of the root (e.g. *\*piti* 'to drink' ~ 'to water, i.e. to cause to drink' *\*poiti*, *\*sěděti*/*\*sěsti* 'to sit' ~ *\*saditi* 'to make sit', *\*rasti* 'to grow' ~ *\*rastiti* 'to raise') is characteristic of the first type. The second type includes denominal verbs, i.e. the verbs derived from nouns or adjectives. Verbs which enter minimal pairs differ by their suffixes. The suffix *-ě-* denotes an intransitive verb of state or of change of state, and the suffix *-i-*, a transitive verb which marks a conscious action performed on an object (Gołąb 1968), e.g. *\*němъ* 'mute' → *\*oněmiti* 'to silence': *\*oněmēti* 'to lose the faculty of speech'; *\*zelenъ* 'green' → *\*zeleniti* 'to make green': *\*zelenēti* 'to become green'. The third type of the causative formation is the

dereflexivization of verbs with reflexive markings. Verbs having a reflexive marking have replaced the intransitive denominal verbs ending in *-ěti/-ějo*, with the reflexive verbs ending in *-iti se* (Gołąb 1968). The causative meaning is obtained when the reflexive marking is removed from the verbs which have been formed in this manner, e.g. *žena* ‘woman’ → *ženiti se* ‘to marry’ → *ženiti* ‘to marry’ + <A>. According to Gołąb (1980), this causative category is especially transparent in marking the psychological verbs in Slavic languages. In that formation type, the causative meaning has been formed subsequently, and the original meaning describes the state and follows the morphological form of the *reflexiva tantum* verbs. The argument for such an interpretation is, according to Gołąb (1968), not the criterion of the form, but of the meaning. Gołąb believed that the less complex meaning is diachronically the original one. The transitive and the causative meanings are considered to be more complex, because they have been derived from the meaning expressing a state by adding a semantic constituent ‘to cause’, and not vice versa. The syntactic-contextual causative is the fourth type, in which the verb becomes transitive and sometimes causative, if a noun in the accusative case is added to it. Several of the stated types are not productive in the contemporary Slavic languages. Plungân (2011) states that, in Russian, the morphological causative is more or less productively formed only from adjectives (*zelënyj* ‘green’ → *zelenit’* ‘to make green’); and several non-productive pairs have been preserved in the domain of verbal causatives (*gasnut’* ‘to fade’ ~ *gasit’* ‘to extinguish’, *sohnut’* ‘to become dry’ ~ *sušit’* ‘to make dry’), and even fewer unchanged pairs of the following type *gnit’* ‘to rot’ ~ *gnoit’* ‘to let rot’. The meaning of these derived morphological causatives is often idiomatic in the contemporary language, i.e. the derivational relationship can only be determined by an etymological analysis (Plungân 2011: 206). In Slavic languages, the reflexive pronoun *\*se* is used to morphologically indicate a decrease in the valency (Nedálkov & Sil’nickij 1970: 40), i.e. a valency-decreasing derivation. The syntactic/semantic differences in the reflexive and the non-reflexive pair, i.e. the valency-decreasing derivation, are a widely described phenomenon in the Slavic languages (Ānko-Trinickaâ 1962; Geniušienė 1987; Oraić Rabušić 2018; Knâzev 2007; Hrdlička 2020). The position of psychological verbs in the semantic/syntactic classification of the changes caused by reflexive marking being added to the verb, is the most important area of research in this work. In the cases when the verb used to be transitive and causative in the earlier period, a diachronic change of the verb valency pattern with regards to

the verb becoming a *reflexiva tantum*, will be interpreted as the end of the productivity of the depicted diathetic change. The reduced productivity of the described pattern shows that there has been a gradual disappearance of the productivity of causative patterns in the history of language, i.e. that there has been a gradual prevalence of valency-decreasing derivation.

## 4. Causativity

According to the lexical valency orientation, Slavic languages belong to the few linguistic groups in which detransitivization is prevalent (Nichols, Peterson & Barnes 2004). The basic or lexical valency orientation is a typological parameter by which languages differ from each other by their preferred model of derivation, responsible for the valency change. The researchers who have postulated this typological parameter (Nichols, Peterson & Barnes 2004) consider that in the lexicon of a language there are unmarked and frequent valency patterns, related to transitivity and causativity of verbs. Languages can be sorted in different categories according to the preferred valency pattern criterion. Eighteen semantic causative pairs in 80 languages have been studied within this research. One pair member has a plain meaning, and the other has an induced meaning. An induced meaning is the meaning which differs from the plain meaning only by the additional semantic causative component (Nichols, Peterson & Barnes 2004). The parameter relies on a dichotomy between the lexical meaning and the formal manner of its marking. The dominant language type, transitive or intransitive, shows which formal, morphological means are used to express the plain verb meaning, and thus also by which other formal, morphological means the induced, causative meaning is expressed. In Slavic languages, the plain meaning is expressed by the reflexive marking, which marks the valency reducing, i.e. the detransitivization. Based on the data acquired from this research, the validity of this thesis will be examined, as well as the diachronic stability of detransitivizing basic valency orientation of Croatian as a Slavic language. A drawback of the definition of the basic lexical valency orientation parameter (Nichols, Peterson & Barnes 2004) is the fact that the transitivity and causativity categories are not completely clearly distinguished. The need for such a distinction has proven to be necessary in this research paper in the context of describing the changes of the valency properties of verbs. A sample analysis has shown

that one group of verbs has changed its meaning and/or semantic category throughout time, and the other group has lost only its causative component. In other words, the inability to form a reflexive pair is followed exclusively by the loss of a causative meaning for some verbs, and not with any other changes of meaning. In order to distinguish between the cases of causative meaning loss, it is important to explain what exactly will be considered to be a causative meaning, i.e. determine which criteria the verb has to adhere to in order to be considered as causative. The loss of a causative meaning will be shown using the logical structure system of Role and Reference Grammar (henceforth RRG, Van Valin & LaPolla 1997). Even though the depiction manner is very elaborate in RRG, the criteria for showing that a verb meaning is causative have not been described clearly enough. Matasović (2018: 104) claims that within the current *Aktionsarten* system in RRG, each reflexive verb referring to a physical activity can be described as causative and that there are no clear criteria that would distinguish reflexive activity verbs from their causative counterparts. The presence of a force which is not inherent to the participant of an activity has been suggested as a distinguishing criterion. Additional criteria for distinguishing causative and reflexive verbs will be defined along with the aforementioned criteria. For such a distinction, a causative relations' classification will be used, as conceived by the Russian semanticist, E V. Padučeva. According to Padučeva (1994, 2004), there is a correlation between causativity and taxonomic verb categories. The term taxonomic category (T-verbs) is a variant of the term *Aktionsart*, of which the goal is to encompass the possibility of the formation of verbal aspectual pairs and their transition from one thematic class to another. Thematic classes are determined according to the characteristics: 1) dynamics; 2) terminativity; 3) controllability, and 4) telicity (Padučeva 1996: 197). Padučeva (1994) highlights that causative relations can be analysed only when a verb belongs to the taxonomic categories of agentive accomplishment (*dejstviâ obyčnye*) and agentive achievement (*dejstvie s akcentom na rezul'tate*)<sup>94</sup> Padučeva conceptualized this as follows (1994: 2): the meaning of all verbs from the agentive category (both agentive accomplishment and agentive achievement) can be explained by the

<sup>94</sup> The category of agentive accomplishment is characterized by the following characteristics: [+dynamics], [-terminativity]; [+controllability], [+telicity], and the category of agentive achievement by [+dynamics], [+terminativity]; [+controllability], [+telicity]. Terminativity is equivalent to punctuality as described in Van Valin & LaPolla (1997).

following group of constituents: 1) the causator | Agent acted with a Goal; 2) causation | that caused; 3) a result matching the Goal. The three criteria which have to be fulfilled in order for the verb to be considered causative are stated below, and they are needed so that once this has been established, its predicate composition can be represented using the operator CAUSE.

(A) Transitivity. In this work, transitivity will be defined as a binary category, and not as a scalar category. Such a simplified understanding of the transitivity category has been chosen, so that a higher level of generalization can be achieved. However, one example of using a verb with an accusative argument will not be used as an example of verb transitivity, and that is the use of a verb with a cognate object. An internal or a cognate object is a ‘highly-borderline transitivity type’ (Belaj & Tanacković Faletar 2017) and cannot be considered a real argument, as it is only an adjunct, because a manner interpretation is characteristic for it (Birtić & Matas Ivanković 2009).

(B) Difference between the referents. If in a logical decomposition there are two participants of a situation, *x* and *y*, in order to make possible for the verb’s semantic structure to be represented by the operator CAUSE, the first actant *x* cannot be contained in the second argument’s composition, i.e. in the logical structure of the fundamental sentence predicate.

- (1) a. *Marija je naučila hrvatski.*  
 b. ‘Marija has learned Croatian.’  
 c. \* [do’ (**Marija**, Ø)] CAUSE [BECOME learned’ (**Marija**, Croatian)]
- (2) a. *Marija je naučila svoje studente hrvatski.*  
 b. ‘Marija has taught her students Croatian.’  
 c. [do’ (Marija, Ø)] CAUSE [BECOME learned’ (students, Croatian)]

(C) The verb has to belong to the taxonomic categories of agentive accomplishment and agentive achievement.

## 5. Goals and methodology

The extent to which psychological *reflexiva tantum* verbs are diachronically stable will be determined in this work, as well as whether the Croatian language has changed its basic valency orientation, i.e. whether it has been detransitivized to the same degree throughout its history. It will also be

analysed whether the causative alterations of psychological verbs were more frequent in the earlier linguistic periods of Croatian. The relation of verb valency change and the verb meaning in the diachronic approach to the verb valency can, according to M. Samardžija (1994), fall into one of the three following categories: 1) the meaning of the verb changes, but this does not affect the valency; 2) the valency change occurs without a change of the meaning; 3) the valency and the meaning change simultaneously in that a verb has one valency in one meaning, and another valency in another meaning. In this work, three additional parameters will be studied along with the relation of the change of the meaning and of the arguments: transitivity, causativity and the presence of a reflexive marking with a verb. As most verbs show only slight changes of meaning with different arguments, the change of meaning criterion is somewhat extended. Smaller meaning expansion or narrowing will not be considered a change of meaning, as whether a verb has changed its thematic class will be looked into. If the meaning of the verb has changed only in that the verb no longer has a causative component, while it remains a part of the same thematic class, the meaning will not be considered to have changed. Additionally, the cases of regular polysemy, when psychological verbs switch to the classes of mental verbs or of verbs of speaking, will also not be considered a change of meaning. Verb transitivity will be analysed as a binary opposition, and the causativity criteria have been described in the previous chapter. In Table 1, all logical possibilities are stated, considering the 1) change of the presence of the reflexive marker *se* (RM); 2) thematic class to which the verb belongs (SC); 3) the morphosyntactic properties of arguments (MS); 4) transitivity (T); 5) causativity (C). These parameters are used for the typology of change and will be analysed in this work. One of the criteria for defining causativity is transitivity, suggesting that a combination where the verb is both causative and intransitive will not be found. Therefore, eight possible combinations are automatically excluded. These combinations are marked with asterisk (\*).

Table 1. Possible combinations of the values of the parameters for valency change

№	RM	SC	MS	T	C
1.	+	+	+	+	+
2.	+	+	+	+	-
3.*	+	+	+	-	+
4.	+	+	+	-	-
5.	+	+	-	+	+
6.	+	+	-	+	-
7.*	+	+	-	-	+
8.	+	+	-	-	-
9.	+	-	+	+	+
10.	+	-	+	+	-
11.*	+	-	+	-	+
12.	+	-	+	-	-
13.	+	-	-	+	+
14.	+	-	-	+	-
15.*	+	-	-	-	+
16.	+	-	-	-	-
17.	-	+	+	+	+
18.	-	+	+	+	-
19.*	-	+	+	-	+
20.	-	+	+	-	-
21.	-	+	-	+	+
22.	-	+	-	+	-
23.*	-	+	-	-	+
24.	-	+	-	-	-
25.	-	-	+	+	+
26.	-	-	+	+	-
27.*	-	-	+	-	+
28.	-	-	+	-	-
29.	-	-	-	+	+
30.	-	-	-	+	-
31.*	-	-	-	-	+
32.	-	-	-	-	-

The research of Croatian from the earlier period is difficult due to a lack of language resources. Currently a corpus and a dictionary of Old-Croatian are not available online. This lack of materials is avoided by analysing Croatian in comparison to the earlier periods of other Slavic languages and by using the available lexicographic materials, which include the Old-Croatian sources. The main source was the Dictionary of Croatian or Serbian language (henceforth ARj). The earlier linguistic period of Croatian, studied as part of this work, is equivalent to the first and second period of pre-standard Croatian, as defined in Brozović (2008). The first pre-standard period includes the period from the 9th/10th century to the end of the 15th century, and the second includes the 16th century (Brozović 2008). The meaning of words as recorded in the Croatian sources prior to the 16th century was issued from the ARj. Other than that, and taking into consideration that Slavic languages had been linguistically more similar to each other in the earlier periods, the forms and meanings of etymologically equivalent verbs in Old East Slavic, Old Czech, and Old Polish have also been considered. The dictionaries and the corpora for the mentioned periods contain textual sources written prior to the 15th century. In the dictionaries of Proto-Slavic (Etymological Dictionary of Slavic Languages: Proto-Slavic Lexical Stock (henceforth ÈSSÂ), Proto-Slavic Dictionary (henceforth SP), Etymological Dictionary of the Slavic Inherited Lexicon (henceforth EDSL), and Etymological Dictionary of Croatian Language (henceforth ERHJ)) it is possible to check only the etymological matches, but it is not possible to check the data on transitivity and valency patterns, because the reconstructed forms and verb meanings are attested in some Slavic languages, but the examples of their use are not. The examples of use are prerequisites for a contextual analysis, which is, along with the lexicographic sources in the diachronic semantics, the core of the methodology of reconstruction, because it removes the researcher's subjectivity, as the person doing the research can easily include their own knowledge of the world in the linguistic data from the distant past (Raffaelli 2009: 185-189). Due to this reason, the stated dictionaries of Slavic languages and the historical sub-corpus of *The Russian National Corpus* have been used as additional sources along with the etymological dictionaries. In the Old Polish, Old East Slavic, and Old Czech dictionaries, it was first checked whether the attested verb forms have a reflexive marking and if they do not, it was checked whether they had been transitive, and whether they used to have a causative meaning. The attestation of transitive forms is necessary in order to establish whether the Slavic languages in the earlier linguistic

periods allowed for causative forms of verbs corresponding to contemporary *reflexiva tantum* verbs in Croatian.

The sample of 30 *reflexiva tantum* verbs was collected by an elimination method. In order for the verbs to be listed in the sample, they had to fit two criteria: 1) the semantic criterion, i.e. the verb had to belong to the group of psychological verbs; 2) the diachronic criterion, by which it is understood that there had to be records of the verb in the selected sources from the earlier linguistic periods. In the first step, the *reflexiva tantum* verbs were extracted, and afterwards the selection of psychological verbs were performed on this group. As two distant time periods are compared in this work, both periods can equally be used as starting points to select the *reflexiva tantum* verbs. A verb which appears in the earlier period (marked as  $t_1$ ) as a *reflexivum tantum* verb is not necessarily such in the contemporary language (the contemporary language or a more recent period language state will be marked as  $t_2$ ) and vice versa. A greater distribution of *reflexiva tantum* verbs in  $t_1$  or  $t_2$  has not been assumed previously. For that reason, the verbs have been extracted from two distant time periods in different states of the language.

In order to bypass the problem of the non-availability of a corpus and of an Old Croatian dictionary, the indices of two Old East Slavic dictionaries – Dictionary of the Russian language XI-XVII centuries (henceforth SRÂ) and Materials for the Dictionary of the Old Russian Language in the Written Records (henceforth MSR) – have been used as the starting points for excerpting verbs as they had been in the earlier linguistic periods. From these dictionaries all verbs which have their etymological equivalent in the contemporary Croatian language have been extracted. The etymological equivalence has been checked by examining three dictionaries of the Proto-Slavic language and the etymological dictionary of other Slavic languages (ÈSSÂ, SP, EDSL, and ERHJ). The extracted verbs have then been compared to the verbs found in two contemporary dictionaries of Croatian – Large Dictionary of the Croatian Standard Language (henceforth VRH) and the Croatian Language Portal (henceforth HJP) –, in order to check whether a certain verb belongs to the category of *reflexiva tantum* verbs.

Taking into account the possibility that a certain verb found in Croatian is not also necessarily present in Old East Slavic, a list of *reflexiva tantum* verbs has been taken from Ivana Oraić Rabušić's work (2018), and the list has been compared with the three Proto-Slavic dictionaries. In case the verbs were not alphabetically listed in dictionaries, it was checked whether

a certain verb occurs in the ARj in the pre-standard period of the Croatian language. For each verb, its reconstructed form was taken from the three Proto-Slavic dictionaries, and then its meanings have been taken from the Old East Slavic dictionary (SRĀ), and in case there were no records of the verb there, it was further searched in the MSR. The next step was the inclusion of the data extracted from the Old Polish and Old Czech sources. A more detailed description of the meanings and arguments is obtained from the Old Croatian verbs taken from the ARj. In case a verb has been analysed in the Croatian Verb Valency database (henceforth BHGV), the data of the BHGV have been included, and if no valency analysis was available, the meaning and the valency characteristics have been extracted from VRH and HJP. If a verb was not found in a dictionary, that dictionary's title has not been noted.

## 6. Results and analysis

In this section, results and analysis of 30 examined verbs are presented. For every verb, characteristics based on five parameters are provided first (reflexive marking (RM), semantic class (SC), morphosyntactic properties of arguments (MS), transitivity (T), and causativity (C)). Then, the reconstructed form of the verb is noted, and data from Old East Slavic, Old Czech, and Old Polish are presented. A detailed overview of verb valency is given only for Old Croatian. Imperfective verbs are listed first, followed by perfective verbs.

### A) Imperfective verbs

#### 1. BOJATI SE ('TO FEAR')

RM: –	SC: –	MS: +	T: –	C: –
-------	-------	-------	------	------

ÈSSĀ: \**bojati* (sę) / SP: \**bojati* sę / EDSL and ERHJ: \**bojati* sę. / In the SRĀ, OCZ and SS, neither a transitive nor a causative use of the verb has been recorded. / ARj: *Bojati se*: 1) to feel fear: a) Ø; b) <G>, the argument marks the cause of the fear; c) *od* 'from' + <G>; d) <D>; e) *za* 'for' + <A>; f) <Inf>; g) <Sent>. / BHGV: 1) to feel fear: a) Ø; b) <G>; <Sent>. 2) to fear for one's safety: a) *za* 'for' + <A>. 3) to not dare: a) <Inf>. 4) to assume something feeling afraid: a) <Sent>.

In Old Croatian, the fear for the other person's wellbeing could be expressed with dative and prepositional argument *za* 'for' + <A>, and in the contemporary language only the prepositional argument has remained. The cause was expressed by the genitive or a prepositional argument *od* 'from' + <G>. L. Hudeček (2003) has shown, using the Croatian language materials from the 17th to the middle 19th century that argument *od* 'from' + <G> accompanying verbs of speaking and thinking had been widespread in all Croatian varieties. This argument came to Croatian due to a Latin and Italian influence (Hudeček 2003). It is possible that this argument had first appeared even earlier than the 17th century and not only with verbs of speaking and thinking. The use without a reflexive marker in Old Polish and Old Croatian is possible exclusively with participles, but even when used like that, the verb is not used in transitive nor in the causative meaning.

## 2. DIVITI SE ('TO ADMIRE')

RM: +	SC: -	MS: +	T: +	C: +
-------	-------	-------	------	------

ÈSSÂ: \**diviti* (*se*) / SP: \**diviti* / ED SL: \**diviti* / ERHJ: *diviti se* / SRÂ: *Diviti*: 1) to be wondering (12th-13th century): a) Ø. 2) to make one wonder: a) <A>. / ARj: *Diviti se*: to admire a) <D>, the argument expresses the cause of admiration; b) *o* 'about' + L; c) <Sent>. / BHGV: *Diviti se* 'to admire': 1) to be in awe by something beautiful or proportional: a) <D>. 2) to feel and/or express respect or admiration: a) <D>; b) <Sent>. 3) to express by words that something is considered beautiful or proportional: a) Ø.

By its origin, the verb is a denominal causative ending in *-iti*, derived from the nouns \**divo* 'a wonder' (ERHJ I: 171). In SP, it is stated that the original meaning in Proto-Slavic has been to 'look at/watch someone with admiration' and that the verb had a causative meaning 'to make one awestruck'. Both meanings have been reconstructed, whereby the meaning 'to look, to watch someone with admiration' has not been found in the researched lexicographic sources, and the causative meaning has been confirmed in Old East Slavic. In Old East Slavic the verb had to express the state without the obligatory reflexive marker, and it could be both reflexive and causative. In this case, exclusively a causative constituent can be added to its meaning. In Old Croatian, the verb could have one additional prepositional argument (*o* 'about' + <L>).

## 3. DURITI SE ('TO SULK')

RM: +	SC: +	MS: +	T: -	C: -
-------	-------	-------	------	------

ÈSSÂ: \**duriti* (*se*) / SP and ERHJ: \**duriti* / EDSL: \**duriti* / SRÂ: *Durit'*: 1) to be crazy: a) Ø. In Old East Slavic, only the intransitive use has been confirmed, but in contemporary Russian, the verb is transitive and it means 'to cheat', 'to lie' + <A>. / VRH: *Duriti se*: 1) to interpret one's words as an insult and to act in accordance with it: a) Ø; b) *na* + <A>.

The verb is denominal according to its word formation process, formed by the suffix *-i-* from the adjective \**durǔ* 'ugly' (ERHJ I: 208). According to the data from the SP, the verb appears in a transitive and causative meaning from the 18th century in Polish in the meaning 'to deceive' 'to trick' + <A>. The oldest confirmation is found in Old East Slavic. In Old East Slavic the verb does not have a reflexive marking, but it is not transitive nor causative.

## 4. GNUŠATI SE ('TO ABHOR')

RM: +	SC: -	MS: +	T: +	C: +
-------	-------	-------	------	------

SP: \**gnōšati* / EDSL: \**gnōšati* / ERHJ: \**gnōšati* (*se*) / SRÂ: *Gnušati*: 1) to detest: a) <A>. *Gnušatisâ*: 1) to detest. / VRH: *Gnušati se*: 1) to be disgusted: a) <N> + V + <D>; b) <N> + V + *nad* 'by' + <I>. 2) to abhor: a) <N> + V + <D> + <A>.

The Proto-Slavic verb \**gnušati* (*se*) is derived from the verb \**gnusiti* (*se*) (ERHJ I: 279). The verb \**gnusiti* is derived from the noun \**gnūsǔ* 'dirt', 'a nauseating thing' using the suffix *-i-* (ibid.). In Old East Slavic, the causative and transitive meaning of the verb *gnušati* (*se*) 'to abhor' have been confirmed. In contemporary Croatian, only the iterative verb *gnušati se* has remained, while the original form *gnusiti* has not been preserved.

## 5. GROZITI SE ('TO DETEST')

RM: +	SC: -	MS: +	T: +	C: +
-------	-------	-------	------	------

ÈSSÂ and SP: \**groziti* / EDSL: \**groziti* / ERHJ: *groziti* (*se*) / SRÂ: *Groziti*: 1) to threaten: a) <D>. 2) to scare: a) <A>. *Grozitisâ*: 1) to threaten someone: a) <I>. 2) to abhor: a) <G>. 3) to get angry a) *na* 'with' + <A>. / OCZ: *Hroziti*: 1) to scare: a) <A>. 2) to be in fear: a) Ø, 3) to threaten: a) <D> <I>. *Groziti se*: 1) to be afraid of: a) Ø. / SP: *Grozić*: 1) to scare: a) <A>. *Grozić się*: 1) to

threaten. / ARj: *Groziti*: 1) to threaten. 2) to cry: a) Ø; b) <A>, a cognate object. 3) to detest: a) <A>. *Groziti se*: 1) to be disgusted, to be afraid. / VRH: *Groziti se*: 1) to feel great fear, to fill with discomfort and fear: a) <G>. 2) to express threats: a) <D>.

The verb was formed from the noun \**groza* ‘dread’, with a causative suffix (ERHJ I: 305). The transitive and causative meaning has been confirmed in all the earlier periods of the analysed Slavic languages. Additionally, in all the found results, the difference in the meaning is achieved using different arguments. As a psychological verb, it is used with the accusative case, and in this case the verb has a causative meaning, and when it is used with the dative, it has the meaning of a verb of speaking. In Old Croatian, the verb might have had the meaning ‘to cry’ and it had an internal object. In Old Croatian, the verb was used without a reflexive marking in the meaning ‘to threaten’, as in the other Slavic languages, which means that its use with a reflexive marking with the same meaning is from a more recent period.

## 6. KAJATI SE (‘TO REGRET’)

RM: +	SC: –	MS: +	T: +	C: –
-------	-------	-------	------	------

ÈSSÂ: \**kajati* (*se*) / EDSL: \**kàjati* / ERHJ: \**kàjati* (*se*) / SRÂ: *Kaâti*: 1) to reprimand: a) <A>. 2) to confess: a) <A>. *Kaâtisâ*: to regret. / OCZ: *Kâti sě*: to regret. / SP: *Kajać* (*się*): 1) to be sorry for one’s sins 2) to admit one’s sins, to confess. It appears only as a participle without the reflexive marking. / ARj: *Kajati*: 1) to confess: a) <A>. 2) to lament: a) <A>. *Kajati se*: 1) to regret one’s bad doing: a) Ø; b) *od* ‘from’ + <G>; c) <G>; d) *za* ‘for’ + <A>; e) <Sent>; f) <Inf>. / VRH: *Kajati se*: 1) to regret one’s bad doing: a) *zbog* ‘because of’ + <A>.

The verb *kajati se* is derived from the Indo-European root \**k<sup>w</sup>ōy-* (ERHJ I: 416). In the analysed sources, only in Old East Slavic there are confirmations of a transitive, but not of a causative use. In its meaning ‘to confess’ in Old East Slavic and Old Croatian, the verb may have had a cognate object (a sin, an evil doing), which would make it an intransitive verb. In these examples of its use, the verb acquires the meaning of a verb of speaking, to which its arguments also point. It seems that in the earlier language period, the verb could have had more arguments.

## 7. NADATI SE ('TO HOPE')

RM: -	SC: -	MS: +	T: -	C: -
-------	-------	-------	------	------

ÈSSÂ: \**nadati* (*se*) / SRÂ: *Nadeâti*: 1) to count on someone, to rely on: a) *na* + <A>. *Nadeâtisâ*: 1) to hope. / ARj: *Nadati se*: 1) to expect the desired will be fulfilled: a)  $\emptyset$ ; b) <G>; c) <D>; d) <A>; e) <N> + V + <D> + *od* 'from' + <G>; f) *u* 'in' <A>; g) <Inf>; h) <Sent>. / BHGV: *Nadati se*: 1) to hope that one's wishes will come true: a) <N> + V + <D>; b) <N> + V <Sent>; c) N + V + <I>; d) <N> + V + *u* 'in' <A>. 2) to expect: a) <N> + V + <D> + *od* 'from' + <G>.

The verb *nadati se* is a denominal verb, formed from the noun \**nada* 'hope'. It has not changed its *reflexiva tantum* class. One example of use without a reflexive marking has been found in Old East Slavic, but it is from a more recent period and it neither has a transitive nor a causative meaning.

## 8. PONOSITI SE ('TO BE PROUD OF')

RM: +	SC: +	MS: +	T: +	C: -
-------	-------	-------	------	------

SRÂ: *Ponositi*: 1) to swear: a) <N> + V + <D> + <A>. 2) to reproach, to chastise, to reprimand: a) <D>. 3) to slander a) *na* + <A>, 4) to carry for a short period of time: a) <A>. *Ponositisâ*: 1) to winnow, *in the wind*. 2) to spread around. 3) to be proud of. / ARj: *Ponositi*: 1) to carry for a short period of time: a) <A>. *Ponositi se*: 1) to treat/ behave to: a) <I>. 2) to take pride: a)  $\emptyset$ ; b) <I>; c) *u* 'in' <L>, d) <Sent>. / BHGV: *Ponositi se*: 1) to feel pride because of someone/something: a) <N> + V + <I>; b) <N> + V + <Sent>.

The verb is etymologically an imperfective form of the verb *ponijeti se* (ERHJ II: 168); Nowadays it is an orphan imperfective (Polančec 2018: 123). The records of this verb have been found in Old East Slavic and, without the reflexive marking, the verb has the meaning of a speaking verb 'to chastise', 'to reprimand', and of a verb of 'carrying'. When the verb is in the category of verbs of speaking, it has an argument in dative, similar to the verb *groziti* 'to abhor'. In Old Croatian the verb can also belong to the category of verbs of speaking, and that use is also marked by an argument in dative. The additional meaning, which does not appear in Old East Slavic, *postupati* 'to treat' has a similar meaning to the verb *ponašati se* 'to behave'. The reason for that similarity can be a formative and semantic relation of the verb *ponositi se* 'to be proud of' and *ponašati se* 'to behave', which was still visible in the earlier period language. The verb was transitive in Old East Slavic and Old Croatian, but it was not causative.

## 9. SNEBIVATI SE ('TO WONDER')

RM: +	SC: +	MS: +	T: +	C: -
-------	-------	-------	------	------

ARj: *Snebivati*: 1) to die, to wither a) Ø. 2) to ruin, to bring down: a) <A>, the face doing the action is inanimate. / *Snebivati se*: 1) to wonder: a) <D>; b) *na* 'on' <A>. 2) to disappear, to deteriorate, to waste: a) Ø. 3) to hesitate, a) Ø. VRH: *snebivati se*: 1) to be overly surprised: a) <D>.

The verb appears exclusively in Old Croatian, and not in other Slavic languages analysed in this work. Its meaning changed with time, as it had been used as a verb of integration and disintegration, similar to the verb *snebiti se*. It was a transitive verb and it had several arguments in the earlier period language. In contemporary Croatian language, it kept only its psychological meaning.

## 10. SRAMITI SE ('TO FEEL ASHAMED')

RM: +	SC: -	MS: +	T: +	C: +
-------	-------	-------	------	------

SRĀ: *Sramiti*: 1) to embarrass: a) <A>, 12th century. *Sramitisâ*: 1) to embarrass oneself: a) <I>; b) <G>. / ARj: *Sramiti*: 1) to embarrass: a) <A>. *Sramiti se*: a) <G>. B.) *od* 'in front of' + <I>; c) <G>; d) <Inf>; e) <Sent>. / BHGV: To feel ashamed: 1) to feel shame: a) Ø; b) *pred* 'in front of' <I>; c) <G>, d) *pred* 'in front of' <I>, e) <Inf>; f) <Sent>. 2) to not dare to do something out of shame: a) <Inf>.

The verb is denominal, formed from the noun \**sôrmъ* 'shame' and the suffix *-i-* (ERHJ II: 378). In Old East Slavic, the verb could be both transitive and causative. The verb did not change its meaning through time.

## 11. SVIÐATI SE ('TO LIKE')

RM: +	SC: +	MS: +	T: +	C: -
-------	-------	-------	------	------

SRĀ: *Svidatisâ*: 1) to encounter one another (17th century) / SS: *Zwiedzać*: 1) to watch someone: a) <A>. / ARj: *Sviđati*: 1) to perform, to do: a) <A>. 2) to understand, to realize: a) <A>. *Sviđati se*: 1) to appear a certain way: a) <N> + V + <D> + <A>. 2) to become aware of something: a) Ø. 3) to be likeable to someone: a) <N> V <D>. / BHGV: *Sviđati se*: 1) to make a positive impression on someone, to be likeable to someone: a) <N> + V + <D>; b) <D> + V + <Adv>; c) <D> + V + <Sent>; d) <D> + V + <Inf>. 2) to become attractive to someone: a) <N> + V + <D>.

The verb *svidjati se* is an imperfective pair to the verb *svidjeti se* (ERHJ II: 413). In Old Polish and Old Croatian, the meaning relationship with the verb *vidjeti* 'to see' has remained preserved. Its earliest confirmed appearance in Russian is very late and dates from the 17th century, the reflexive marking as a verb constituent has a reciprocal meaning, and the relation to the verb *vidjeti* 'to see' is also noticeable. Its semantic class changed in the way that the verb could have the following meanings: 'to seem', 'to watch', 'to understand' and 'to feel'. The verb follows the same valency pattern characteristic for its contemporary use, where it takes the meaning 'to like'. In Old Polish and Old Croatian, the verb was transitive, but it was not causative.

## 12. STIDJETI SE ('TO FEEL ASHAMED')

RM: +	SC: -	MS: +	T: +	C: +
-------	-------	-------	------	------

SRĀ: *Styditi*: 1) to embarrass: a) <A>. *Styditisâ*: 1) to feel ashamed. / ARj: *Stidjeti*: 1) to embarrass: a) <A>. *Stidjeti se*: a) Ø; b) <G>; c) *od* 'from' <G>; d) *pred* 'in front of' <I>; e) <Inf>; f) <Sent>. / BHGV: *Stidjeti se*: 1) to be ashamed: a) Ø; b) *pred* 'in front of' + <I>; c) <G>; d) <Inf>; e) <Sent>. 2) to not dare to do something because of shame: a) <Inf>. 3) to hide one's private matters: a) Ø.

As in Czech, the verb was formed from the noun *\*stydvĕ*, by adding a suffix *-i-* to it (ERHJ II: 389). In Old Czech and Old Polish, the verb has not been found in the sources, but it has in Old East Slavic, however, there are no records earlier than the 17th century. The earlier and more frequent was the verb which had a reflexive marking. In Old Croatian, one causative and transitive use of the verb has been found in the 16th century. The verb did not change its meaning, but there has been a confirmed causative and transitive use in the earlier period.

## 13. UZDATI SE ('TO RELY ON SOMEONE')

RM: +	SC: +	MS: +	T: +	C: +
-------	-------	-------	------	------

ERHJ: *\*vĕzdati se* / SRĀ: *Vozdati*: 1) to give a) <A>. 2) to do, a verb of incomplete meaning which is used with nouns to express the meaning of some action: a) *vozdati vĕst'* 'to give the news', to inform; b) to give revenge, to avenge. *Vozdatisâ*: 1) to surrender: a) Ø. / OCZ: *Vzdāti*: 1) to give a) <A>. *Vzdāti sĕ*: 1) to surrender: a) <D>. / SS: *Wzdać, wezdać*: 1) to gift: a) <N> + V

+ <D> + <A>. 2) to return: a) <N> + V + <D> + <A>. 3) to give: a) <N> + V + <A> + <D>. 4) to make, a verb of incomplete meaning which is used with nouns to express the meaning of some action: a) *wzdać modłø*. 5) to mark, to allot: a) <N> + V + <D> + <A>. / ARj: *Uzdati*: 1) to give: a) <N> + V + <D> + <A>. 2) to do. 3) to allow a) <A>; b) <Inf>; c) <D>. 5) to have trust in someone, to believe: a) *u* 'in' <A>; b) <Inf>. 6) to dare: a) <Inf>. *Uzdati se*: 1) to rely on, to believe in: a) Ø; b) <D>; c) *od* 'from' + <A>; d) *u* 'in' + <A>; e) <Inf>. 2) to hope, to expect a) <Inf>; b) <Sent>. / VRH: *Uzdati se*: 1) to have trust: a) *u* 'in' + <A>.

The verb was formed from the prefix \**vbz* and the verb \**dàti* (ERHJ II: 532). It has been found in the records of all analysed Slavic languages in their earlier linguistic periods. Its meaning is extremely extensive, and it can even occur as a verb of incomplete meaning, when it has the meaning 'to do something'. The verb might have appeared without a reflexive marking in a very early linguistic period, with the argument *u* 'in' + <A>. The verb could have other meanings when combined with an infinitive argument such as 'to expect', 'to dare', which have not been preserved in contemporary Croatian. An infinitive argument is probably responsible for the modification of the meaning 'to dare', because it is characteristic for the modal verbs *drznuti se*, *usuditi se*, *osmjeliti se*, all meaning 'to dare'.

## B) The perfective verbs

### 14. DRZNUTI SE ('TO DARE')

RM: +	SC: -	MS: -	T: -	C: -
-------	-------	-------	------	------

ÈSSÂ: \**dbrznŋti* / EDSL: \**dbrznŋti* / In SRÂ, OCZ, SS and ARj: *Drznuti se*: 1) to dare: a) <Inf>. / VRH: *Drznuti se*: 1) to dare: a) <Inf>.

According to the ÈSSÂ, the verb was formed from the adjective \**dbrzŋ*, by adding a suffix *-ŋ-* to it (ie. \**-neŋ-*). It is closely related to Old Indo-Aryan verb *dṛṣ-*, √5., *dṛṣŋoti* by its meaning and formation process. The verb has been recorded in all the analysed Slavic languages with the same meaning and the same argument. Furthermore, in all of them it had the function of a modal verb and it did not have a reflexive marking. The Croatian *reflexivum tantum* verb is undoubtedly of a more recent origin. It is possible that it had started appearing as a *reflexivum tantum*, analogous to the verb *usuditi se*, which has a similar meaning. The verb did not have a transitive nor a

causative meaning, the infinitive argument has remained unchanged, and the only difference with regards to the earlier state is in the obligatory reflexive marking.

### 15. NASLADITI SE ('TO GLOAT')

RM: +	SC: -	MS: +	T: +	C: +
-------	-------	-------	------	------

SRĀ: *Nasladiti*: 1) to make sweet: a) <A>. 2) to fill with delight: a) <A>. *Nasladitisâ*: 1) to eat something sweet, to eat a lot of sweets: a) <G>. 2) to feel joy. / OCZ: *Nasladiti*: 1) to make sweet: a) <A>. / ARj: *Nasladiti*: 1) to make joyful, sweet: a) <A>. *Nasladiti se*: 1) to feel joy, delight: a) <I>; b) *u* 'in' + <L>; c) *od* 'from' + <G>. / VRH: *Nasladiti se*: 1) to experience delight: a) Ø. 2) to enjoy someone's misfortune: a) <I>.

The verb is composed from the prefix *\*na* and the verb *\*solditi* 'to make sweet' (ERHJ I: 677). In Old Polish this verb is not found, but in Old Czech it exists, and is used exclusively as a transitive and a causative verb. In Old East Slavic, the verb appears as a transitive and a causative verb, but also as a reflexive. In Old Croatian it can be found as a causative, transitive and a reflexive verb. The concrete use in Old Croatian with the meaning 'to taste something sweet' has not been recorded. The shift in meaning went from concrete to abstract, from 'taste something sweet' to 'enjoy', and according to ERHJ it had occurred in the Proto Slavic language.

### 16. OKAMENITI SE ('TO FREEZE IN SHOCK')

RM: +	SC: -	MS: +	T: +	C: +
-------	-------	-------	------	------

SRĀ: *Okameniti*: 1) to become hard as a rock: a) <A>. *Okamenitisâ*: 1) to get petrified: Ø. 2) to lose all feelings; 3) to become immobile. / OCZ: *Okameněti*: 1) to turn into stone. / ARj: *Okameniti*: 1) to turn into stone: a) <A>. *Okameniti se*: 1) to turn into stone: a) Ø. 2) to become emotionless. 3) to become immobile: a) *od* 'from' + <G>. / VRH: *Okameniti se*: 1) to turn into stone: a) Ø. 2) to stiffen from being overcome with feelings: a) Ø. 3) to stop changing: a) Ø. 4) to lose liveliness and creativity: a) Ø. 5) to be astonished, to be surprised a) Ø; b) *od* 'from' + <G>.

In Old East Slavic and Old Croatian, the verb was transitive and causative and it was used both with the literal and metaphorical meaning. There are no records of the verb having a metaphorical meaning in transitive use. The verb marks an emotional state exclusively in its reflexive use.

## 17. OSLADITI SE ('TO ACQUIRE A TASTE FOR SOMETHING')

RM: +	SC: -	MS: +	T: +	C: +
-------	-------	-------	------	------

SRĀ: *Osladiti*: 1) to make pleasant: a) <A>. 2) to bring joy a) <A>. *Osladitisâ*: 1) to become sweet. / OCZ: *Osladiti*: 1) to make sweet: a) <A>. *Osladiti sě*: 1) to become sweet. / SS: *Oslodzić*: 1) to make something sweet. / ARj: *Osladiti*: 1) to make something sweet: a) <A>. *Osladiti se*: 1) to become sweet: a) Ø. 2) to get enticed by: a) *na* 'on' + <A>. / HJP: *Osladiti se*: 1) to eat something sweet: a) <N> + V + <I>. 2) to like something: a) <N> + V + <D>.

The verb has been found in all earlier periods of the analysed Slavic languages. It was used as a transitive and a causative verb. Its meaning as a psychological verb 'to like something' is certainly from a more recent period of the Croatian language. The meaning development is assumed to be as follows: 1) *nasladiti* 'to make something sweet/lovely' → 2) *nasladiti se* 'to become sweet' → 3) *nasladiti se* 'to like something'. In the second meaning under it has the same valency frame as the verbs *svidjeti se*, *sviđati se* 'to like'.

## 18. POKAJATI SE ('TO REPENT')

RM: +	SC: -	MS: +	T: +	C: -
-------	-------	-------	------	------

ERHJ: \**pokajati se* / SRĀ: *Pokaâti*: 1) to confess: a) <A>, the object is animate. *Pokaâtisâ*: 1) to repent. / OCZ: *Pokâti sě*: 1) to repent: a) <G>. 2) to confess one's sins, to confess. / SS: *Pokajać się*: 1) to repent: a) <G>. / ARj: *Pokajati se*: 1) to feel guilt and remorse because of something bad one has done: a) Ø; b) <G>; c) *od* 'from' + <G>; d) *za* 'for' + <A>; e) <Sent>. / VRH: *Pokajati se*: 1) to feel guilt and remorse because of something bad one has done.

The verb is made of the prefix \**po* and the verb \**kajati se* 'to regret' (ERHJ II: 156). A transitive use has only been recorded in Old East Slavic, as late as the 15th century, and only for one example. This points to the fact that the transitive use was not the primary one. In all of the earlier linguistic periods of the analysed languages, it is possible to have an argument in the genitive case with the reflexive verbs. However, in the examples found in the sources, there is often a noun which marks the content of the spoken words ('a sin'), i.e. the cognate object. When accompanied by the cognate object 'a sin', the verb transitions into the semantic class of the verbs of speaking.

## 19. PRIČINITI SE ('TO SEE SOMETHING WHICH IS NOT THERE')

RM: +	SC: +	MS: +	T: +	C: +
-------	-------	-------	------	------

SRĀ: *Pričiniti*: 1) to add something: a) <A>. 2) to make, to build: a) <A>. *Pričinitisâ*: 1) to get dressed: a) *v* 'in' + <A>. 2) to happen. / OCZ: *Přičiniti*: 1) to get close: a) <N> + V + <A> + *k* 'to' + <D>. 2) to add: a) <N> + V + <A> + <D>. 3) to give: a) <N> + V + *k* + <D>. 4) to make bigger, to expand. *Přičiniti sě*: 1) to get close to: a) (*k*) + <D>. 2) to join: a) (*k*) + <D>. 3) to focus on: a) <D>, 4) to make an effort: a) *k* 'to' + <D>; b) *o* 'about' + <L>. / SS: *Przyczynić*: 1) to add: a) <N> + V + <A> + <D>. 2) to confirm: a) <A>. 3) to incite. *Przyczynić się*: 1) to advocate: *za* 'for' + <I>. 2) to cooperate. 3) to undertake. / ARj: *Pričiniti*: 1) to make, to build. <A>. *Pričiniti se*: 1) to happen: a) Ø. 2) to seem. / VRH: *Pričiniti se*: 1) to have a feeling that something which is not there can be seen or heard, to be seeing things: a) <N> + V + <D>.

In all of the earlier linguistic periods of the analysed languages, the verb kept the meaning 'to make', from which it has been derived, and in each of the compared languages it has developed very different meanings. The verb often denotes the meaning 'to add' and 'to build'. In Croatian, from the 18th century onwards, it has become a psychological verb, which means that this meaning is secondary.

## 20. PRISJETITI SE ('TO RECALL')

RM: +	SC: -	MS: +	T: +	C: +
-------	-------	-------	------	------

SRĀ: *Prisjetiti*: 1) to visit a) <A>. 2) to loathe: a) <A>. 3) to remind: a) <A>. / VRH: *Prisjetiti se*: 1) to bring to memory something from the past a) <G>; b) <Sent>.

The verb is formed from the prefix *\*pri* and the verb *\*sjetiti se*. The verb *\*sjetiti se* has probably been derived from the same Indo-European root *\*sh<sub>2</sub>ey-t-* 'rage, pain' as the noun *\*sěta* 'sadness' (ERHJ II: 156). The only language in which the verb has been recorded is Old East Slavic and the verb had a transitive and a causative meaning there. In the contemporary Croatian language, it appears only with a reflexive marking.

## 21. SAŽALITI SE ('TO FEEL SORRY FOR')

RM: +	SC: -	MS: +	T: +	C: -
-------	-------	-------	------	------

ERHJ: \**sobžaliti* / SRĀ: *Sžaliti*, *sžaleti* 1) to enrage someone: a) <A>. 2) to embitter: a) <A>. *Sžalitisâ*, *sžalitisâ*: 1) to become bitter. 2) to become enraged. 3) to get sorry for: a) <D>; b) *o* 'about' + <L>. / OCZ: *Sžěleti*: 1) to feel sorry for: a) <A> *Sžěliti sě*: 2) to feel sorry for: a) <D>; b) *o* 'about' + <L>. / ARj: *Sažaliti*: 1) to feel sorry for: a) <A> *Sažaliti se*: a) <D>; b) *radi* 'for' + <G>; c) *nad* 'over' + <I>. / HJP: *Sažaliti se*: 1) to feel remorse, compassion: a) *na* 'on' + <A>. 2) to get sad: a) *nad* 'over' + <I>.

In Old East Slavic, the verb marked an emotional state, not the feeling of being sorry for something or someone, but it does express other feelings, such as anger and bitterness. In Old East Slavic, Old Czech and Old Croatian, the verb was transitive. According to the analysed data, the use with a reflexive marking appeared earlier.

## 22. SMILOVATI SE ('TO SHOW MERCY')

RM: +	SC: -	MS: +	T: +	C: -
-------	-------	-------	------	------

SRĀ: *Smilovati*: 1) to take pity: a)  $\emptyset$ . *Smilovatisâ*: 1) to show mercy: a) *o* 'about' + <L>; b)  $\emptyset$ . / OCZ: *Smilovati*: 1) to take pity a) <A> *Smilovati sě*: 2) feel sorry: a) for <I>. / ARj: *Smilovati*: 1) to pardon: a) <A>. *Smilovati se*: a)  $\emptyset$ ; b) <D>; c) *na* 'on' + <A>; d) *vrh*, *vrhu* + <G> 'above, over, on top of'.

The verb has been derived from \**somiliti se* (ERHJ II: 362). In Old East Slavic, Old Croatian and Old Czech, the verb was transitive, but a causative use cannot be established as the meaning 'to make one merciful' has not been found. In Old East Slavic, only the use without a direct object has been found. In Old Croatian, the number of possible arguments was much greater. The preposition *vrh*, *vrhu* 'above' was used with verbs of speaking and thinking in Croatian from the 17th to the middle 19th century. That preposition, used with a genitive, is a translation of the Latin prepositional argument *super* + <Abl> (Hudeček 2003: 117-118). In the contemporary language, this prepositional argument has not been preserved.

## 23. SMUČITI SE ('TO GET NAUSEATED BY')

RM: +	SC: +	MS: +	T: +	C: +
-------	-------	-------	------	------

ERHJ: \**сѣмуčiti* / SRĀ: *Smučiti*: 1) to struggle: a) <A>. / OCZ: *Zmučiti*: 1) to struggle: a) <A>. / SS: *Smęcić*: 1) to torture: a) <A>. 2) to make sad: a) <A>; b) <G>; c) *przez* + <A>; d) *za* 'for' + <A>, 3) to agitate. *Smęcić się*: 1) to become sad. 2) to struggle. / VRH: *Smučiti se*: 1) to feel nauseated in one's stomach. 2) to feeling loathing or disgust for something: <N> + V + <D>.

In Old East Slavic, Old Czech and Old Polish, the records of the verb in its transitive meaning 'to torture someone' have been found. In the ARj, the verb has not been found, so nothing can be claimed about the change of the argument or the meaning. In contemporary Croatian, it has the meaning of a psychological verb and it denotes a physical reaction to an uncomfortable external sensation. The verb has the same valency pattern as the verbs *svidjeti se* and *gaditi se*. As in the earlier linguistic periods of the studied Slavic languages, the meaning as a psychological verb is not found, so it can be concluded that it has developed more recently.

## 24. SNEBITI SE ('TO WONDER')

RM: +	SC: +	MS: +	T: +	C: -
-------	-------	-------	------	------

ERHJ: \**сѣнебыти* / ARj: *Snebiti*: 1) to astonish: a) <A>, there is only one record of the verb being used without a reflexive marking, in the 17th/18th century. *Snebiti se*: 1) to become scared, to become confused, to become bewildered: a) Ø. 2) to disappear: a) Ø. / HJP: *Snebiti se*: 1) to become very unpleasantly surprised: a) Ø.

The verb has not been found in any of the earlier linguistic periods of Slavic languages, except for Croatian. In Old Croatian, the verb was polysemous and it did not belong only to the psychological verbs' class, but also to the verbs of appearance and disappearance (*zbiti se*, 'to happen' *pojavit se* 'to appear').

## 25. STRPJETI SE ('TO PATIENTLY WAIT')

RM: +	SC: -	MS: +	T: +	C: -
-------	-------	-------	------	------

ERHJ: \**сѣтврпѣти* (*se*) / SRĀ: *Sterpěti*: 1) to withstand: a) <A>. 2) to endure: a) <A>. *Sterpětisâ*: 1) It can be endured, a faceless use as part of a saying. / OCZ: *Strpěti*: 1) to endure, to withstand: a) <A>. *Strpěti se*: 1) to show the ability to

endure: a) s ‘with’ + <I>. / SS: *Ścierpieć, ścierpieć, ścierzpieć*: 1) to endure: a) <A>. / ARj: *Strpjeti*: 1) to withstand: a) <A>. 2) to wait, to withstand a) Ø; b) <A>. *Strpjeti se*: 1) to endure: a) Ø. 2) to allow: a) <N> + V + <D> + <A>. 3) to preserve. / HJP: *Strpjeti se*: 1) to patiently endure something or to wait for something to happen: a) Ø; b) *do* ‘until’ + <G>.

The verb has been found in all studied Slavic languages, and it was transitive in all of them, while it was not causative in any of them. In Old Czech and Old East Slavic, the verb appears with a reflexive marking, but its use is faceless in Old East Slavic, and marking behaviour in Old Czech. Some aspect of the discomfort experienced by the experiencer is expressed by an argument in the dative case. The verb became a *reflexiva tantum* verb in the more recent linguistic history, while its earlier use is transitive. In contemporary Croatian, the verb stopped being transitive and it is mostly used without an object.

## 26. SVIDJETI SE (‘TO LIKE’)

RM: +	SC: +	MS: +	T: +	C: -
-------	-------	-------	------	------

ERHJ: \**sъviděti se* / SRĀ: *Sviděti, s'viděti*: 1) to see: a) <A>. *Svidětisâ*: 1) to encounter: a) <I>. / SS: *Zwidzić się*: 1) to appear, to seem: a) Ø; 2) to approve, to agree: a) Ø. / ARj: *Svidjeti se*: 1) to become aware of something: a) Ø. 2) to be aware: a) <G>. 3) to think, to decide: Ø. 4) to remember: a) <Sent>. 5) to be likeable to someone: a) <N> + V + <D>. 6) to humble yourself: a) Ø. / BHGV: *Svidjeti se*: 1) to make a positive impression on someone a) <N> + V <D>; b) <D> + V + <Adv>; c) <D> + V + <Sent>; d) <D> + V + <Inf>. 2) to become attractive: a) <N> + V + <D>.

Among the researched languages, the verb was transitive only in Old East Slavic, but it was not causative. The transitive use is a more recent one, and it keeps the semantic relation to the verb *vidjeti* ‘to see’, from which this verb has been derived by prefixation. In Old Polish, the verb belongs to the *reflexiva tantum* verb class. It used to have a reciprocal meaning and the meaning of a verb of perception. In Old Croatian, the verb was also polysemic, but in all its meanings it belonged to psychological verbs.

## 27. UMILITI SE (‘TO BECOME LIKED BY SOMEONE’)

RM: +	SC: +	MS: +	T: +	C: +
-------	-------	-------	------	------

ERHJ: \**umiliti (se)* / MSR: *Umiliti*: 1) to make somebody feel mercy for one, to arouse someone’s mercy: a) <A>. *Umilitisâ*: 1) to grow sad: a) Ø. 2) to show

mercy to someone: a)  $v + \langle A \rangle$ , 12. 3) to be delighted: a)  $\langle I \rangle$ . / ARj: *Umiliti*: 1) to alleviate, to silence: a)  $\langle A \rangle$ . 2) to make dear: a)  $\langle A \rangle$ . 3) to humble a)  $\langle A \rangle$ . *Umiliti se*: 1) to flatter someone to make them like one: a)  $\langle N \rangle + V + \langle D \rangle$ . 2) to calm down: a)  $\emptyset$ . 3) to humiliate oneself: a)  $\emptyset$ . / VRH: *Umiliti se*: 1) to become dear to someone, to get in their good books: a)  $\langle N \rangle + V + \langle D \rangle$ .

The verb is composed of the prefix *\*u-* and the verb *\*militi*, which was formed by a denominal verb formation process from the adjective *\*milz* (ERHJ II: 509). It has been confirmed in Old East Slavic and Old Croatian. It was transitive and causative.

## 28. USUDITI SE ('TO DARE')

RM: +	SC: +	MS: +	T: +	C: -
-------	-------	-------	------	------

ERHJ: *\*usŏditi se* / SRĀ: *Vsuditi*: 1) to allot. *Vsuditisâ*: 1) to be a part of a certain composition: a)  $v + \langle A \rangle$ . / OCZ: *Usúditi*: 1) ordain. / ARj: *Usuditi*: 1) to condemn, to ordain:  $\langle A \rangle$ . *Usuditi se*: 1) to have the courage to do something a)  $\langle \text{Inf} \rangle$ ; b)  $\langle \text{Sent} \rangle$ . / BHGV: *Usuditi se*: 1) to have the daring or the courage to do something: a)  $\langle \text{Inf} \rangle$ .

In Old East Slavic, Old Czech, and Old Croatian, the verb was transitive, but not causative. In the transitive use, the verb meaning can be depicted as a sum of the meanings of the prefixes and the root. In the contemporary Croatian language, the verb became modal and it can appear exclusively with an argument in the infinitive, and its meaning is almost the same as the meaning of the verb *drznuti se* 'to dare'. Based on the fact that, in other Slavic languages, in the earlier linguistic periods, the verb has been recorded in transitive use, and that in that transitive use it has preserved meaning close to the meaning of its verb root, its *reflexiva tantum* semantic class is a more recent development, as well as its meaning 'to have courage to do something'.

## 29. UZDRŽATI SE ('TO ABSTAIN')

RM: +	SC: +	MS: +	T: +	C: -
-------	-------	-------	------	------

ERHJ: *\*vŏzdržati (se)* / SRĀ: *V”zd’ržati, v”zderžati, v”zdr”žati*: 1) to withhold, to stop: a)  $\langle A \rangle$ . 2) to support, raise: a)  $\langle A \rangle$ . 3) to hold: a)  $\langle A \rangle$ . *V”zd’ržatisâ, v”zderžatisâ, v”zdr”žatisâ*: 1) to abstain from:  $\langle G \rangle$ . / OCZ: *Vzdržěti*: 1) to keep: a)  $\langle A \rangle$ . *Vzdržěti sě*: 1) to abstain from a)  $\langle G \rangle$ . / SS: *Wzdzierzeć, wdzirzeć*: 1)

to keep, to imprison <A>. *Wzdzierżec się*: 1) to abstain from: a) *ot* + <G>. 2) to withstand: a) <A>. / ARj: *Uzdržati*: 1) to keep: a) <A>. 2) to preserve: <A>. 3) to receive: a) <A>. 4) to carry out, to fulfil: a) <A>. 5) to feed, to withstand, to refrain from: a) <A>. 6) to contain: a) <A>. *Uzdržati se*: 1) to linger: a) Ø. 2) subsist, to survive: a) Ø. 3) to withhold, to resist something: a) *od* 'from' + <G>; b) <Sent>; c) <G>. / HJP: *Uzdržati se*: 1) to manage to control oneself a) Ø. 2) to abstain from something: a) *od* 'from' + <G>. 3) to keep silent about something: a) Ø.

In the earlier periods of all Slavic languages, the verb was transitive. It had multiple meanings, all related to the verb *držati* 'to hold'. In all the analysed languages, the meaning *zadržati* 'to keep', and the other meanings have been derived from it. Similarly, the meaning it has when used with a reflexive marking, *ustegnuti se* 'to withhold', has also been confirmed in all languages and the meaning of the reflexive marking is related to the true reflexivity, i.e. the meaning can be described as 'to keep oneself away from something'.

### 30. ZALJUBITI SE ('TO FALL IN LOVE')

RM: +	SC: -	MS: +	T: +	C: -
-------	-------	-------	------	------

ERHJ: \**zaljubiti* / SRĀ: *Zalûbiti*: 1) to come to love someone: a) <A>. / OCZ: *Zalûbiti, zalíbiti*: 1) to come to love someone: a) <A>. 2) to promise. *Zalûbiti sě*: 1) to like: a) <N> + V + <D> + <A>. 2) to fall in love: a) *v* + <L>: to fall in love. / SS: *Zalubić*: 1) to make an agreement: a) *s* + <I>. 2) to come to love someone: a) <A>. *Zalubić się*: 1) to commit to someone. / ARj: *Zaljubiti*: 1) to come to love someone: a) <A>. *Zaljubiti se*: 1) to start loving: a) *u* 'in' + <A>; b) *s* 'with' + <A>. 2) to start loving one another. / BHGV: *Zaljubiti se*: 1) to start loving: a) Ø; b) *u* 'in' + <A>. 2) to reciprocate being in love. 3) to get carried away by something: a) *u* 'in' + <A>.

In all earlier periods of the researched languages, the verb was transitive and non-causative and its meaning was 'to come to love'. In Old Czech and Old Polish, its one possible meaning was also 'to commit to'. In Old Czech, it also had the meaning 'to like', and in that meaning it had the same valency pattern as the verb *sviđjeti se* 'to like' in Croatian. In the earlier period of Croatian, the verb had a greater number of arguments, but it also had the same meanings as in the contemporary language, the only change is the fact that the transitive meaning has not been preserved in the contemporary Croatian language.

Table 2 summarizes changes in five parameters (reflexive marking (RM), semantic class (SC), morphosyntactic properties of arguments (MS), transitivity (T) and causativity (C)) for every verb included in the research. Pluses (“+”) and minuses (“-”) in Table 2 mark the change of a parameter. The analysis of every individual verb explains which kind of change happened.

Table 2. Classification of the *reflexiva tantum* psychological verbs according to the parameters for valency change

Nº		RM	SC	MS	T	C
<b>IMPERFECTIVE VERBS</b>						
1.	BOJATI SE ('TO FEAR')	-	-	+	-	-
2.	DIVITI SE ('TO ADMIRE')	+	-	+	+	+
3.	DURITI SE ('TO SULK')	+	+	+	-	-
4.	GNUŠATI SE ('TO ABHOR')	+	-	+	+	+
5.	GROZITI SE ('TO DETEST')	+	-	+	+	+
6.	KAJATI SE ('TO REGRET')	+	-	+	+	-
7.	NADATI SE ('TO HOPE')	-	-	+	-	-
8.	PONOSITI SE ('TO BE PROUD OF')	+	+	+	+	-
9.	SNEBIVATI SE ('TO WONDER')	+	+	+	+	-
10.	SRAMITI SE ('TO FEEL ASHAMED')	+	-	+	+	+
11.	SVIDATI SE ('TO LIKE')	+	+	+	+	-
12.	STIDJETI SE ('TO FEEL ASHAMED')	+	-	+	+	+
13.	UZDATI SE ('TO RELY ON SOMEONE')	+	+	+	+	+
<b>PERFECTIVE VERBS</b>						
14.	DRZNUTI SE ('TO DARE')	+	-	-	-	-
15.	NASLADITI SE ('TO GLOAT')	+	-	+	+	+
16.	OKAMENITI SE ('TO FREEZE IN SHOCK')	+	-	+	+	+
17.	OSLADITI SE ('TO ACQUIRE A TASTE FOR SOMETHING')	+	-	+	+	+
18.	POKAJATI SE ('TO REPENT')	+	-	+	+	-
19.	PRIČINITI SE ('TO SEE SOMETHING WHICH IS NOT THERE')	+	+	+	+	+
20.	PRISJETITI SE ('TO RECALL')	+	-	+	+	+
21.	SAŽALITI SE ('TO FEEL SORRY FOR')	+	-	+	+	-

22.	SMILOVATI SE ('TO SHOW MERCY')	+	-	+	+	-
23.	SMUČITI SE ('TO GET NAUSEATED BY')	+	+	+	+	+
24.	SNEBITI SE ('TO WONDER')	+	+	+	+	-
25.	STRPJETI SE ('TO PATIENTLY WAIT')	+	-	+	+	-
26.	SVIDJETI SE ('TO LIKE')	+	+	+	+	-
27.	UMILITI SE ('TO BECOME LIKED BY SOMEONE')	+	+	+	+	+
28.	USUDITI SE ('TO DARE')	+	+	+	+	-
29.	UZDRŽATI SE ('TO ABSTAIN')	+	+	+	+	-
30.	ZALJUBITI SE ('TO FALL IN LOVE')	+	-	+	+	-

## 7. Discussion

From the possible combinations of the values of the parameters for valency change shown in Table 1, from 32 available combinations of default parameters, only seven have been confirmed. Since the shift of verbs from transitive to intransitive also includes the change of the verb arguments' composition, it can be assumed that there are rare or unconfirmed cases of a reflexive marker change, while the same verb arguments are preserved. In Table 3 the seven attested combinations of parameters and the number of verbs displaying the relevant combination are shown.

Table 3. Attested combinations of parameters for valency change in *reflexive tantum* psychological verbs

№	RM	SC	MS	T	C	
1.	+	+	+	+	+	4
2.	+	+	+	+	-	7
3.	+	+	+	-	-	1
4.	+	-	+	+	+	9
5.	+	-	+	+	-	6
6.	+	-	-	-	-	1
7.	-	-	+	-	-	2

A plus in Table 3 means that the following parameter changed over time, and a minus means that it remained the same. It is essential to highlight that change is understood as loss and as the appearance of a feature. In the

Discussion for every group of verbs, it is explained what kind of change happened. The seven groups can be divided into two groups according to the criterion of the presence of transitivity and then also a causative use in the analysed Slavic languages. Three groups (Table 3, rows 3, 6 and 7) contain four verbs that have not changed their transitivity feature. The verb *drznuti se* (Table 3, row 6) changed only the presence of a reflexive marker. In all of older examined Slavic languages, the verb was used without the reflexive marker. The verb *duriti se* (Table 3, row 3) has been found without a reflexive marking only in Old East Slavic, where it marks a state, and in all other examined older periods of Slavic languages, it is used with the reflexive marker. A transitive and a causative form of the verb *duriti se* exist in contemporary Russian, Polish and Ukrainian, but it has not been recorded in the earlier periods of the analysed Slavic languages. The verbs *bojati se* 'to fear' and *nadati se* 'to hope' belong to the third group (Table 3, row 7), as they have not changed their meaning nor their semantic class, but only their arguments.

The fact that only two verbs have been *reflexiva tantum* throughout their entire linguistic history, and that the other *reflexiva tantum* verbs, 28 of them, could be used in their transitive meaning, points to the disappearance of transitivity. The opposite case, in which primarily *reflexiva tantum* verbs become transitive verbs, has not been found in Croatian. However, it is possible that there have been such examples, but have not been recorded due to the methodology of excerpting verbs which has been used. Furthermore, only one semantic class of verbs has been included in the research, i.e. not all *reflexiva tantum* verbs have been included. Regardless of this lack, a tendency for formal reflexivization can be noticed in modern Croatian. The described situation concerns first and foremost Croatian and not necessarily the other Slavic languages. The dominance of a certain tendency in other Slavic languages could be determined with research similar to this one, in a sample including a greater number of languages.

Four groups consist of verbs whose transitivity has changed. Among these four groups, we can further distinguish between the verbs which have been causative in the earlier linguistic period and those which have not. In the category of verbs that used to be causative, there is an additional difference among those which have belonged to different semantic classes and have changed their meaning significantly. Verbs that changed all parameters are shown in row 1 in Table 3. In that group, only one causative verb (*pričiniti se* 'to appear to be') has not been derived from a noun or an adjective, and the rest (*smučiti se*, *umiliti se*, *uzdati se*) have.

The verbs whose meaning has changed by losing a causative component (Table 3, row 4) belong to a similar group. The verbs of this type are the most numerous verb group in this sample. In that group, almost all verbs, except for the verb *gnušati se*, which is historically derived from the verb \**gnusiti se* (ERHJ I: 279), are formed from a denominal base (*diviti se*, *groziti se*, *sramiti se*, *stidjeti se*, *nasladiti se*, *okameniti se*, *osladiti se*) and the causative suffix. Among the recorded verbs, 11 of them fit the causative meaning formation pattern from the adjective or nominal verb base by adding a suffix *-i-*, described by Gołąb (1968), which is highly productive in Russian as well (Plungân 2001). This pattern shows the reduced productivity of causative dichotomies in the contemporary Croatian language, in comparison to the earlier period verb states in the analysed Slavic languages, which can be depicted by the following schema:

- (3)  $\{(groziti \text{ 'abhor' } [do' (x, \emptyset)] \text{ CAUSE } [feel' (y), [zgrožen \text{ 'abhorred'}]]\} t_1$   
 $\rightarrow \{groziti se \text{ 'to be abhorred' } (x, y)\} t_1/t_2$

The schema shows that the decausative meaning, formally marked by a reflexive marking, was present in both time periods, and the causative meaning exclusively in the earlier period. The causative formation reduced productivity, and it can be characteristic only for contemporary Croatian and no other contemporary Slavic languages. Based on the results acquired exclusively for Croatian, it is not possible to draw conclusions on the reduced productivity of this pattern in the history of Slavic languages in general, and the possibility that Slavic languages differ from each other by this parameter, needs to be taken into consideration. The results of the preliminary comparison of Slavic denominal verbs which have entered causative pairs do not give unambiguous answers about the diachronic primary state of the (non)reflexive valency pattern. However, some conclusions can still be made. Most non-prefixed verbs which follow the described pattern (*duriti se*, *diviti se*, *sramiti se*, *stidjeti se*) show diachronic primary state of the non-causative form in the analysed sources. The non-prefixed verb *groziti se* 'to abhor' does not show a notable diachronic primary state for neither of the causative pair members. A scheme opposite to the one applicable to the Croatian language verbs is relevant for these verbs:

- (4)  $\{stiditi se \text{ 'to be ashamed' } (x, y)\} t_1 \rightarrow \{(stiditi \text{ 'to shame' } [do' (x, \emptyset)]$   
 $\text{ CAUSE } [to feel (y), [posramljen \text{ 'ashamed'}]]\} t_1/t_2$

Prefixed verbs more often than non-prefixed verbs had transitive and causative forms (*prisjetiti se*, *smučiti se*, *umiliti se*). Also, in some cases, prefixed verbs do not show a diachronic primary transitive or causative form (*nasladiti se*, *okameniti se*, *osladiti se*). This points to the fact that the non-prefixed psychological verbs could be diachronically first in their reflexive form and decausative meaning, and the prefixed could be diachronically first in their transitive and causative meaning. Such a relationship is not unusual, because verb prefixation is used as a syntactic means of the transitivization of certain verb groups (Krysko 2006: 292). The verbs which changed meaning kept the meaning of the base from which they had been derived in the earlier linguistic period, so that, for example, the verb *smučiti se* ‘to be disgusted by’ had not had the meaning of the feeling of disgust recorded, but only of a short period of torturing the second actant in the semantic role of the patient.

The verbs which changed transitivity, but not causativity, can be classified into two groups. The first group includes the verbs which have changed their semantic class (Table 3, row 2). The verb *ponositi se* used to have the meaning ‘to move’ and ‘to put’, in its transitive use (Brač & Bošnjak Botica 2015: 114) and it could make a transition to the category of verbs of communication. The change in meaning is marked by an argument in dative which denotes the face at which what is being said is directed. The verb *groziti se* is the same. The difference is in the fact that in contemporary Croatian, the verb *ponositi se* can no longer be used as a verb of speaking, and the verb *groziti se* can. The verb *uzdržati se* used to a verb of ‘keeping’ and ‘holding’ in its transitive use (Brač & Bošnjak Botica 2015), and in its reflexive use, preserved in the contemporary Croatian language, the verb is exclusively psychological. The verb *usuditi se* ‘to dare’ used to have a transitive performative meaning ‘to condemn’, and in the contemporary language, it marks the meaning ‘to have courage’ and it exclusively has an infinitive argument and a modal meaning. The reflexive marking with all three stated verbs can be related to the meaning of the reflexive pronoun, which is unusual for the other verbs. The verbs *snebiti se* and *sne bivati se* could have belonged to the group of verbs of occurrence (Mikelić Preradović 2020: 79) characterized by complete decausativity, i.e. by a non-specified cause of the action. Such shift in meaning is probably the result of their similarity to the verbs *zbiti se* ‘to happen’, *zbivati se* ‘to be happening’. The verbs *svidjeti se* and *sviđati se* belonged to the semantic class verbs of perception in the earlier periods and to the group with which the verbs of perception have a relation of regular polysemy – the verbs of thinking, learning, and knowing.

Such a transition is possible due to their preservation of the relationship with the non-prefixed verbs *vidjeti* and *vidati*.

The second subgroup of verbs which changed transitivity, but not causativity, (Table 3, row 5) includes the verbs which have not changed their semantic class. The transitive use of the verb *kajati se*, 'to regret', has been recorded only in Old East Slavic in one example from the 12th century, which shows that the transitive use is secondary. A similar scenario is also applicable to the verb *pokajati se* 'to repent', for which the reflexive use has only been recorded in Old East Slavic in the 15th century, and in the other Slavic languages the verb was intransitive prior to the 15th century. The verb *strpjeti se* 'to patiently wait' used to be transitive, as it had an argument in the accusative case which marked the discomfort felt by the experiencer, and such a use was undoubtedly its primary one. In contemporary Croatian, the verb became a *reflexiva tantum* verb and it almost completely lost nearly every possibility of having a direct object added and the verb can now only have an adjunct. The verbs *sažaliti se*, *smilovati se*, *zaljubiti se* have changed their valency pattern by preserving the semantic role of the patient, i.e. the face on which the action is performed, but in the way that the patient is not expressed by a direct object in the contemporary language, but with a dative argument instead (*sažaliti se*, *smilovati se*) or a prepositional argument (*zaljubiti se*). Other than that, the verbs *smilovati se* and *sažaliti se* have a pronounced semantic role of the cause in the contemporary language, which is expressed by a prepositional argument.

Verb arguments in some verb groups have not been studied in detail due to several reasons. Most sources in the ARj, from which the data have been extracted, are written in Chakavian dialects, which is not the direct ancestor to the contemporary Croatian language, and also, in the dictionary itself, the arguments are not described in the equally detailed manner for all dictionary entries. However, some more general conclusions can be made. It has been noticed that in the earlier period of Croatian, verbs could have more arguments, as 25 verbs had more arguments in Old Croatian, not including the accusative argument which is the transitivity criterion, while for four verbs there are no data, and only one verb, *svidjeti se* 'to like', in the meaning of a psychological verb, had a fewer number of arguments when compared to the contemporary language. Other than that, more frequent bare arguments have been noticed with some verbs. The analysis of the five parameters of psychological verbs makes it possible to create a hierarchy of their frequency and diachronic stability in Croatian. The properties of arguments marking are the least stable, only one verb *drznuti se* 'to dare' has been found with

unchanged arguments. All the other verbs changed their arguments, usually by reducing their number. The reflexive marking is the second least stable feature. Only two verbs have remained part of the *reflexiva tantum* category throughout time. Most verbs which could be marked without a reflexive marking, 26 of them, were transitive. Thirteen verbs in total might have been causative in the studied sample and none of them preserved causativity in the contemporary language. The thematic class to which the verb belongs has shown to be the most stable parameter. Twelve verbs belonged to the psychological verbs' group the entire time. This does not include the regular polysemy cases, typical for the psychological verbs thematic class.

## 8. Conclusion

In the history of Croatian, the psychological verbs used to be able to more frequently enter reflexive oppositions, which shows that most *reflexiva tantum* verbs have a more recent origin. In other Slavic languages, the given category does not have to be so widespread nor typical for psychological verbs. The chronology of the development of the *reflexiva tantum* category could in theory be reconstructed, so that the earliest recorded use of the non-prefixed psychological verbs derived from nouns and adjectives using the suffix *-i-* is recognized as a reflexive one, followed by a reflexive marking drop, which indicates the secondary, causative meaning. This causative meaning would afterwards have gradually been lost throughout the history of Croatian, which does not necessarily have to also be applicable to other Slavic languages, for which there may not have occurred a loss of causative opposition. Prefixed verbs, as a rule, originally had a transitive meaning or the primary state of transitivity has not been confirmed, which can be related to the transitivity feature, which appears as a result of the prefixation. The analysis of psychological verbs points to the fact that in the earlier linguistic history, Croatian used to be a little more transitive than it is today. However, future research should also include verbs other than the psychological ones, as it would make this statement more reliable. The position of Croatian within the framework of other Slavic languages, which can show, to a varying degree, the characteristics of detransitivizing languages, remains uncovered. Reflexive marking, as a formal indicator, usually, but not always, marks the end of transitivity, and it can sometimes entail a change in the causative meaning. The data from other Slavic languages need to be compared in future research, and other semantic classes should be included.

## Appendix

### Acknowledgments

We are grateful for the valuable comments given by two anonymous reviewers.

### List of abbreviations

<A>	accusative
<Abl>	ablative
C	causativity
<D>	dative
<G>	genitive
<I>	instrumental
<Inf>	infinitive
<L>	locative
MS	morphosyntactic properties of arguments
<N>	nominative
∅	absence of an object
RM	reflexive marker <i>se</i>
SC	thematic class
<Sent>	sentence
T	transitivity
V	verb

### Acronyms for lexicographic sources

ARj	Daničić <i>et al.</i> (1880–1976).
ERHJ	Matasović <i>et al.</i> (2016–21).
ÈSSÂ	Trubačev (1974–).
BHGV	Baza hrvatskih glagolskih valencija.
HJP	Hrvatski jezični portal.
EDSL	Derksen (2008).
OCZ	<i>Slovníky staré češtiny</i> .
SP	Sławski (1974–2001).
SS	Urbańczyk (1953–2002).
MSR	Sreznevskij (1893–1912).
SRÂ	Avanesov (1975–).
VRH	Jokić, Ljiljana (2015).

## References

- Ānko-TrinickaĀ, N. A. (1962). *Vozvratnye glagoly v sovremenom ruskom āzyke*. Moskva: Izdatel'tstvo Akademii Nauk SSSR.
- Apresjan, Ju. D. (1974). Regular Polisemy. *Linguistics*, 12 (143), 5–32.
- Avanesov, P. I. (ed) (1975–). *Slovar' russkogo āzyka XI–XVII vv*. Moskva: Nauka.
- Baza hrvatskih glagolskih valencija. (Available online: <http://valencije.ihjj.hr/>, Accessed on 2022–09–02)
- Belaj, B. (2001). Prototipno-kontekstualna analiza povratnih glagola u hrvatskom jeziku. *Suvremena lingvistika*, 51/52(1/2), 1–11.
- Belaj, B. & TanackoviĀ Faletar, G. (2017). *Kognitivna gramatika hrvatskoga jezika. Sintaksa jednostavne reĀenice*. Zagreb: Disput.
- Beram breviaries (Available online: <https://beram.stin.hr/>, Accessed on 2023–08–12).
- BirtiĀ, M. & Matas IvankoviĀ I. (2009). Akuzativne dopune uz neprijelazne glagole: ŀto su unutraŀnji objekti? *Rasprave: Āasopis Instituta za hrvatski jezik i jezikoslovlje*, 35(1), 1–19.
- BraĀ, I. & Boŀnjak Botica, T. (2015). SemantiĀka razdioba glagola u Bazi hrvatskih glagolskih valencija. *Fluminensia*, 1, 105–121.
- BrozoviĀ, D. (2008). *Povijest hrvatskog knjiŀevnog i standardnog jezika*. Zagreb: ŀkolska knjiga.
- DaniĀiĀ, Đ., Valavac, M., Budmani P., MaretiĀ, T., Bosanac, S., GrdeniĀ, D., Hamm, J. & Jedvaj, J. (eds) (1880–1976). *RjeĀnik hrvatskoga ili srpskoga jezika*. Zagreb: JAZU.
- Derksen, R. (2008). *Etymological Dictionary of the Slavic Inherited Lexicon*. Leiden: Brill.
- GeniuŀienĀ, E. (1987). *The Typology of Reflexives*. Berlin: Mouton de Gruyter.
- GoĀab, Z. 1968. The grammar of Slavic causatives. In: Kucera, H. (ed) *American Contributions to the Sixth International Congress of Slavists*, 71–94. The Hague: Mouton.
- Grimshaw, J. (1990). *Argument structure*. Cambridge (MA): MIT Press.
- Hansack, E., Hansen, B., Wald, V., Horvat, M. & PeriĀ GavranĀiĀ, S. (2016). Regensburŀski dijakronijski korpus hrvatskoga jezika – CroDi. *Rasprave: Āasopis Instituta za hrvatski jezik i jezikoslovlje*, 42 (1), 1–19.
- HrdliĀka, M. (2020). *Kategorija povratnosti u hrvatskom i poljskom jeziku*. PhD thesis, University of Zagreb.
- Hrvatski jeziĀni portal. (Available online: <https://hjp.znanje.hr/>, Accessed on 2022-09-29).
- HudeĀek, L. (2003). Dopune glagolima govorenja, miŀljenja i srodnih znaĀenja u hrvatskome knjiŀevnom jeziku od 17. do polovice 19. stoljeĀa – strani sintaktiĀki utjecaji. *Rasprave Instituta za hrvatski jezik i jezikoslovlje*, 29, 103–129.
- JokiĀ, Lj. (ed). (2015). *Veliki rjeĀnik hrvatskoga standardnog jezika*. Zagreb: ŀkolska knjiga.
- KnĀzev, ŀ. (2007). *GrammatiĀeskaĀ semantika: Russkij āzyk v tipologiĀeskoj perspektive*. Moskva: Āzyki slavĀnskih kul'tur.
- Krysko, V. V. (2006). *IstoriĀeskoj sintaksis russkogo āzyka. Ob"ekt i perehodnost'*. Moskva: Izdatel'skiĀ centr "Azbukovnik".
- Levin, B. & Rappaport Hovav, M. (2005). *Argument Realization*. Cambridge: Cambridge University Press.
- LjubeŀiĀ, N. & KlubiĀka, F. (2014). {bs,hr,sr}WaC – Web Corpora of Bosnian, Croatian and Serbian. In: Bildhauer, F. & SchĀfer, R. (eds) *Proceedings of the 9th Web as Corpus Workshop (WaC-9)*, 29–35. Association for Computational Linguistics.

- Matasović, R. (2018). Aspect and Aktionsart in Slavic, Inflection and Derivation in RRG. In: Kailuweit, R.; Künkel, L. & Staudinger, E. (eds) *Applying and expanding role and reference grammar*, 95–109. Freiburg: Freiburg Institute for Advanced Studies, Albert-Ludwigs-Universität Freiburg.
- Matasović, R., Ivšić Majić, D., Pronk, T. & Brozović Rončević, D. (2016–2021). *Etimološki rječnik hrvatskoga jezika, I-II*. Zagreb: Institut za hrvatski jezik i jezikoslovlje.
- Mikelić Preradović, N. (2020). *CROVALLEX: Valencijski leksikon glagola hrvatskoga jezika*. Zagreb: Zavod za informacijske studije.
- Nedálkov, V. P. & Sil'nickij, G. G. (1969). Tipologija morfoložičeskogo i leksičeskogo kauzativov. In: Holodovič A. A. (ed) *Tipologija kauzativnyh konstrukcij: Morfoložičeskij kauzativ*, 20–51. Leningrad: Nauka.
- Nichols, J., Peterson, D. A. & Barnes, J. (2004). Transitivity and detransitivizing languages. *Linguistic Typology*, 8(2), 149–211.
- Oraić Rabušić, I. (2018). *Struktura povratnih glagola i konstrukcije sa se u hrvatskome jeziku*. Zagreb: Institut za hrvatski jezik i jezikoslovlje.
- Padučeva, E. V. (1994). Tipy kauzal'nyh otnošenij v semantičeskoj strukture leksemy. *Russian Linguistics*, 18(1), 1–16.
- Padučeva, E. V. (1996). *Semantičeskie issledovaniâ: semantika vremeni i vida v russkom âzyke, semantika narrativa*. Moskva: Âzyki russkoj kul'tury.
- Padučeva, E. V. (2004). *Dinamičeskie modeli v semantike sintaksisa*. Moskva: Âzyki slavânskoj kul'tury.
- Pesetsky, D. (1995). *Zero Syntax: Experiencers and Cascades*. Cambridge (MA): MIT Press.
- Plungân, V. A. (2011). *Vvedenie v grammatičeskuû semantiku: grammatičeskie značeniâ i grammatičeskie sistemy âzykov mira*. Moskva: Rossijskij gosudarsvennyj gumanitarnyj universitet.
- Polančec, J. (2018) Osamostaljeni izvedeni nesvršeni glagoli u hrvatskom jeziku. *Suvremena lingvistika*, 44(85), 113–138.
- Raffaelli, I. (2009). *Značenje kroz vrijeme. Poglavlja iz dijakronijske semantike*. Zagreb: Disput.
- Samardžija, M. (1994). Valentnost i semantičke mijene hrvatskih glagola. *Fluminensia*, 6(1/2), 49–53.
- Slovníky staré češtiny*. (Available online: <https://vokabular.ujc.cas.cz/hledani.aspx>, Accessed on 2022-08-05).
- Sreznevskij, I. I. (1893–1912). Materialy dlâ slovarâ drevne-russkago âzyka po pis'mennym "pamâtnikam". Sankt-Peterburg: Tipografija imperatorskoj akademij nauk".
- The Russian National Corpus*. (Available online: <https://ruscorpora.ru/>, Accessed on 2022-08-05).
- Trubačev, O. N. (ed) (1974–). *Ėtimologičeskij slovar' slavânskikh âzykov*. Praslavânskij leksičeskij fond, t. 1-42. Moskva: Nauka.
- Urbańczyk, S. (ed) (1953–2002). *Słownik staropolski, t. 1-11*. Kraków: Instytut Języka Polskiego Polskiej Akademii Nauk.
- Van Gelderen, E. (2018). *The Diachrony of Meaning. Aspect and Argument Structure*. New York: Routledge.
- Van Valin, R. D. & LaPolla, R. J. (1997). *Syntax: Structure, Meaning, and Function*. Cambridge: Cambridge University Press.

## 7. Alternating constructions with *mittō* 'send': the encoding of third arguments

MARTINA GIULIANI

Keywords: *Latin; ditransitive construction; argument alternation; animacy; semantic role; affectedness*

Some Latin ditransitive verbs of physical transfer (e.g. *dō* 'give', *mittō* 'send') feature construction alternation: animate third arguments (i.e. recipient-like arguments) can be expressed either by the dative case or by an allative prepositional phrase with *ad* and the accusative. This paper investigates argument structure and alternations of the Latin verb *mittō* 'send', focusing on the encodings of recipient-like arguments, with a corpus-based approach. My analysis of a sample of ditransitive constructions shows that construction alternation of *mittō* 'send' can be motivated by considering variables taken into account in typologically-oriented studies. I argue that the distribution of the encodings of recipient-like arguments can be related to: (i) animacy of theme arguments; (ii) semantic roles of third arguments (e.g. animate directions vs. concrete recipients); (iii) degree of affectedness of third arguments.

*Original scientific paper*

### 1. Introduction

The issue of ditransitive constructions has been at the centre of linguistic research, in different theoretical frameworks (s. Haspelmath 2015 for a review of previous studies). Ditransitive verbs show high construction variability both cross-linguistically and intra-linguistically. Scholars have also devoted attention to Latin ditransitives (e.g. Baños 1998, 2000; Fedriani & Prandi 2014; Napoli 2018, among others). A group of Latin ditransitives has been acknowledged to alternate in the expression of the third argument (i.e. the recipient-like argument) since Early Latin (s. Bennett 1914; Baños 1996, 1998; Luraghi 2010; Adams 2013; Fedriani & Prandi 2014; Pinkster 2015; Adams & de Melo 2016; Napoli 2018). The third argument might be encoded by either the dative case or the allative prepositional phrase (PP) with *ad*

and the accusative (*ad* + ACC). Although several studies have looked at these verbs (e.g. among others Fedriani & Prandi 2014; Adams & de Melo 2016), the factors motivating this construction alternation have not yet been fully explored. To further investigate this topic, in this paper I examine the argument structure and the alternations of the Latin ditransitive verb *mittō* 'send' with a corpus-based approach. My analysis follows Goldberg's (1995, 2006) constructionist approach to argument structure within a usage-based perspective (s. Perek 2015). Argument structures are constructions «pairing an abstract event description containing an array of arguments, with a specification of the morphosyntactic encoding of these arguments» (Perek 2015: 24). Constructions exist independently of particular verbs, that is, they themselves carry meaning independently from the lexical items (i.e. verbs and their arguments) instantiating them (Goldberg 1995). Alternations arise when a verb is compatible with more than one construction and result in variations in number and coding patterns of the arguments (Goldberg 2006).<sup>95</sup> The main goal of this paper is to describe and propose motivation for construction alternation with *mittō* 'send' considering parameters taken into account in typologically-oriented studies, such as animacy, semantic roles and affectedness (see e.g. among others, Haspelmath 2007; Kittilä 2008; Malchukov *et al.* 2010).

The paper is structured as follows. Section 2 provides theoretical remarks concerning the definition of a ditransitive construction and the encodings of its arguments. Section 3 is focused on the ditransitive construction of the Latin verb *mittō* 'send' and on previous studies investigating its construction alternation. In Section 4 I discuss the methodology, describing the corpus, data extraction, and annotation. In Section 5 I show the results of my analysis, which are discussed in Section 6. Finally, in Section 7 I summarize my findings and suggest directions for future research.

---

<sup>95</sup> See Ježek (*this volume*, Chapter 2, section 3) for a survey of approaches to the semantic classification of verbs based on verb alternations.

## 2. Ditransitive constructions: definition and types of alignment

In the typological framework, ditransitive constructions are regarded as constructions «consisting of a (ditransitive) verb, an agent argument (A), a recipient-like argument (R), and a theme argument (T)» (Malchukov *et al.* 2010: 2; see also Korn & Malchukov 2018). Such a definition, making reference to the semantics rather than to the syntactic properties of the construction, is cross-linguistically tenable.<sup>96</sup> According to Malchukov *et al.* (2010: 2) prototypically<sup>97</sup> ditransitive constructions «contain a verb of physical transfer (e.g. 'give', 'hand') and describe a scene in which A causes an object to pass into the possession of an animate receiver (= recipient)» (see also Goldberg 1995; Newman 1996; Kittilä 2006; Haspelmath 2015). An example of a prototypical ditransitive construction of the verb 'give' is shown in (1).

- (1) *Mary gave John a pen.*  
       A                  R      T

From a semantic perspective, participants of prototypically ditransitive events differ in terms of (i) animacy; (ii) volition/control properties. While the A and the R are animate, typically instantiated by human referents, the T is low in animacy. The A and the R are characterized by high control and volition over the verbal process: the A volitionally causes the transfer of the T, which the R is supposed to accept willingly. Denoting an inanimate entity, the T does not exert any control over the event: in this sense it is similar to a patient (P) of a monotransitive construction. However, the T is less affected than the P by the verbal process, as it undergoes a change of location rather than a change of state (on the scalar notion of transitivity

<sup>96</sup> As ditransitive verbs show high construction variability both cross-linguistically and intra-linguistically, syntactically-based definitions are not cross-linguistically applicable (Malchukov *et al.* 2010). However, some theoretical frameworks adopt definitions based upon formal properties. In traditional English studies (e.g. Goldberg 1995, 2006; Levin & Rappaport Hovav 2005) the term ditransitive refers to a double object construction (e.g. *John sent Mary a letter*), in contrast to a prepositional construction (e.g. *John sent a letter to Mary*).

<sup>97</sup> A prototypical view of categories is borrowed from Cognitive Linguistics and goes back to Rosch's (1975) experimental studies in Cognitive Psychology.

s. Hopper & Thompson 1980; Tsunoda 1981, 1985). Some verbs of physical transfer (e.g. 'bring' and 'send') can be constructed with non-prototypical T and R arguments. Consider examples (2)–(3).

(2) *Mary brought his son to the grandmother.*

A                      T              R

(3) *Mary sent a letter to Rome.*

A                      T              R

In example (2) containing 'bring', the T argument has a human referent. In example (3) instead, the verb 'send' expresses the transfer of the T towards a place, bearing the semantic role of direction, i.e. the endpoint of a trajectory (Luraghi 2010, 2011). The semantic roles of recipients and directions are distinguished based on (i) animacy and (ii) presence/absence of the feature of possession. A recipient is typically animate and, as a result of a transfer event, gains possession or control over the T. A direction, instead, is typically inanimate and is not characterized by the feature of possession (see Kittilä & Ylikoski 2011 for a comprehensive view).

A closely related type of construction is the benefactive construction. In many languages benefactive constructions are expressed as ditransitive constructions (Malchukov *et al.* 2010). Beneficiaries and/or maleficiaries are participants that are, advantageously or not, affected by an event without being its obligatory participants (Zúñiga & Kittilä 2010: 2). Recipients and beneficiaries have some properties in common, as recipients usually benefit from the transfer of the T. However, reception is not a necessary feature of beneficiaries (see Kittilä 2005 and Zúñiga & Kittilä 2010 for further details). The most salient difference between these two semantic roles is that only beneficiaries occur with intransitive verbs (Malchukov *et al.* 2010: 2).

Ditransitive constructions represent the core of three-argument constructions (Malchukov *et al.* 2010). The argumenthood of the third participant (R) of some ditransitives (e.g. 'send') is controversial. The argument/adjunct distinction, especially with spatial constituents, is not clear-cut and benefits from a scalar approach (s. Haspelmath 2014; Haspelmath & Hartmann 2015). According to Kittilä (2006) events expressed by the verbs 'give' and 'send', despite inherently involving three participants, differ in respect to the prominence accorded to these participants. An event of 'giving' is completed only if the R has accepted the transfer of the T. In contrast, an event of 'sending' can be considered completed even if the R

has not received the T. As highlighted by Kittilä (2006: 56), Rs «may even be completely unaware of the event, and (they) may refuse to accept the transfer». Whereas an event of giving is conceptualized as a whole, with the three participants activated simultaneously, an event of sending consists of two phases, which can be separated by a temporal interval. Moreover, the A and the R of an event of giving are physically implied, as the T is directly handed over to the R (Newman 1996; Kittilä 2006). Such a condition does not apply to an event of 'sending'. The lexical semantics of 'send' entails a motion component: the A and the R are not physically implied, and the transfer of the T requires a means of transportation.

Differences in the encodings of the three arguments of ditransitive events are captured by the notion of alignment (s. Malchukov *et al.* 2010; Haspelmath 2011; Haspelmath 2015). In ditransitive constructions whereas the A is always treated as the subject of prototypical monotransitive constructions, differences arise in terms of the alignment of the coding of T and R arguments, that is, «whether it is the R or the T of a ditransitive clause that is coded like the P of monotransitive clauses» (Haspelmath 2015: 22). The resulting basic alignment types are: (i) indirective alignment, if the R is treated differently from the P and the T ( $T = P \neq R$ ), (ii) secundative alignment, if the T is treated differently from the P and the R ( $T \neq P = R$ ), and (iii) neutral alignment: if the T, the R, and the P are encoded in the same way ( $T=P=R$ ). Besides the three basic alignment types mentioned, languages often show splits and alternations (Malchukov *et al.* 2010: 18-20). According to Malchukov *et al.* (2010: 18) «a lexical split is the situation where different verbs use different constructions, while an alternation is the situation where one and the same verb occur with different constructions with roughly the same meaning». The choice between alternating constructions can be conditioned by several factors, such as affectedness of R-like arguments and prominence (i.e. animacy/topicality) of T arguments (s. Haspelmath 2007: 83–85; Malchukov *et al.* 2010: 20–21).

### 3. The ditransitive construction of the Latin verb *mittō* ‘send’

#### 3.1. Argument structure and alternations of *mittō* ‘send’

Like other Latin verbs of physical transfer (e.g. *dō* ‘give’ and *ferō* ‘bring’), *mittō* ‘send’ features an indirective alignment (for a typological approach towards Latin ditransitives, see Napoli 2018). Whereas the A and the T are encoded as the A and the P of a prototypical monotransitive construction, marked with the nominative and accusative cases, respectively, the R is expressed differently, allowing for both case marking and prepositional constructions. The animacy of Rs implies a split in the coding patterns (Luraghi 2010; Pinkster 2015; see also Kittilä 2008, 2022 on differential marking of Rs more in general). Inanimate Rs are expressed with direction PPs: *ad* + ACC (s. (4)); *in* + accusative (*in* + ACC), (s. (5)).

- (4) *meretrices ... ad portum mittunt seruos ancillulas*  
 prostitute:NOM.PL to harbour:ACC.SG send:PRS.3PL slave:ACC.PL slave\_girl:ACC.PL  
 ‘Prostitutes ... send slaves and slave girls to the harbour.’ (Plaut. *Men.* 338–339)<sup>98</sup>
- (5) *nunc hinc parasitum in Cariam misi meum ...*  
 now hence parasite:ACC.SG in Caria:ACC.SG send:PRF.1SG POSS.ACC.SG  
 ‘Now I sent off my parasite to Caria ...’ (Plaut. *Curc.* 67)

The preposition *in* implies containment: the T moves towards a destination, and eventually it will occupy a portion of the space of the landmark (s. Luraghi 2011). However, the allative PP: *ad* + ACC instead, does not entail the spatial coincidence of the T with its final destination. The latter also expresses directions with animate referents, as in example (6) containing the two-place verb of motion *eō* ‘go’ (s. Luraghi 2011: 218; see also example (7) and the discussion in section 6.2).

<sup>98</sup> Glosses follow the Leipzig Glossing Rules (Bickel *et al.* 2015) with some adjustments. On verbs mood is specified only if not indicative, and voice is indicated if not active; on nouns gender is not indicated unless it is relevant to the discussion. Abbreviations for authors and works quoted follow the *Thesaurus Linguae Latinae* (TLL; <https://thesaurus.badw.de/en/tll-digital/index/a.html>). The full list of abbreviations is shown in Appendix. Examples provide the Loeb translations with some adjustments made by the author.

- (6) *ego*            *ibo*            *ad*            *fratrem*  
 1SG.NOM    go:FUT.1SG    to            brother:ACC.SG  
 'I will go to (my) brother..' (Plaut. *Capt.* 126)

Prototypical landmarks (city names, names of small islands, and the word *domus* 'home') require less morphological coding and occur in plain accusative (see Luraghi 2011 on Latin; Haspelmath 2019 more in general).

Animate Rs are encoded either with the allative PP: *ad* + ACC (example (7)) or with the dative case (example (8)).

- (7) ... [*ad Pistoclerum*            *meum*            *sodalem*]<sub>R</sub>    [*litteras*]<sub>T</sub>    *misi*  
 to Pistoclerus:ACC.SG POSS.ACC.SG friend:ACC.SG letter:ACC.PL send:PRF.1SG  
 'I sent letters to my friend Pistoclerus.' (Plaut. *Bacch.* 389-390)

- (8) [*epistulam*]<sub>T</sub>    [*Lucceio*]<sub>R</sub>            *quam*            *misi*  
 letter:ACC.SG    Lucceius:DAT.SG    REL.ACC.SG    send:PRF.1SG  
 'The letter that I sent to Lucceius ...' (Cic. *Att.* 4.6.4)

In many Indo-European languages the prepositionless dative consistently encodes animate participants encompassing different semantic roles. In Latin the dative marks the third argument of transfer verbs such as *dō* 'give' (i.e. a recipient) and the third argument of communication verbs such as *dicō* 'say' (i.e. an addressee). As an adverbial, the dative case most frequently expresses the roles of beneficiaries and/or maleficiaries (Blake 1994; Luraghi 2010; see also Fedriani & Prandi 2014: 574-578 for a comprehensive account of the Dative role complex).

Passivization in ditransitive constructions often follows the same alignment of encoding (Malchukov *et al.* 2010: 28). Latin ditransitives with indirective alignment display indirective passivization (s. Napoli 2018). As stated by Napoli (2018: 78) «the argument which is passivized invariably corresponds to the direct object, i.e. to the T, whereas the R, if expressed, is retained as a dative». The alternation between dative and *ad* + ACC in the encoding of Rs is also attested in the passive voice. Take as an example the sentence in (9) containing a passive form of *mittō* 'send' and an R argument in the dative case.

- (9) *liber*            *tibi*            *mittetur*  
 book:NOM.SG    2SG.DAT    send:FUT.PASS.3SG  
 'The book will be sent to you ...' (Cic. *Att.* 1.13.5)

It is debated whether the third participant of *mittō* 'send' should be considered part of the argument structure or whether it should be better viewed as an adjunct (e.g. Pinkster 2015). This issue raises interesting questions, which, however, will not be addressed in detail in this paper.

### 3.2. Previous proposals for accounting for the alternation in the encoding of Rs

Latin ditransitive verbs have been studied from a synchronic perspective (see Baños 1998, 2000 on Classical Latin; Sznajder 2012 on the Vulgate; Fedriani 2020 on Merovingian Latin) and within a typologically-oriented approach (s. Napoli 2018). A smaller number of studies have investigated Latin ditransitives in diachrony (s. Fedriani & Prandi 2014; Adams & de Melo 2016; for a comprehensive account on previous studies see also Fedriani & Napoli 2020). The alternation between dative and *ad* + ACC in the encoding of Rs has been attested with transfer verbs (e.g. *dō* 'give', *ferō* 'bring') and communication verbs (e.g. *dicō* 'say', *nuntiō* 'announce') since Early Latin (s. Bennett 1914; Palmer 1988; Baños 1996, 1998; Luraghi 2010; Adams 2013; Fedriani & Prandi 2014; Pinkster 2015; Adams & de Melo 2016; Napoli 2018). According to Lindsay (1907: 20), the constructions with dative Rs and with *ad* + ACC Rs of the verb *dō* 'give' are equivalent. In Plautus, the use of *ad* + ACC instead of the standard dative case is seen as a colloquial trait foreshadowing Romance development. In fact, modern Romance languages continue the prepositional construction (s. Fedriani & Prandi 2014; Adams & de Melo 2016). In Late Latin, *ad* + ACC was increasingly used to express Rs of transfer and communication verbs. The prepositional construction gradually replaces the dative, «taking over all of its uses in Romance languages» (Fedriani & Prandi 2014: 567; see also e.g. Maiden 1996; Zamboni 2000; Napoli 2020). Other scholars acknowledge that transfer verbs have different meanings if used in different constructions (Adams 2013; Fedriani & Prandi 2014; Adams & de Melo 2016). The dative case is used when recipients are in the presence of the sender, whereas *ad* + ACC implies motion over a distance (Adams 2013; Fedriani & Prandi 2014; Adams & de Melo 2016). Pinkster (1990) argues that the PP: *ad* + ACC have been extended to the expression of the third argument (the addressee) of communication events, according to the conduit metaphor (see Reddy 1979 and 'information transfer' in Luraghi 2016), which conceives of words and messages as objects moving during an act of communication. Similarly, Fedriani & Prandi (2014: 596) suggest

that the «gradual spread of the prepositional strategy is diachronically motivated by a metaphorical projection of the Goal onto the Recipient and the Addressee». Luraghi (2010) examined construction alternation of *mittō* 'send' considering referential properties (e.g. animacy) and semantic roles of arguments. She observed that in Caesar's and Cicero's corpora Rs are more often encoded with *ad* + ACC when Ts have human referents. The alternation is related to a difference in the semantic roles of Rs: recipients are marked by the dative, while directions are expressed with *ad* + ACC. Finally, Pinkster (1990) explains construction alternation with *mittō* 'send' making reference to the argumenthood of Rs. He argues that when *mittō* 'send' is constructed as a three-place verb, Rs are marked with the dative. Instead, when it is constructed as a two-place verb, Rs are optional adjuncts taking the PP: *ad* + ACC.

#### 4. Data and methodology

As previously mentioned, the main goal of this work is to describe the construction alternation of *mittō* 'send' with a corpus-based approach. With this purpose, I selected a sample of texts that ranges from Plautus (3rd century BCE) to the Vulgate (5th century CE). The corpus comprises the Latin texts in the PROIEL treebank (Haug & Jøhndal 2008) and in the Ancient Greek and Latin Dependency Treebank (AGLDT; Bamman & Crane 2006), available in the Universal Dependencies (UD) version (de Marneffe *et al.* 2021), as well as the Plautus' corpus. The Latin data in the PROIEL treebank and in the AGLDT include texts from Classical Latin (ca. 90 BCE – 14 CE), and Late Latin (ca. 200 CE – ca. 600 CE). The Plautus' corpus is one of the chief sources of Early Latin (ca. 240 BCE – ca. 90 BCE).<sup>99</sup>

The sample includes different textual genres showing aspects of diastratic and diaphasic variation (s. Palmer 1988; Poccetti *et al.* 1999; Cuzzolin & Haverling 2009; Clackson 2011). In regard to the former, Plautus' comedies have features of Latin spoken at around 200 BCE, which seem to represent a low linguistic variety (Cuzzolin & Haverling 2009: 31). Regarding the latter, Cicero's letters provide evidence of «the language of everyday conversation of the educated élite in Rome» (Cuzzolin & Haverling 2009: 33). Scholars

---

<sup>99</sup> See e.g. Cuzzolin & Haverling (2009) on the periodization of the Latin language.

have pointed out differences between the Latin Cicero uses in his letters and that in his more formal writings, assuming that «some such features in his letters represent the informal variety» (Cuzzolin & Haverling 2009: 33). Late Latin is mostly represented by Jerome's Vulgate. Regarding the Latin language of the Vulgate, it is worth mentioning the contact with Greek and Hebrew sources (see Luraghi & Cuzzolin 2007 on contact-induced phenomena in the early translations of the Gospels and Sznajder 2012 on verbs of saying in the Vulgate).

Considering a diachronic sample comprising different textual genres allows to test Luraghi's (2010) predictions on construction alternation of *mittō* 'send' (see section 3.2), against a larger and more diversified corpus. Moreover, this study could give some insights on the evolution of the ditransitive construction of *mittō* 'send' over the different stages of Latin. Even though I will provide some data regarding the distribution of the two constructions (ad + ACC vs. dative), a systematic diachronic account of the constructional variation of *mittō* 'send' falls beyond the scope of this work.

First, I collected all the occurrences of the lemma *mittō* 'send' in the sample of texts. The occurrences of *mittō* 'send' in Plautus' corpus were extracted with the online resource *Perseus*.<sup>100</sup> As to the texts in the treebanks, I automatically extracted the occurrences of *mittō* 'send' constructed with an object argument in the accusative and an argument expressed either by the dative or by the PP: *ad* + ACC, using *Udeasy*, a tool for querying treebanks in CoNLL-U format (Brigada Villa 2022). Secondly, after having revised the sample, I manually extracted all the ditransitive constructions of *mittō* 'send' with animate third arguments (Rs). I considered finite and non-finite verb forms (i.e. participles and infinitives in predicative function) taking R arguments expressed either by the dative or by *ad* + ACC. The final sample resulted in 157 instances of *mittō* 'send' constructed ditransitively with animate Rs, in the following texts: Plautus' comedies (53x), Caesar's *Commentarii belli Gallici* (42x), Cicero's *Epistulae ad Atticum* (32x), Sallust's *In Catilinam* (1x), and Jerome's Vulgate (29x). This sample comprises texts of Early Latin (the Plautus' comedies), Classical Latin (Caesar's *Commentarii belli Gallici*, Cicero's *Epistulae ad Atticum* and Sallust's *In Catilinam*), and Late Latin (Jerome's Vulgate) stages.

---

<sup>100</sup> The Perseus resource (Crane 2022) is available at <http://www.perseus.tufts.edu/hopper/>.

The ditransitive constructions, together with their contexts of occurrence, have been examined qualitatively. For each construction I focused on the morphosyntactic encoding of Rs (i.e. dative vs. *ad* + ACC) and on the animacy of the T arguments. Regarding the latter, I considered whether each T argument denotes an animate or an inanimate entity. Some quantitative data are provided to give a picture of (i) the distribution of the two constructions in the corpus and (ii) the distribution of the two constructions according to the animacy of T arguments.

## 5. Results

In this Section I show the results of my analysis, which are extensively discussed in Section 6. Consider first the overall distribution of the coding patterns of animate R arguments in my sample.

Table 1. Distribution of the encodings of R arguments in different stages of Latin.

	Dative	<i>ad</i> + ACC	Total
Early Latin	19 (35.8%)	34 (64.2%)	53 (100%)
Classical Latin	15 (20%)	60 (80%)	75 (100%)
Late Latin	6 (20.7%)	23 (79.3%)	29 (100%)
<i>Overall</i>	40 (25.5%)	117 (74.5%)	157 (100%)

As seen in Table 1, R arguments are more frequently encoded with the PP: *ad* + ACC, rather than with the dative case. Out of 157 total occurrences, 117 (74.5%) of Rs are expressed with the PP and 40 (25.5%) of Rs are marked datively. Let us now compare the distributions of the encodings of Rs in the different periods of Latin.

The distributions of encodings in Early Latin, Classical Latin and Late Latin (s. Table 1), show that the prepositional construction is always preferred. In Early Latin, Rs encoded with *ad* + ACC represent 64.2% of the occurrences (34 out of 53), while the dative Rs represent 38.5% of the occurrences (19 out of 53). In Classical and Late Latin, the percentage of dative Rs is still lower: 20% in Classical Latin (15x out of 75) and 20.7% in Late Latin (6x out of 29). Rs with *ad* + ACC, instead, represent 80% of the

occurrences (60 out of 76) in the Classical Latin corpus and 79.3% of the occurrences in the Late Latin corpus (23 out of 29).

It can be pointed out that the percentage of dative Rs decreases from Early to Late Latin. However, dative Rs do not disappear in Late Latin: they occur in similar percentages in Classical Latin and Late Latin. I argue that the survival of the dative pattern is not surprising. Literary texts from late periods do not show much evidence of syntactic changes, as «the written form remained strongly influenced by the language in the literary classics from the Late Republic and the Early Empire» (Cuzzolin & Haverling 2009: 34). As already mentioned, the influences of Hebrew and Greek passages (s. Luraghi & Cuzzolin 2007) should also be taken into account when dealing with the Vulgate's language. These aspects have not been investigated in the present study and should be addressed in future research.

Secondly, consider the overall distribution of the coding patterns of R arguments according to the animacy of T arguments.

Table 2. Encodings of Rs co-occurring with animate/inanimate Ts.

	Dative	<i>ad</i> + ACC	Total
<i>Animate Ts</i>	12 (7.6%)	90 (57.3%)	102 (65%)
<i>Inanimate Ts</i>	28 (17.83%)	27 (17.2%)	55 (35%)
	40 (25.5%)	117 (74.5%)	157 (100%)

As we can see from Table 2, animate Ts more frequently co-occur with Rs encoded by the PP: *ad* + ACC, rather than with dative Rs. Out of 102 instances containing an animate T, Rs are marked with the dative only 12 times (7.6%), while they are expressed with *ad* + ACC 90 times (57.3%). Inanimate Ts instead, co-occur with Rs encoded either by the dative or by *ad* + ACC in similar percentages. Out of 55 instances containing an inanimate T, Rs are marked with the dative 28 times (17.83%), while they are expressed with *ad* + ACC 27 times (17.2%). Let us now compare the distributions of the encodings of Rs according to the animacy of Ts in Early Latin, Classical Latin, and Late Latin.

Table 3. Encodings of Rs co-occurring with animate/inanimate Ts in Early Latin.

	Dative	<i>ad</i> + ACC	Total
<i>Animate Ts</i>	1 (1.9%)	27 (50.9%)	28 (52.8%)
<i>Inanimate Ts</i>	18 (34%)	7 (13.2%)	25 (47.2%)
	19 (35.9%)	34 (64.1%)	53 (100%)

Table 4. Encodings of Rs co-occurring with animate/inanimate Ts in Classical Latin.

	Dative	<i>ad</i> + ACC	Total
<i>Animate Ts</i>	8 (10.7%)	40 (53.3%)	48 (64%)
<i>Inanimate Ts</i>	7 (9.3%)	20 (26.7%)	27 (36%)
	15 (20%)	60 (80%)	75 (100%)

Table 5. Encodings of Rs co-occurring with animate/inanimate Ts in Late Latin.

	Dative	<i>ad</i> + ACC	Total
<i>Animate Ts</i>	3 (10.3%)	23 (79.4%)	26 (89.7%)
<i>Inanimate Ts</i>	3 (10.3%)	0 (0%)	3 (10.3%)
	6 (20.6%)	23 (79.4%)	29 (100%)

As we can see from Table 3, Table 4, and Table 5, in Early, Classical, and Late Latin, the pattern regarding Rs occurring with animate Ts reflects the general tendency: animate Ts are more frequently found with Rs encoded with *ad* + ACC: 50.9% (27 out of 28 occurrences) in Early Latin, 53.3% (40 out of 48x) in Classical Latin and 89.7% (23 out of 26x) in Late Latin. More interesting is the figure concerning inanimate Ts. Whereas in Early and Late Latin inanimate Ts combine more frequently with dative Rs, in Classical Latin inanimate Ts combine more frequently with Rs prepositionally constructed. In Early Latin, out of a total of 25 occurrences, 18 (34%) of inanimate Ts are marked with the dative, whereas only 7 (13.2%) are encoded with the PP: *ad* + ACC. Moreover, inanimate Ts with *ad* + ACC Rs are not found in the Vulgate (s. Table 5). In contrast, in Classical Latin, out of a total of 27 occurrences, 20 (26.7%) of inanimate Ts are encoded with *ad* + ACC, whereas 7 (9.3%) are marked with the dative.

## 6. Discussion

In this section I discuss the results displayed in section 5, focusing on three parameters: (i) animacy of T arguments, (ii) semantic roles of R arguments, and (iii) affectedness of R arguments. To further elaborate on these issues, I provide examples of ditransitive constructions of *mittō* 'send' from my corpus.

### 6.1. Animacy of T arguments

Matching previous findings in the literature (e.g. Luraghi 2010; see also Napoli 2018: 79–84), my sample of occurrences has confirmed that animate Ts tend to co-occur with Rs encoded by the PP: *ad* + ACC. This figure is clear in all the stages of Latin. Consider the sentence in example (10).

- (10) *legatos*            *ad Dumnorigem*    *Haeduum*            *mittunt*  
 deputy.ACC.PL    to Dumnorig:ACC.SG    Aeduan:ACC.SG    send.PRS.3PL  
 '(The Helvetii) send deputies to Dumnorig the Aeduan.' (Caes. *Gall.* 1.9.2)

In (10) the T is expressed by a noun with an indefinite specific human referent. Deputies, messengers, and servants are typical examples of animate Ts co-occurring with Rs encoded with *ad* + ACC. Different animate Ts are found with *ad* + ACC Rs: first- and second-person pronouns (i.e. speech act participants), third-person pronouns, proper nouns, and common (human) nouns (on animacy-hierarchies see e.g. Comrie 1989 and Croft 1990).<sup>101</sup> In my sample, first- and second-person pronouns in the T function never co-occur with Rs marked with the dative. If the T is a first- or second-person pronoun, the R is always expressed by *ad* + ACC (s. (11)).

- (11) *meus*            *gnatus*            [*me*]<sub>T</sub>            [*ad te*]<sub>R</sub>            *misit*  
 POSS.NOM.SG    son:NOM.SG    1SG.ACC    to    2SG.ACC    send:PRF.3SG  
 'My son sent me to you.' (Plaut. *Trin.* 442)

<sup>101</sup> Different types of animacy-hierarchies have been proposed in the literature. I quote here the extended animacy hierarchy: first/second-person pronouns < third-person pronoun < proper names < human common noun < nonhuman animate common noun < inanimate common noun (s. Silverstein 1976; Dixon 1979: 85; Croft 1990: 30).

This pattern has cross-linguistic parallels in modern Romance languages (e.g. French and Italian). In such languages, combinations of bound pronouns are disfavoured when T is a first- or second-person pronoun (see Haspelmath 2004, 2007 on the ditransitive person-role constraint). Consider examples in (12) from French.

French (Haspelmath 2004: 2)

- (12) a. *Agnès me la présentera*  
 Agnès 1SG.DAT 3SG.ACC.F introduce:FUT.3SG  
 'Agnès will introduce her to me.'
- b. \**Agnès me lui présentera*  
 Agnès 1SG.ACC 3SG.DAT.F introduce:FUT.3SG  
 'Agnès will introduce me to her.'
- c. *Agnès me présentera à elle.*  
 Agnès 1SG.ACC introduce:FUT.3SG to 3SG.F  
 'Agnès will introduce me to her.'

In French, a construction with two bound pronouns is possible if the T is a third-person pronoun and the R is a first- or second-person pronoun (example 12a). Different combinations of pronouns (e.g. T as a first-person pronoun and R as a third-person pronoun) result in the ungrammaticality of the sentence (example 12b). When the T referentially outranks the R, the latter is prepositionally constructed (example 12c). In my sample, when the T is a pronoun (including third-person pronouns) and the R is a nominal expression (proper or common nouns), the latter is most often encoded by *ad* + ACC. Out of 21 combinations of pronouns, Rs are expressed with *ad* + ACC 18 times. This finding suggests that the higher the T is in animacy, the more likely the R is to be prepositionally constructed. These predictions are interesting, but need to be pursued further with additional data.

In my sample inanimate nominal themes occur in similar percentages with Rs encoded either with the dative or with *ad* + ACC. However, these percentages differ if we consider Early Latin or Classical and Late Latin. While in Early and Late Latin, inanimate nouns are more frequently found with dative R arguments, in Classical Latin, the pattern is reversed. Regarding Classical Latin, we can consider in Table 6 and Table 7 the encodings of Rs co-occurring with animate/inanimate Ts in Caesar's and Cicero's corpora respectively.

Table 6. Encodings of Rs co-occurring with animate/inanimate Ts in Caesar.

	Dative	<i>ad</i> + ACC	Total
<i>Animate Ts</i>	7 (16.7%)	34 (81%)	41 (97.6%)
<i>Inanimate Ts</i>	0 (0%)	1 (2.4%)	1 (2.4%)
	7 (16.7%)	35 (83.3%)	42 (100%)

Table 7. Encodings of Rs co-occurring with animate/inanimate Ts in Cicero.

	Dative	<i>ad</i> + ACC	Total
<i>Animate Ts</i>	1 (3.1%)	5 (15.6%)	6 (18.7%)
<i>Inanimate Ts</i>	7 (21.9%)	19 (59.4%)	26 (81.3%)
	8 (25%)	24 (75%)	32 (100%)

As we can see from Table 6, in Cesar's corpus, the verb *mittō* 'send' is always found with animate Ts (except for one occurrence of an inanimate T). In fact, in Caesar's *Commentarii belli Gallici mittō* 'send' is used in contexts where human beings (mostly deputies and delegations) are sent to other human beings (mostly generals; s. (10)). The Rs of animate Ts are encoded by the PP: *ad* + ACC (regarding the occurrences of animate Ts with dative Rs see later, this section). What is interesting is that in Cicero's corpus (i.e. *Epistulae ad Atticum*), inanimate Ts are more frequently found with Rs prepositionally constructed (s. Table 7). Out of 26 occurrences, 19 (59.4%) of inanimate Ts are encoded by the PP, *ad* + ACC, whereas only 7 (21.9%) are marked with the dative.<sup>102</sup>

Moreover, to better understand the figure of inanimate Ts in Latin stages, it is worth to examine more closely the referents from a semantic point of view, together with the contexts of occurrences. Inanimate nominal themes denoting entities generally considered beneficial, such as gifts or sums of money (see *munus* 'gift' in Plaut. *Pseud.* 781, *anulus* 'ring' in Plaut. *Curc.* 656 and *pecunia* 'money' in Vulg. *Matth.* 25.27) occur only with dative

<sup>102</sup> As an anonymous reviewer pointed out, it might be interesting to compare the distribution of the two constructions (*ad* + ACC vs. dative) in Cicero's letters with the distribution of the two constructions in more formal writings (e.g. orations). This comparison could help to assess if the preference for the prepositional construction is a feature of Cicero's informal texts, such as the letters.

R arguments (s. Baños 1996; Fedriani & Prandi 2014: 582). Also consider example (13):

- (13) ... *qui*            ***mihi***        *munera*        *mittunt*  
 REL.NOM.PL    1SG.DAT     gift:ACC.PL    send:PRS.3PL  
 'Those who send me gifts...' (Plaut. *Mil Gl.* 710)

Only in one case is a T denoting a gift found with a prepositional R (s. Plaut. *Truc* 589). Abstract entities are attested only with dative Rs (s. Fedriani & Prandi 2014: 582). Consider, for instance, the passage in (14) taken from the Vulgate, where the verb *mittō* 'send' is used with a non-literal meaning and it is combined with an abstract noun.

- (14) *ideo*        *mittit*        ***illis***        *Deus*        *operationem erroris* ...  
 therefore send:PRS.3SG DEM.DAT.PL God:NOM.SG work:ACC.SG mistake:GEN.SG  
 'And for this reason, God shall send them strong delusion.'  
 (Vulg. *II Thess.* 2.11)<sup>103</sup>

The dative is consistently found in collocations or *formulae*, such as in the expression *mittere salutem alicui* 'to send greetings to someone' (see (15), and also Plaut. *Pseud.* 41–43).<sup>104</sup>

- (15) *manu*        *salutem*        *mittunt*        *bene*        ***volentibus***  
 hand:ABL.SG greeting:ACC.SG send:PRS.3PL well wish:PTCP.PRS.DAT.PL  
 'With their hand they send greetings to their well-wishers.' (Plaut. *Pseud.* 1005)

Fedriani (2020: 74) found the analogous greeting formula with a dative R, *alicui salutem mittere (in Domino)* 'to send greetings to someone (in the name of God)', in Merovingian Latin, and regards it as an «archaizing crystallized pattern». Dative Rs are also found in collocations with specific Ts such as *auxilia* 'auxiliary troops' and *copia* 'troops' (see *auxilia* in Caes.

<sup>103</sup> It can be noted that the correspondent Greek passage has the dative of the third-person pronoun *αὐτός*. Regarding the Greek and Semitic influences on the encoding of the addressee of verbs of saying, see Sznajder (2012).

<sup>104</sup> According to reference dictionaries (e.g. the Lewis-Short Dictionary available at the *Perseus Digital Library*, <http://www.perseus.tufts.edu/hopper/text?doc=Perseus:text:1999.04.0059>), the noun *salus* 'health' is attested in various *formulae* of greetings in combinations with different verbs (e.g. *dicō* 'say'). Moreover, these *formulae* are often found at the beginning of letters (e.g. in the headings of Cicero's letters).

*Gall.* 6.32.1 and *copia* in *Caes. Gall.* 7.5.3). In military contexts (e.g. in the passages of Caesar's *Commentarii belli Gallici*) the expression *mittere auxilia alicui* 'to send auxiliary troops to someone' occurs very frequently. Also consider example (16):

- (16) *quod auxilia ...Treveris miserant*  
 that auxiliary:ACC.PL Treveri:DAT.PL send:PPF.3PL  
 'That (the Germans) had sent auxiliaries to (or: for) the Treveri...'  
 (*Caes. Gall.* 6.9.1)

Finally, inanimate nouns denoting objects of communication events, such as letters, poems, writings, and messages (e.g. *littera*, *epistula* 'letter', *scriptum* 'writing', *nuntius* 'message') show the high construction variation, as they occur with Rs expressed either in the dative or with *ad* + ACC in similar proportions (see Fedriani & Prandi 2014, and section 6.2.)

## 6.2. Semantic roles of R arguments

Rs are encoded with *ad* + ACC when they are not in the presence of the sender and Ts are transported across space (Adams 2013: 279, Fedriani & Prandi 2014; see also the caused-motion construction in Goldberg 2006). Rs expressed with *ad* + ACC can be conceived as (animate) directions, since they can be «metonymically viewed as places where animate participants reside» (Kittilä 2014: 171; see also Fedriani & Prandi 2014: 580). In example (10), deputies are sent to the headquarters of Dumnorix, the Aeduan general. In some contexts, a direction reading of Rs is corroborated by the co-occurrence with spatial constituents, such as direction or source constituents. In Latin the source is encoded with prepositions *ex*, *ab*, and *de* and the ablative. With city names, names of small islands, and sometimes names of countries, the source can be expressed with the plain ablative (s. Luraghi 2010: 32; Pinkster 2015). Consider examples (17) and (18):

- (17) *...te in Alidem mittam ad patrem*  
 2SG.ACC in Elis:ACC.SG send:SBJV.PRS.1SG to father:ACC.SG  
 '(It was arranged that) I send you to Elis, to (my) father.' (Plaut. *Capt.* 379)
- (18) *quas ad te Thessalonica misit*  
 REL.ACC.PL to 2SG.ACC Thessalonica:ABL.SG send:PRF.3SG  
 '(In regard to the letters) that he sent you from Thessalonica.' (Cic. *Att.* 1.17.4)

The R is combined with a direction constituent expressed by *ad* + ACC in example (17) and with a source constituent marked by a prepositionless ablative in example (18). The dative marks Rs conceived as (concretely) receiving the theme after a dispatchment, i.e. prototypical recipients. In contexts such as example (19), the verb *mittō* 'send' expresses a prototypical event of transfer: (i) the T is an inanimate concrete entity, (ii) the dispatching of the T implies a transfer of its possession/control properties towards a R, and (iii) the R is high in animacy and control/volition.

- (19) *quae mihi antea signa misisti ...*  
 REL.ACC.PL 1SG.DAT earlier statue:ACC.PL send.PRF.2SG  
 'The statues that you sent me earlier ...' (Cic. *Att.* 1.4.3)

In sentence (19), *mittō* 'send' is used in a sense similar to 'give': the initiation of the event and its completion are construed as a single event with the three participants activated at the same time (Kittilä 2006).

In some contexts, the two constructions might be used to disambiguate between semantic roles.

- (20) *Cato ad me litteras misit*  
 Cato:NOM.SG to 1SG.ACC letter:ACC.PL send.PRF.3SG  
 'Cato sent me the letters.' (Cic. *Att.* 7.1.17)

- (21) *litteras quas ad Pompeium*  
 letter:ACC.PL REL.ACC.PL to Pompey:ACC.SG  
  
*scripsi tibi misi*  
 write.PRF.1SG 2SG.DAT send.PRF.1SG  
 'I sent you the letters that I wrote to Pompey'. (Cic. *Att.* 3.8.3)

In (20) *ad* + ACC marks the endpoint of the event of dispatching. In (21) the dative encodes the argument to whom the T will be handed over, whereas *ad* + ACC encodes the addressee of the event of writing (s. Martín Rodríguez 1991: 97; Adams 2013: 280; Fedriani & Prandi 2014: 583–584). In other languages, direction markers are also used to express targets of communication events, i.e. addressees (on which see the metaphor of the transfer of information in Reddy 1979, Luraghi 2016).

Ts similar from a semantic point of view (especially in writings and letters) are found with Rs marked differently. Compare examples (22) and (23).

- (22) *poemata ...            ad            me            mittas*  
 poem:ACC.PL        to            1SG.ACC        send:SBJV.PRS.2SG  
 '(I wish) you would send me (any) poems ...' (Cic. *Att.* 1.16.18)
- (23) *... illa                    scripta                    mittebam            tibi*  
 DEM.ACC.PL        writing:ACC.PL        send:IMPF.1SG        2SG.DAT  
 '...(When) I sent you those writings.' (Plaut. *Epid.* 138)

When the third argument is both the endpoint and the recipient of the event, the two constructions are considered interchangeable (Adams & de Melo 2016: 94). The choice of one construction rather than the other might depend on the intention to highlight one facet of the meaning of the verb: the *ad* + ACC construction highlights the movement of T whereas the dative construction highlights the reception of the T. The use of *ad* + ACC to express true recipients is motivated by a «metaphorical mapping featuring the recipient as if it was a metaphorical goal» (Fedriani & Prandi 2014: 590).

### 6.3. Other parameters: affectedness of R arguments

Another parameter that might contribute to the alternation between dative and *ad* + ACC is the affectedness of the R. As already mentioned, dative Rs are found in contexts of prototypical transfer events. These Rs are highly affected by the events as they become possessors of the received T (see examples (13) and (19), for instance). Dative recipients also make some use of the T, so that the transfer event positively (or negatively) affects them. In contrast, Rs encoded with *ad* + ACC are less affected by a transfer event because they do not gain possession/control of the T. They are not supposed to make use of the T either, as they might be unaware of the transfer and even refuse to accept it (Kittilä 2006). The affectedness of Rs might be related to the inherent temporal nature of the transfer. Prototypical transfer events with dative recipients are (generally) permanent as they are conceived as more resultative and as more focused on the completion of the event (the phase of reception of the T) rather than on its initiation (the phase of movement of the T; cf. Kittilä 2007). The permanent nature of the event is particularly evident when the T transferred is a gift. Transfer events with *ad* + ACC Rs are often temporary, instead, because they are conceived less as resultative and more as focused on the initiation of the event (the movement of T) rather than on its completion (Kittilä 2007). In a related fashion, Daniel

(2014: 225-226) argues that in some East Caucasian languages temporary Rs feature an allative marking, while permanent Rs are datively-marked.

The affectedness of Rs can also be considered in non-prototypically transfer events. Consider example (24).(24)

- (24) *tempestatem* *Neptunus* ... *nobis* *misit* ...  
 storm:ACC.SG Neptune:NOM.SG 1PL.DAT send:PRF.3SG  
 '(Immortal gods), what a storm Neptune sent us (last night)!' (Plaut. *Rud.* 83-84)

Despite not being the recipient of a prototypical transfer event, the R in (24) is marked with the dative case. This R is characterized by a high degree of affectedness, as it can be conceived of as the maleficiary of the event. Nevertheless, the presence of the dative in (24) is not surprising, as the T is an inanimate entity.

I suggest that the affectedness of the R could explain instances contradicting the tendencies observed. Consider example (25):

- (25) *et* *tibi* *ego* *misi* *mulierem*  
 and 2SG.DAT 1SG.NOM send:PRF.1SG girl:ACC.SG  
 'And I sent you the girl...' (Plaut. *Asin.* 171)

Differently from what was observed in most cases, in (25) a human T occurs with a R in the dative. In this passage taken from the *Asinaria*, Clereeta is speaking with Diabolus, who is infatuated with her daughter. Clereeta claims that she has sent her daughter to Diabolus, to fulfil his sexual desires, in exchange for money.<sup>105</sup> I argue that, in such contexts, the dative might be used to stress the high involvement of the R in the event. The transfer event is focused on its final phase (the reception of the T) and the R "benefits from the T received". In brief, in my sample the dative marks Rs which make use of the T transferred, or that are affected in some way by a transfer event (physical or more figurative as in examples (24) and (25)).

<sup>105</sup> The rest of the passage is quoted here. Plaut. *Asin.* 172: *pari datum hostimentum est, opera pro pecunia*. 'A fair return has been given for a fair price, service for money.'

## 7. Conclusion

In this paper I have examined argument structures and alternations of the Latin ditransitive verb of physical transfer *mittō* 'send'. I have focused on the alternation between the dative and the PP: *ad* + ACC in the encoding of animate Rs. I tried to describe and explain this alternation by considering variables taken into account in the typological literature (e.g. Haspelmath 2007; Kittilä 2008; Malchukov *et al.* 2010), and in previous studies concerned with argument alternations of Latin ditransitives (Fedriani & Prandi 2014; Napoli 2018). I have also based my predictions on corpus data. This preliminary analysis led to promising results regarding three dimensions: (i) animacy of T arguments; (ii) semantic roles of R arguments, and (iii) affectedness of R arguments.

- i. Whereas animate Ts consistently co-occur with Rs encoded by the PP: *ad* + ACC (s. Luraghi 2010), inanimate Ts are constructed either with dative or *ad* + ACC Rs in similar percentages. In Early Latin, however, inanimate Ts are more frequently attested with dative Rs. Moreover, in Late Latin inanimate Ts are never found with Rs prepositionally constructed. To better understand this pattern, I looked at inanimate Ts to see if some types of referents were related to construction alternations. I argue that the animacy of the T plays a role in the construction alternation of R arguments. I suggest that the default encoding of Rs occurring with animate Ts is the PP: *ad* + ACC and that the default marking of inanimate Ts is the dative case. However, inanimate Ts allow for a prepositional construction of Rs under some circumstances. For instance, the Rs of objects of communication of events (e.g. letters and writings), being both the endpoint and the concrete recipient of the transfer, are frequently encoded with *ad* + ACC. Nevertheless, the hypothesis that a differential marking of R arguments is based upon the animacy of the T should be tested with other ditransitives, to determine whether it is independent of the lexical semantics of the verb (i.e. the motion component entailed by a verb such as *mittō* 'send').
- ii. Differences in the semantic roles of R-like arguments are mirrored in the alternating use of the plain dative and of *ad* + ACC: the former marks (true) recipients; the latter marks (animate) directions (Adams & de Melo 2016).
- iii. Whereas the dative marks highly affected R arguments, the PP: *ad* + ACC encodes less affected R arguments. The dative is frequently attested with Rs that make use of the Ts transferred (e.g. in combinations with nouns denoting gifts).

Further corpus-based investigations are needed to deepen our understanding of construction alternation with Latin transfer verbs. In regard to referential properties, it would be interesting to more closely examine construction alternation with Ts that are high in animacy (e.g. first- and second-person pronouns). As animacy, affectedness, and semantic roles are closely interrelated, it would be important to consider the interplay of these three factors to better account for the construction alternation of the transfer verb *mittō* 'send'. Parameters not addressed here, such as the pragmatic properties of T and R arguments and constituent order, should be investigated as well. Besides typologically-related variables, aspects concerning the genre(s) and the chronological period(s) of the texts included in the sample, should also be taken into account. Considering a sample with different textual types might help to understand whether one variant is tied to a particular genre. For instance, the high presence of Rs prepositionally constructed might be a colloquial feature typical of Cicero's letters that is avoided in more formal texts. Enlarging the sample with Cicero's more formal texts, such as the orations, could shed light on this aspect. Finally, the diachronic dimension should be further investigated in order to increase our understanding of the evolution of the two constructions from Early Latin to Romance (s. Fedriani & Prandi 2014; Fedriani & Napoli 2020).

## Appendix

### Acknowledgments

I would like to thank two anonymous reviewers for the useful comments and Luca Brigada Villa for the help in the automatic extraction of the data.

### Abbreviations

1	first person
2	second person
3	third person
A	agent argument
ABL	ablative
ACC	accusative
DAT	dative
DEM	demonstrative
F	feminine

FUT	future
GEN	genitive
IMPF	imperfect
NOM	nominative
P	patient argument
PASS	passive
PL	plural
POSS	possessive
PP	prepositional phrase
PPF	pluperfect
PRF	perfect
PRS	present
PTCP	participle
R	recipient-like argument
REL	relative pronoun
SBJV	subjunctive
SG	singular
T	theme argument

### **Authors and works quoted in the examples**

Asin.	Asinaria
Att.	Epistulae ad Atticum
Bacch.	Bacchides
Caes.	Caesar
Capt.	Captivi
Cic.	Cicero
Curc.	Curculio
Gall.	Commentarii belli Gallici
Epid.	Epidicus
Matth.	Evangelium sec. Matth(a)eum
Men.	Menaechmi
Mil. Gl.	Miles Gloriosus
Plaut.	Plautus
Pseud.	Pseudolus
Rud.	Rudens
Trin.	Trinummus
Truc.	Truculentus
Vulg.	biblia sacra iuxta vulgatam versionem
I. II Thess.	Pauli epistulae ad Thessalonicenses

## References

- Adams, J. (ed) (2013). *Social Variation and the Latin Language*. Cambridge: Cambridge University Press.
- Adams, J. & de Melo, W. (2016). Ad versus the dative: from early to late Latin. In: Adams, J. & Vincent, N. (eds) *Early and Late Latin: Continuity or Change?*, 87–131. Cambridge: Cambridge University Press.
- Bamman, D. & Crane, G. (2006). The design and use of a Latin dependency treebank. In: Hajič, J. & Nivre, J. (eds) *Proceedings of the Fifth International Workshop on Treebanks and Linguistic Theories (TLT 2006)*, 67–78. Prague: ÚFAL MFF UK.
- Baños Baños, J. M. (1996). Litteras Neroni / ad Neronem mittere: ¿alternancia dativo / ad + acus?. In: Agud, A., Fernández Delgado, J. A. & Ramos Guerreira, A. (eds) *Las lenguas de corpus y sus problemas lingüísticos*, 217–236. Madrid: Ediciones Clásicas.
- Baños Baños, J. M. (1998). Sintaxis y semántica del dativo objeto indirecto: su concurrencia con ad-acus en latín clásico. In: Torrego, M. E. (ed) *Nombres y funciones: estudios de sintaxis griega y latina*, 11–41. Madrid: Ediciones Clásicas.
- Baños Baños, J. M. (2000). Vulgarismos sintácticos en Plauto (II): Quae ad patrem vis nuntiari (CAP.360). In: García-Hernández, B. (ed) *Latín vulgar y tardío. Homenaje a Veikko Väänänen (1905–1997)*, 1–15. Madrid: Ediciones Clásicas.
- Bennett, Ch. E. (ed) (1914). *Syntax of Early Latin. Volume 2: The Cases*. Leipzig: Stauffer.
- Bickel, B., Comrie, B. & Haspelmath, M. (2015). *Leipzig Glossing Rules*. Available online: <https://www.eva.mpg.de/lingua/resources/glossing-rules.php>, Accessed on 2023-12-06).
- Blake, B. J. (ed) (1994). *Case*. Cambridge: Cambridge University Press.
- Brigada Villa, L. (2022). UDeasy: a Tool for Querying Treebanks in CoNLL-U Format. In: Banski, P., Barbaresi, A., Clematide, S., Kupietz, M. & Lungen, H. (eds) *Proceedings of the CMLC-10 Workshop (@LREC 2022)*, 16–19. Paris: European Language Resources Association.
- Clackson, J. (ed) (2011). *A Companion to the Latin Language*. Oxford: Blackwell.
- Comrie, B. (ed) (1989). *Language Universals and Linguistic Typology. 2nd revised edition*. Chicago: University of Chicago Press.
- Crane, G. (2022). The Perseus Digital Library and the future of libraries. *International Journal on Digital Libraries*, 24(2), 117–128.
- Croft, W. (ed) (1990). *Syntactic Categories and Grammatical Relations*. Chicago: University of Chicago Press.
- Cuzzolin, P. & Haverling, G. (2009). Syntax, sociolinguistics, and literary genres. In: Baldi, P. & Cuzzolin, P. (eds) *New Perspectives on Historical Latin Syntax. Volume 1*, 19–64. Berlin: Mouton de Gruyter.
- Daniel, M. (2014). Against the Addressee of speech – Recipient metaphor: Evidence from East Caucasian. In: Luraghi S. & Narrog, H. (eds) *Perspectives on semantic roles*, 201–239. Amsterdam: John Benjamins.
- de Marneffe, M.C., Manning, C. D., Nivre J. & Zeman D. (2021). Universal Dependencies. *Computational Linguistics*: 47(2), 255–308.
- Dixon, R. M. W. (1979). Ergativity. *Language*, 55(1), 59–138.
- Fedriani, C. & Prandi, M. (2014). Exploring a diachronic (re)cycle of roles: The Dative complex from Latin to Romance. *Studies in Language*, 38(3), 566–604.

- Fedriani, C. (2020). The spread of the ad construction in Merovingian Latin: identifying semantic paths in the domain of ditransitives. In: Fedriani, C. & Napoli, M. (eds) *The Diachrony of Ditransitives*, 61–96. Berlin: Mouton de Gruyter.
- Fedriani, C. & Napoli, M. (eds) (2020). *The Diachrony of Ditransitives*. Berlin: Mouton de Gruyter.
- Goldberg, A. E. (ed) (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.
- Goldberg, A. E. (ed) (2006). *Constructions at work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Haspelmath, M. (2004). Explaining the Ditransitive Person-Role Constraint: A usage-based approach. *Constructions*, 1, 1–49.
- Haspelmath, M. (2007). Ditransitive alignment splits and inverse alignment. *Functions of Language*, 14(1), 79–102.
- Haspelmath, M. (2011). On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology*, 15, 535–567.
- Haspelmath, M. (2014). Arguments and Adjuncts as Language-Particular Syntactic Categories and as Comparative Concepts. *Linguistic Discovery*, 12(2), 3–11.
- Haspelmath, M. (2015). Ditransitive constructions. *Annual review of Linguistics*, 1, 19–41.
- Haspelmath, M. & Hartmann, I. (2015). Comparing verbal valency across languages. In: Malchukov A. & Comrie, B. (eds) *Valency Classes in the World's Languages*, 41–72. Berlin: Mouton de Gruyter.
- Haspelmath, M. (2019). Differential place marking and differential object marking. *STUF – Language Typology and Universals*, 72(3), 313–334.
- Haug, D. T. & Jøhndal, M. L. (2008). Creating a Parallel Treebank of the Old Indo-European Bible Translations. In: Sporleder, C., van den Bosch, A. & Grover, C. (eds) *Proceedings of the Second Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2008)*, 27–34. Association for Computational Linguistics.
- Hopper, P. J. & Thompson, S. A. (1980). Transitivity in grammar and discourse. *Languages*, 56(2), 251–299.
- Kittilä, S. (2005). Recipient-prominence vs. beneficiary-prominence. *Linguistic Typology*, 9, 269–297.
- Kittilä, S. (2006). The anomaly of the verb 'give' explained by its high (formal and semantic) transitivity. *Linguistics*, 44, 569–612.
- Kittilä, S. (2007). On the encoding of transitivity-related features on the indirect object. *Functions of Language*, 14(1), 149–164.
- Kittilä, S. (2008). Animacy effects on differential Goal marking. *Linguistic Typology*, 12(2), 245–268.
- Kittilä, S. (2014). The (non-)prototypicality of Direction: The (allative and illative) case(s) of Finnish. In: Luraghi S. & Narrog, H. (eds) *Perspectives on semantic roles*, 151–179. Amsterdam: John Benjamins.
- Kittilä, S. (2022). A typology of Goal-Source marking in transfer events. *Studies in Language*, 46(2), 352–375.
- Kittilä, S. & Ylikoski, J. (2011). *Remarks on the coding of Goal, Recipient and Vicinal Goal In European Uralic*. In: Kittilä, S., Västi, K. & Ylikoski, J. (eds) *Case, Animacy and Semantic Roles*, 29–64. Amsterdam/Philadelphia: John Benjamins.

- Korn, A. & Malchukov, A. (eds) (2018). *Ditransitive Constructions in a Cross-Linguistic Perspective*. Wiesbaden: Reichert.
- Levin, B. & Rappaport Hovav, M. (eds) (2005). *Argument realization*. Cambridge: Cambridge University Press.
- Lindsay, W. M. (ed) (1907). *Syntax of Plautus*. Oxford: Clarendon Press.
- Luraghi, S. (2010). Adverbials. In: Baldi, Ph. & Cuzzolin, P. (eds) *New Perspectives on Historical Latin Syntax. Volume 2*, 19–96. Berlin: Mouton de Gruyter.
- Luraghi, S. (2011). The coding of spatial relations with human landmarks: From Latin to Romance. In: Kittilä, S., Västi K. & Ylikoski, J. (eds) *Case, Animacy and Semantic Roles*, 209–234. Amsterdam: John Benjamins.
- Luraghi, S. (2016). The mapping of space onto the domain of benefaction and some unpredicted trends in semantic change. *Linguistics*, 54(2), 339–379.
- Luraghi, S. & P. Cuzzolin, (2007). Mediating culture through language. Contact-induced phenomena in the early translations of the Gospels. In: Ramat, P. & Roma, E. (eds) *Europe and the Mediterranean as linguistic areas: convergencies from a historical and typological perspective*, 133–158. Amsterdam: John Benjamins.
- Maiden, M. (ed) (1996). *A Linguistic History of Italian*. New York: Longman.
- Malchukov, A., Haspelmath, H. & Comrie, B. (eds) (2010). *Studies in Ditransitive Constructions. A Comparative Handbook*. Berlin: Mouton de Gruyter.
- Martín Rodríguez, A. M. (ed) (1999). *Los verbos de 'dar' en latín arcaico y clásico: Análisis estructural de un campo semántico*. Las Palmas de Gran Canaria: Universidad de Las Palmas de Gran Canaria.
- Napoli, M. (2018). Ditransitive Verbs in Latin: A Typological Approach. *Journal of Latin Linguistics*, 17(1), 51–91.
- Napoli, M. (2020). Old Italian ditransitive constructions: between alternation and change. In: Fedriani, C. & Napoli, M. (eds) *The Diachrony of Ditransitives*, 97–136. Berlin: Mouton de Gruyter.
- Newman, J. (ed) (1996). *Give: A Cognitive Linguistic Study*. Berlin: Mouton de Gruyter.
- Palmer, L. R. (ed) (1988). *The Latin Language*. Norman Oklahoma: University of Oklahoma Press.
- Perek, F. (ed) (2015). *Argument structure in usage-based construction grammar*. Amsterdam: John Benjamins.
- Pinkster, H. (1990). The development of cases and adpositions in Latin. In: Pinkster, H. & Genee, I. (eds) *Unity in diversity*, 195–209. Dordrecht: Foris.
- Pinkster, H. (ed) (2015). *The Oxford Latin Syntax*. Oxford: Oxford University Press.
- Pocchetti, P., Poli, D. & Santini, C. (eds) (1999). *Una storia della lingua latina*. Roma: Carocci.
- Reddy, M. (1979). The Conduit metaphor. In: A. Ortony (ed) *Metaphor and Thought*, 164–201. Cambridge: Cambridge University Press.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104(3), 192–233.
- Silverstein, M. (1976). Hierarchy of features and ergativity. In: Dixon, R. M. W. (ed) *Grammatical Categories in Australian Languages*, 112–171. New Jersey: Humanities Press.
- Sznajder, L. (2012). Dixit autem serpens ad mulierem / Mulieri quoque dixit: la double expression de l'allocutaire dans les propositions introductrices de discours directs dans

- la Vulgate. In: Biville F., Lhommé M. K. & Vallat D. (eds) *Latin vulgaire-latin tardif IX, Actes du IXe colloque international 'latin vulgaire-latin tardif'*, (Lyon, 2–6 septembre 2009), 271–290. Lyon: Collection de la maison de l'Orient et de la Méditerranée 49.
- Tsunoda, T. (1981). Split case-marking in verb types and tense/aspect/mood. *Linguistics*, 19, 389–438.
- Tsunoda, T. (1985). Remarks on Transitivity. *Journal of linguistics*, 21, 385–396.
- Zamboni, A. (ed) (2000). *Alle origini dell'italiano. Dinamiche e tipologie della transizione dal latino*. Roma: Carocci.
- Zúñiga, F. & Kittilä, S. (eds) (2010). *Benefactives and Malefactives: Typological perspectives and case studies*. Amsterdam: John Benjamins.

## 8. The encoding of the manipulee in Spanish: the competition between dative and accusative

BOJANA MIKELENIĆ, METKA BEZLAJ

Keywords: *Spanish; clitic variation; manipulation verbs; leísmo*

In this study, we analyse the alternation between dative and accusative clitics in nine manipulation verbs in Spanish. We use a reference corpus *CORPES XXI* to extract all examples of these verbs in constructions with a prepositional object in order to observe the current status of clitic alternation, or *leísmo*, and its geographical distribution in different Spanish-speaking areas. We look into the variables of type of clause, gender, and number of the clitics' referent as well. The analysis has shown that the process of replacing the etymological dative with the accusative with these verbs is uneven, that is, it varies from area to area and from verb to verb.

*Original scientific paper*

### 1. Introduction

The present study deals with selected Spanish verbs that show change in their government, i.e. we focus on a certain number of verbs that can present their pronominal object as accusative or dative. The literature review (e.g. Fernández-Ordóñez 1999) showed that these verbs are fairly frequent in Spanish. However, due to practical reasons, as well as the homogeneity of our analysis, we decided to consider only those that can be syntactically and semantically categorized into a coherent verb group. More specifically, we looked at verbs such as *animar* 'to encourage', *autorizar* 'to authorize', *forzar* 'to force', etc. that could be included in the so-called manipulation verbs (Givón 2001: 151–152). The reinterpretation of the clitic with these verbs is not consistent and it depends not only on dialectological factors but can vary from verb to verb. The data was gathered from a corpus: *Corpus del Español del Siglo XXI – CORPES XXI*, which is representative of the Spanish

language spoken in Spain and in Latin America, so it allows for diatopic analyses.

In that sense, the overall goal of our work is to examine the current status of clitic alternation within this group of verbs, which is a topic that is also related to the relationship between direct and indirect object in Spanish. The aim of our analysis is three-fold: firstly, we want to examine which third-person clitics are used most frequently with manipulation verbs selected for the study; secondly, our goal is to describe various types of this usage, especially with respect to different characteristics of the object (such as case, number, gender, person, and type of complement clause). Finally, our work also sets out to account for substantial regional variation of the phenomenon, which is, as already stated, an important factor when it comes to dealing with this topic. In other words, the study includes language variety as an important variable, and we investigate how different Spanish-speaking countries compare in this regard.

## 2. Theoretical background

The behaviour of verbal clitics in Romance is different than that of a noun phrase in its place; essentially, their distribution is not the same. As Halpern (2017) shows on a French example, but the same can also be applied to Spanish, a clitic precedes the verb it modifies, while a nonclitic object follows it. As a consequence, these are not the so-called *simple clitics*, like reduced auxiliaries and pronouns in English that are basically just accentless words but could be named *special clitics* (Halpern 2017). There is substantial research being done regarding the similarities and differences of these clitics with the inflectional affixes (agreement morphology) and their status in the argument structure of a verb (cf. Halpern 2017; Mayer 2017; Spencer & Luís 2012). This also connects to the relationship between direct and indirect object in general.

In his paper on coding properties of Spanish objects, García-Miguel (2015) investigates clitic doubling, *a*-marking, and clitic case as a means of understanding the boundaries between direct and indirect object. Clitic doubling (or object doubling), a frequent topic in studies about clitics in Spanish, is possible only when the object (direct or indirect) is introduced by the preposition *a* (*a*-marking; see section 3 for more details). The author considers these properties as variables and observes their frequency

in a corpus and their distribution in different types of constructions. He concludes that, even though all three variables are dependent on several factors (animacy and definiteness of the referent for *a*-marking, topicality and accessibility for clitic doubling, animacy, gender, and process type for clitic case), considerable dialectal variation has to be taken into account for clitic case.<sup>106</sup>

This suggests that one of the principal means of distinguishing between the two objects loses its strength and the boundaries between them become unclear. On the one hand, De Mello (2002) claims that *leísmo*, the use of the clitic *le* for direct object, does not exist in Latin America and that the verbs that alternate between usage of dative and accusative actually alternate between usage of indirect and direct object, respectively. Even though De Mello considers only Latin American data, the line of reasoning is obvious: dative marks the indirect and accusative the direct object. On the other hand, Company Company (2001, 2003) uses diachronic data to identify the switch in case marking, going so far as to argue that the dative promotion displaces the accusative as the usual and more frequent marking of a direct object. A third route would be to accept the murky boundaries and view objecthood as a continuum, with direct and indirect object on either side of it, and with only limited generalizations possible (García-Miguel 2015).

In the following sections, we will be looking more closely into the clitic case. The clitic variation in Spanish is described through three different phenomena: *leísmo*, which we have already mentioned, but also *laísmo*, and *loísmo*. However, to understand them, we first need to introduce the clitic system in Spanish.

## 2.1. The clitic system in Spanish and clitic variation

In Spanish, like in some other Romance languages, first and second person (singular and plural) clitics show syncretism for case. However, case distinctions are present among third person (singular and plural) clitics. Gender is clearly marked in accusative forms, while there is no gender distinction among third person dative clitics (Table 1).

---

<sup>106</sup> See Giuliani (*this volume*, Chapter 7) for an account of the encoding of the third argument of the Latin verb *mittō* 'send' based on the animacy of the referents.

Table 1. A part of personal pronoun paradigm in Spanish<sup>107</sup>

	Strong subject pronouns	Strong object pronouns	Accusative clitics	Dative clitics
1SG	<i>yo</i>	<i>mí</i>		<i>me</i>
2SG	<i>tú</i>	<i>ti</i>		<i>te</i>
3M.SG		<i>él</i>	<i>lo</i>	<i>le</i>
3F.SG		<i>ella</i>	<i>la</i>	
1PL	<i>nosotros / nosotras</i>			<i>nos</i>
2PL	<i>vosotros / vosotras</i>			<i>os</i>
3M.PL		<i>ellos</i>	<i>los</i>	<i>les</i>
3F.PL		<i>ellas</i>	<i>las</i>	

The forms *lo*, *la*, *los*, *las* typically correspond to the direct object and *le*, *les* to the indirect one, a usage known as *etymological*, because it follows the case system of Latin demonstratives (Fernández-Ordóñez 1999: 1319). Nevertheless, in present-day Spanish there are various uses that diverge from this original accusative/dative distinction. This variation can be accounted for by three phenomena mentioned before: *leísmo*, *laísmo*, and *loísmo*. *Leísmo* consists of the use of dative clitics for the direct object (s. Vinja 1998: 149; Alarcos Llorach 1999: 202). For example, in the sentence (1) the clitic *le* is used to double the direct object *A Pedro*. On the other hand, the use of *lo* for indirect object is standardly called *loísmo*,<sup>108</sup> and this usage can be seen in the example (2), where *lo* is used instead of the indirect *le*. Finally, *laísmo* refers to the usage of *la* instead of *le*, as in (3), where *la* refers to the indirect object.

- (1) *A Pedro le vi ayer.*  
 ‘Peter, I saw [**him**] yesterday.’
- (2) *Lo dije que se fuera.*  
 ‘I told **him** to go away.’

<sup>107</sup> This table only shows the part of this paradigm that interests us here; for a more detailed account, including regional variations, cf. *Nueva gramática de la lengua española* – NGLÉ (2009: §16.3).

<sup>108</sup> Fernández-Ordóñez (1999: 1320) mentions that for some authors *loísmo* means the usage of *lo* for the direct object as opposed to *le*. We will not be using the said term in this sense.

(3) *La di un libro.*‘I gave **her** a book.’

Out of these three phenomena, *leísmo* is the most widespread and can be found both in Spain and in Latin America, while *loísmo* and *laísmo* are restricted only to certain areas of Spain (NGLE 2009: §16.10; Fernández-Ordóñez 1999: 1322). This is the reason why we have chosen *leísmo* as the central focus of our research, even though it can be difficult to separate the three, as they have frequently been studied together.

Fernández-Ordóñez (1999, 1993) gives an overview of clitic variation in Spanish, summarizing the most important hypotheses about its origins and the research done during the 19<sup>th</sup> and 20<sup>th</sup> century. She explains that the authors contrast *etymological* and *non-etymological* or *anti-etymological* usage, depending on whether it follows the regimen of the verb in question in Latin or not (Fernández-Ordóñez 1999: 1320). The first traditional hypothesis that was proposed to explain these phenomena is based on the idea that case was eliminated in favour of gender; in other words, the syncretism between first person and second person clitics was extended to third person clitics. In that sense, Fernández Ramírez (1987: 43) and Lapesa (1968: 541) see the origins of clitic variation in a tendency to mirror the demonstrative pronouns model (*este, esta, esto*), where *le* would be used for masculine, *la* for feminine, and *lo* for neutral gender. The second hypothesis suggests that Spanish tends to differentiate between personal and non-personal entities, and this favoured the emergence of *leísmo* (Fernández Ramírez 1964: 280). However, these explanations are not sufficient to account for different facets of these phenomena. For example, the first hypothesis cannot explain why *leísmo* occurs more frequently with personal direct objects, while the second cannot account for examples of *leísmo* with inanimate non-personal objects, nor origins of *laísmo* and *loísmo* (Fernández-Ordóñez 1999: 1321–1322).

This prompted linguists to consider dialectological differences in the Spanish-speaking world, which proved to be a turning point in the interpretation of clitic variation (Fernández-Ordóñez 1999: 1322). Working independently on Spanish spoken data, García González (1981) and Klein-Andreu (1993) concluded that continuity and discontinuity of the object’s referent, or it being countable or uncountable, was as important for the choice of clitic as gender. This system is known as *referential*, as opposed to the *etymological*. In light of this, Fernández-Ordóñez (1999: 1322) concludes that we should distinguish between three situations: 1) *leísmo* in the areas that keep the case distinction, 2) *leísmo* and perhaps *loísmo* in the areas where

Spanish comes in contact with non-Indo-European languages (Quechua, Aymara, Guaraní, and Basque), and 3) clitic confusions (*leísmo*, *laísmo*, and *loísmo*) in the areas where the referential system is at work. Concerning the first situation, Fernández-Ordóñez (1999: 1323–1339) distinguishes different types of verbs that alternate the clitics and have been described in literature as cases of *leísmo*, but do not, strictly speaking, fall into that category. With these verbs dative is not being used where the accusative should be, but on the contrary, these were originally intransitive constructions that are presently being transformed into transitive ones (Fernández-Ordóñez 1999: 1323). In this study we consider this type of verbs: we will explain more about them in section 3.

It is also worth mentioning that there are different types of *leísmo*, depending on animacy, gender and number of clitic's referent (Fernández-Ordóñez 1999: 1319; NGLÉ 2009: §16.8c). The most frequent type of *leísmo* has the clitic referring to a singular masculine and personal direct object (4). It is followed by the singular and masculine clitic which refers to an inanimate object (5), that is in turn followed by two types of plural *leísmo*: clitic in masculine gender referring to a person (6) and clitic in masculine gender referring to an inanimate object (7). Finally, the rarest of all, *leísmo* with a feminine referent, usually a person, both singular (8) and plural (9).

- (4) *Mario viene esta noche, así que le llevaré a cenar.*  
 'Mario is coming this evening, so I'll take **him** to dinner.'
- (5) *Mi madre me envió un paquete y le recogeré mañana.*  
 'My mother sent me a package and I'll pick **it** up tomorrow.'
- (6) *Hace tiempo que no veo a tus hijos, pero Juan les acaba de ver.*  
 'It's been some time since I've seen your children, but Juan just saw **them**.'
- (7) *Compré los pasteles y les probé enseguida.*  
 'I've bought the cakes and tried **them** immediately.'
- (8) *¿Conoces a Marta? No, no le conozco.*  
 'Do you know Marta? No, I don't know **her**.'
- (9) *Las chicas están cantando, les oigo de mi cuarto.*  
 'The girls are singing, I hear **them** from my room.'

A separate subtype of *leísmo* where the referent is a person is the so-called *leísmo de cortesía*, where the clitic *le/les* is used for courtesy address to the second person, that is, where the referent of the object is not the third

person (NGLE 2009: §16.8d). In these cases, the dative clitic agrees with second person pronouns *usted/ustedes*, as in *Yo le conozco a usted* ('I know **you**'). In our research, it was important to tag these examples manually so as to exclude them from the analysis.

## 2.2. Previous research

Apart from numerous studies on written data (s. Fernández-Ordóñez 1993), there are various studies looking into *leísmo* in Spain (Klein-Andreu 1993; Díaz Montesinos 2017), and Latin America (DeMello 2002), which analyse speech samples or are corpus-based. Due to the difficulties in collecting and working with spoken data, those studies tend to be limited to a specific city or region. For example, during three months of fieldwork, Klein-Andreu (1993) collected 150 hours of conversation in five Castilian provinces (Valladolid, Burgos, Soria, Logroño, and Toledo). This became somewhat easier with the development of language resources and technologies, but oral corpora are scarce to this day and tend to be specialized and collected specifically for a certain investigation or project. In this sense, for his study of *leísmo* in Málaga, Díaz Montesinos (2017) used a corpus of 72 semi-structured interviews and DeMello (2002) a corpus of approximately 1.3 million words of educated speech samples from ten Latin American capitals.

One of the important questions is also how different types of this phenomenon can be classified and whether all of them could even be considered *leísmo* (Fernández-Ordóñez 1999; DeMello 2002). As mentioned before, among them are the verbs that interest us: manipulation verbs that can alternate between dative and accusative clitics.

In his study of four of these verbs (*enseñar* 'to teach', *ayudar* 'to help', *obligar* 'to make someone do something', and *invitar* 'to invite'), Rivas (2013: 156) explains that the question of *leísmo* with respect to manipulation verbs has not received a lot of attention in literature. Previous research seems to show that most of the Peninsular Spanish prefers dative clitics with these verbs, while in Andalusia, Canary Islands, and Latin America the accusative clitic is more common (Roegiest 1990: 243;<sup>109</sup> Fernández-Ordóñez 1999: 1326). However, Díaz Montesinos (2017) concludes for Málaga, a city in Andalusia, that with some of these verbs (*obligar* 'to make someone do something',

---

<sup>109</sup> Roegiest (1990) analysed written data from twentieth-century literary works in the Peninsular Spanish.

*animar* ‘to encourage’, *ayudar* ‘to help’ and *enseñar* ‘to teach’) dative is more frequent. In NGLE (2009: §16.9o) it is noted that these verbs show a considerable dialectal variation in Spain, the Antilles, Mexico, and a part of Central America, while for the rest of the areas the accusative form is more frequent. On the other hand, Rivas (2013: 162), who used *Corpus del español*<sup>110</sup> to extract approximately 5 million words of spoken Spanish, observed that the Peninsular Spanish prefers the original dative marker, while the Latin American Spanish more often opts for the accusative. However, his (Rivas 2013: 162) research indicates that this can vary from verb to verb, since in present-day Spanish verbs *obligar* ‘to make someone do something’ and *invitar* ‘to invite’ prefer the accusative form, while *enseñar* ‘to teach’ and *ayudar* ‘to help’ favour the presence of the dative clitic. In our analysis, we will be returning to these previous studies to compare the results for each verb in more detail.

To the best of our knowledge, there are no studies of *leísmo* on a large reference corpus.<sup>111</sup> Even though the corpus we used, *Corpus del Español del Siglo XXI – CORPES XXI*,<sup>112</sup> contains mostly written data, it is comprised of all different kinds of texts that are representative of the usage of Spanish language. Our study is based on the version 0.94 of the corpus which was published in July 2021 and contains more than 327 000 documents and 350 million words.<sup>113</sup> As already noted, the *CORPES XXI* covers mostly written registers (90%), but it also includes oral data (10%). The written registers in the corpus are categorized as books (40%), press (40%), Internet sources (7,5%) and miscellaneous sources (2,5%). When it comes to geographical distribution, 70% of the texts are from Latin America and 30% were produced in Spain. The material from Latin America is classified according to different linguistic regions (see Table 4). The version which was used in the study

<sup>110</sup> The structure of the corpus, as well as access to its search engine, can be found here: <https://www.corpusdelespanol.org>.

<sup>111</sup> García-Miguel (2015) uses the *ARTHUS* corpus in his research. This corpus contains almost 1,5 million words and, while it does contain texts of different types and origin, it is quite small. More information about this corpus can be found here: <https://www.bds.usc.es/corpus.html>.

<sup>112</sup> The detailed structure of the corpus, as well as access to its search engine, can be found here: <https://www.rae.es/banco-de-datos/corpes-xxi>.

<sup>113</sup> Since then, two updated versions of the corpus have been published. The current version (the version 1.0) is available from June 2023 and it represents an increase of more than 45 million forms compared to the version that was used in this study.

covers the period from the beginning of this century up to the year 2020. Furthermore, it is worth mentioning that all the texts in the corpus are lemmatized and morphosyntactically annotated. For these reasons, we believe the *CORPES XXI* to be the state-of-the-art resource to examine the current status of clitic alternation within this group of verbs.

### 3. Data and methodology

The possible verb candidates for the study were first gathered from Fernández-Ordóñez's (1999) two interrelated sections about clitic reinterpretation with different sets of verbs (1999: 1325–1328, 1330–1335).<sup>114</sup> In these sections, the author deals with a number of suasive verbs and verbs of influence, as well as other verbs that show a change in government. She (1999: 1330) explains that all of these verbs governed the dative case in Medieval Spanish and in the present-day Spanish they are in the process of replacing it with the accusative. In this way 32 verbs were selected and further analysed in order to see whether they matched the criteria for inclusion in the group of manipulation verbs.<sup>115</sup>

For a verb to be considered a manipulation verb, we relied on Givón's criteria (2001: 151–152). Semantically, these verbs typically include a person who manipulates (or tries to manipulate) the behaviour of another human, the so-called manipulee. The agent of the complement clause should be coreferential with the manipulee from the main clause and the complement clause expresses the target event which should be performed by the manipulee.

Syntactically, the manipulee of the main clause can be either the direct or indirect object and it is also the coreferential-subject of the complement clause, where it is left unexpressed. The verbs we have chosen all follow the same pattern: *verb + object + preposition + clause* (which can appear in a

<sup>114</sup> It is worth noting that the author also mentions many verbs for which alternation causes a change in structure and/or meaning. She calls these the cases of "apparent *leísmo*". Such verbs were omitted from the analysis.

<sup>115</sup> As mentioned in the previous section, morphosyntactic variation in the complementation of four of these verbs (i.e. *enseñar*, *ayudar*, *obligar*, and *invitar*) has already received substantial interest in Rivas (2013). However, our study is based on a different set of data and more attention is paid to dialectological differences.

nominalized or a non-finite form, as shown below). Moreover, it is crucial to highlight the importance of the preposition in this structure, since it seems that its presence favours the reinterpretation of the clitic, i.e. those manipulation verbs which require a preposition tend to show clitic variation (Fernández-Ordóñez 1999: 1333).

Because of the nature of these verbs, the object expressing the manipulee is [+animate] and [+human], while the prepositional object expressing the target event to be performed by the manipulee is necessarily inanimate. This and the fact that in these constructions the animated object pronominalized in dative can be transformed into the subject of a passive clause (usually reserved only for direct objects) could explain why the original dative is being substituted for the accusative (Fernández-Ordóñez 1999: 1333–1334). The passivization of the sentence (10a) (where the object *a Juan* is doubled by the dative clitic *le*) can be found in (10b).

- (10) a. *A Juan le invitaron a la conferencia.*  
 DAT John CL.3SG invite.PST.3PL to DET.F.SG conference(F).SG  
 ‘They invited John [him] to the conference.’
- b. *Juan fue invitado a la conferencia.*  
 John be.PST.3SG invite.PTCP to DET.F.SG conference(F).SG  
 ‘John was invited to the conference.’

Each of the 32 verbs were examined semantically and syntactically and, in the end, only 9 verbs (*animar* ‘to encourage’, *autorizar* ‘to authorize’, *convencer* ‘to convince’, *forzar* ‘to force’, *incitar* ‘to incite’, *invitar* ‘to invite’, *obligar* ‘to make someone do something’, *ayudar* ‘to help’, and *enseñar* ‘to teach’) were included in the study. For example, the verb *animar* ‘to encourage’ was evaluated on the basis of the example (11a) where Juan is the manipulee and María is the one who tries to manipulate his behaviour. In addition, Juan is the object of the main verb *animar* ‘to encourage’, but at the same time it is the subject of the complement clause, i.e. of the verb *estudiar* ‘to study’. The complement clause clearly represents the target action that should be done by Juan.

- (11) a. *María animó a Juan a estudiar más.*  
 ‘Mary encouraged **John** to study harder.’
- b. *María lo / le animó a estudiar más.*  
 ‘Mary encouraged **him** to study harder.’

It is also important to clarify that in Spanish the syncretic form *a* (as in *a Juan* in (11a)) can introduce both direct objects differentially and indirect objects obligatorily. In other words, when not presented as clitics, all indirect objects are necessarily introduced by *a*, while only direct objects that are [+human], [+animate] and [+definite] are marked by differential object marking, which is homonymous of the dative marker *a* (Mayer 2017: 5). In consequence, the change in the government of the manipulation verbs selected for our study can only be appreciated when their objects are expressed through 3<sup>rd</sup> person clitic pronouns which are marked for case (*lo* for the accusative and *le* for the dative, as in (11b)).

Furthermore, the prepositional object is also introduced by the preposition *a* (except in the case of *convencer + de* ‘to convince’) and it can present itself as a finite noun clause<sup>116</sup> (as in (12)), non-finite noun clause in the infinitive form (as in (13)) or a noun phrase (as in (14)).

- (12) *María animó a Juan a que*  
 Mary encourage.PST.3SG DAT/ACC John to that  
*cantara una canción.*  
 sing.SBJV.3SG INDF.F.SG song(F).SG  
 ‘Mary encouraged John to sing a song.’
- (13) *María obligó a Juan a venir a su fiesta.*  
 Mary make.PST.3SG DAT/ACC John to come.INF to POSS.3SG party.FSG  
 ‘Mary made John come to her party.’
- (14) *María invitó a Juan a un café.*  
 Mary invite.PST.3SG DAT/ACC John to INDF.M.SG coffee(M).SG  
 ‘Mary invited John to get a cup of coffee.’

On the other hand, some verbs were excluded from the analysis because they lacked the necessary syntactic characteristic, namely the preposition (e.g. *prohibir* ‘to forbid’ or *ordenar* ‘to order’ as in (15) and (16)), while others did not fit in the category semantically (e.g. *acusar* ‘to accuse’ or *reñir* ‘to scold’ as in (17) and (18), where the subject of the main clause does not seem to manipulate another person’s behaviour nor the complement clause could be seen as coding the target event to be performed by the manipulee).

<sup>116</sup>In Spanish grammaticography, it is customary to consider the element introducing these clauses as a preposition (e.g. NGLÉ 2009: §36.1e), even though other grammatical traditions may see it as a complementizer.

- (15) *María prohibió a Juan que estudiara más.*  
‘Mary forbade John to study harder.’
- (16) *María ordenó a Juan estudiar más.*  
‘Mary ordered John to study harder.’
- (17) *María acusó a Juan de robar los libros.*  
‘Mary accused John of stealing the books.’
- (18) *María riñó a Juan por robar los libros.*  
‘Mary scolded John for stealing the books.’

Subsequently, we proceeded to look for the selected verbs in a reference corpus for Spanish language, *Corpus del Español del Siglo XXI – CORPES XXI*, which contains more than 350 million words and is representative of the modern 21st century Spanish language spoken in Spain and in Latin America (see section 2.2. for more details about the corpus). The data was collected in the following manner: We looked for the lemma of each verb and its required preposition in an interval 2 to the right followed by either the accusative clitic (i.e. the lemma *lo*) or the dative clitic (i.e. the lemma *le*)<sup>117</sup> in an interval 2 to the left and right.<sup>118</sup> This procedure was selected in order to obtain separate results for *verb + the dative clitic* and *verb + the accusative clitic*. Only the third person clitics (i.e. *lo, la, los, las*) were chosen due to the fact that, as we have seen previously, clitics for first and second person (both singular and plural) show syncretism for case in Spanish. All the concordances were then downloaded, as well as the data about the distribution by region and country for each verb, which is provided automatically by the corpus search engine.

In the next step, a percentage of dative and accusative examples for each verb was calculated from the total number of concordances. However, since we set out to explore different types of clitic alternation according to various above-mentioned variables that could not be searched for automatically in the corpus, a more detailed manual annotation and analysis of the examples

<sup>117</sup> Even though *lo* and *le* are a part of the same paradigm, they are in this corpus lemmatized independently. This is why we refer to them as two lemmas.

<sup>118</sup> The interval for the preposition was chosen on the basis of previous studies on prepositional complementation (Mikelenić 2020) and the preliminary analysis of a smaller set of random concordances showed that accusative and dative clitics are highly unlikely to appear further than two spaces away from the verb (either to the left or to the right).

was warranted. In other words, the annotation of each example included (apart from the case and number tags that were obtained directly from the corpus) the following categories: person (we annotated whether the referent of the clitic was a third person or the second person addressed by means of *usted* or *ustedes*, a phenomenon called *leísmo de cortesía*), gender (we distinguished between masculine and feminine forms whenever it was possible to establish this distinction based on the referent), and type of clause (we marked each complement clause as infinitival, nominalized or finite). This annotation was performed on 15% of random examples for each clitic type and for each verb.<sup>119</sup>

## 4. Results and analysis

We begin with the overall distribution of the clitic alternation with these verbs, to pass onto a general geographical distribution, where we distinguish between frequency of usage of each clitic for each verb in Spain and in Latin America as a whole. We then look into a more detailed diatopic analysis of five most frequent verbs. Lastly, we examine other manually annotated variables, namely type of clause, number, and gender.

### 4.1. Overall distribution of the accusative/dative alternation

Table 2 shows the first set of results obtained from the corpus. The verbs selected for the analysis appeared in 34,881 sentences in total and their overall distribution is shown in Table 2. The verbs are listed in descending order according to their frequency with both the accusative and the dative clitic, and each verb is accompanied by the corresponding preposition. The last two columns display the percentage of these examples that appear with the accusative and the dative clitic. This data already seems to confirm that the process of replacing the etymological dative with the accusative is not nearly finished and, what is more, these changes do not appear to affect all the verbs in the same way. Some verbs, like *ayudar* ‘to help’, *animar* ‘to encourage’ and *autorizar* ‘to authorize’, hesitate between the two possibilities,

---

<sup>119</sup> Since the concordances for each verb were not lacking in number (for verbs like *ayudar* we found almost 10.000 examples, as we will show later on), this percentage was chosen as the highest number of sentences that we could feasibly annotate manually.

i.e. they are used as frequently with the accusative and the dative. On the other hand, the rest of the verbs show a more obvious preference for one of the options, in most cases the accusative object (with the clear exception of the verb *enseñar* ‘to teach’, which favours the dative). For the verbs *obligar* ‘to make someone do something’, *invitar* ‘to invite’, *enseñar* ‘to teach’, and *ayudar* ‘to help’, the results are in line with what Rivas (2013) concluded: The first two verbs occur more often with the accusative, while the last two prefer the dative clitic, even though with *ayudar* ‘to help’ both options are equally frequent. However, as we have mentioned in the theoretical introduction to our paper, the status of these verbs cannot be accounted for without registering their diatopic differences, which are presented in the following section.

Table 2. Overall distribution of the accusative/dative alternation with selected manipulation verbs

verb + preposition	total number of examples	with accusative <i>lo</i>	with dative <i>le</i>
<i>ayudar a</i> ‘to help’	9,943	45.65%	54.35%
<i>obligar a</i> ‘to make someone do something’	9,363	69.34%	30.66%
<i>invitar a</i> ‘to invite’	6,906	76.74%	23.26%
<i>enseñar a</i> ‘to teach’	3,458	14.11%	85.89%
<i>convencer de</i> ‘to convince’	2,653	80.63%	19.37%
<i>animar a</i> ‘to encourage’	1,135	56.30%	43.70%
<i>forzar a</i> ‘to force’	668	72.75%	27.25%
<i>incitar a</i> ‘to incite’	388	67.78%	32.22%
<i>autorizar a</i> ‘to authorize’	367	58.04%	41.96%

## 4.2. Geographical distribution of the accusative/dative alternation

In Table 3, the geographical distribution of the variation dative/accusative is presented. The table displays the normalized frequencies for each verb with lemmas *lo* or *le* in Spain and in Latin America. In order to first present a more general view of the issue, we have not divided the frequencies for Latin America by different regions; instead, they are displayed as ranges. When presented in this way, the frequencies show that most of the examined verbs prefer the dative construction (*le*) in Spain (with the exception of *convencer* ‘to convince’), while Latin America usually opts for the accusative (*lo*). This is in accordance with the overview we find in other studies (s. section 2.2.) that also confirm that Spain, as a whole, favours the older conservative option, that is, the presence of the dative clitic, whereas Latin America disfavours its use.<sup>120</sup>

Table 3. Geographical distribution of the accusative/dative alternation with selected manipulation verbs in Spain and Latin America

verb + preposition	Spain		Latin America	
	norm. freq. <sup>121</sup> accusative <i>lo</i>	norm. freq. dative <i>le</i>	norm. freq. accusative <i>lo</i>	norm. freq. dative <i>le</i>
<i>ayudar a</i> ‘to help’	8.45	16.68	13.99–18.60	5.57–20.48
<i>obligar a</i> ‘to make someone do something’	10.07	16.3	20.52–25.45	2.43–6.31
<i>invitar a</i> ‘to invite’	7.81	8.38	15.78–23.41	1.53–3.64
<i>enseñar a</i> ‘to teach’	1.67	6.43	0.55–2.39	7.33–10.92
<i>convencer de</i> ‘to convince’	4.31	3.35	6.09–8.76	0.36–0.71
<i>animar a</i> ‘to encourage’	2.21	3.41	0.88–2.63	0.32–0.59
<i>forzar a</i> ‘to force’	0.92	1.14	1.17–1.99	0.15–0.2
<i>incitar a</i> ‘to incite’	0.49	0.83	0.59–1.08	0.12–0.14
<i>autorizar a</i> ‘to authorize’	0.16	0.6	0.56–1.45	0.32–0.48

<sup>120</sup> For an analysis by areas, see the next section.

<sup>121</sup> Normalized frequency per 1 million words.

This is most obvious with verbs such as *obligar* ‘to make someone do something’ and *invitar* ‘to invite’, while the strength of the preference for one option over the other decreases with verbs at the end of the table, namely *animar* ‘to encourage’, *convencer* ‘to force’, *incitar* ‘to incite’, and *autorizar* ‘to authorize’. However, we must consider that the last four verbs are generally not as frequent in the corpus, which prompted us to focus on the first verbs in more detail. Another reason for looking at the first five verbs in this table more closely is the fact that some of the data observed here diverges from the general tendencies, as we can see with verbs like *ayudar* ‘to help’, *enseñar* ‘to teach’, and *convencer* ‘to convince’. In addition, some of the frequency ranges for Latin America are wider, which shows us there is a discrepancy between different areas, and only a closer look into the data can help clarify the extent of regional variation.

### 4.3. Analysis of the most frequent verbs by areas

When it comes to the verb *ayudar* ‘to help’ (Table 4), four regions of Latin America exhibit a preference for the accusative. This preference is especially strong in the Río de la Plata region (namely, Argentina). However, it is interesting to note that the continental Caribbean region (i.e. Colombia and Venezuela), as well as Mexico and Central America prefer the dative, which is a more conservative option that is typically characteristic of Spain.<sup>122</sup> This is especially important because previous research shows that Caracas seems to opt for the accusative in these constructions (Aleza Izquierdo 2009: 266–268), and the data on Mexico is mixed. Fernández-Ordóñez (1999: 1331) notes that with this verb Mexico and Central America prefer the dative, while Aleza Izquierdo (2009: 266–268) reports accusative for Mexico City and DeMello’s (2002: 264) data for Mexico City displays no preference between the two clitics.<sup>123</sup> DeMello (2002: 269) goes even further and claims that there is a difference in meaning between *ayudarlo* and *ayudarle*, where *le* signifies a more active participant than *lo*. Rivas (2013: 162–163) uses this distinction to explain the overall preference for dative with the verb *ayudar* ‘to help’

<sup>122</sup> As it was mentioned in section 2.2., a difference exists regarding dative/accusative alternation with these verbs between different Spanish regions. However, since the corpus we used does not provide the data to allow for that kind of distribution, we will not make any assumptions about this in our analysis.

<sup>123</sup> It should be noted that DeMello (2002) does not look into *ayudar* ‘to help’ only with the prepositional object, but in all occurrences. In addition, his corpus is smaller, and he recorded a total of eight examples of this verb, four with *lo* and four with *le*.

in these constructions, saying that «dative marking is preferred in order to show that the referent of the object has an active participation in the activity expressed by the matrix verb». For our data, however, this explanation is incomplete, as it should give more importance to the differences between the varieties of Spanish language.

Table 4. Geographical distribution by areas of the accusative/dative alternation with *ayudar + a* 'to help' and *enseñar + a* 'to teach'

Ayudar a 'to help'			Enseñar a 'to teach'		
Area <sup>124</sup>	norm. freq. accusative <i>lo</i>	norm. freq. dative <i>le</i>	Area	norm. freq. accusative <i>lo</i>	norm. freq. dative <i>le</i>
Río de la Plata	18.6	5.57	Continental Caribbean	1.82	10.92
The Antilles	17.23	11.28	Chilean	0.69	10.17
Continental Caribbean	14.67	19.77	Río de la Plata	0.6	10.16
Andean	14.41	11.55	Mexico & Central America	1.46	9.42
Chilean	14.33	10.12	Andean	0.55	8.28
Mexico & Central America	13.99	20.48	The Antilles	2.39	7.33
Spain	8.45	16.68	Spain	1.67	6.43

The verb *enseñar* 'to teach' (Table 4) prefers the dative both in Spain and in Latin America. This preference is registered in the literature, but the data so far was inconclusive regarding its distribution. Even those countries like Colombia, Uruguay, or Argentina, for which some authors claimed they exhibit a higher degree of alternation (s. Fernández-Ordóñez 1999: 1331),

<sup>124</sup> Countries included in each area: Andean: Bolivia, Ecuador, Peru; the Antilles: Cuba, Puerto Rico, Dominican Republic; Continental Caribbean: Colombia, Venezuela; Chilean: Chile; Mexico & Central America: Costa Rica, El Salvador, Guatemala, Mexico, Nicaragua, Panama; Río de la Plata: Argentina, Paraguay, Uruguay. The corpus also includes data from the United States of America, the Philippines, and Equatorial Guinea, but we chose not to include it in this distribution because of an extremely small number of documents and, by extension, examples for each of these varieties.

seem to prefer the dative. In other words, our data shows that the use of *le* is a general tendency with *enseñar* ‘to teach’ in present-day Spanish. This is in line with Rivas (2013: 162), who also reports a preference for dative in these constructions.

When it comes to its distribution in Latin America, the verb *invitar* ‘to invite’ (Table 5) displays a strong preference for the accusative that is in line with claims made by Fernández-Ordóñez (1999: 1326) and Rivas (2013: 162). Nevertheless, this preference is not always corroborated in previous research, especially in the case of the Antilles, Mexico, and Central America, that, according to NGLE (2009: §16.9o), should show a higher degree of alternation between the two forms. The data for Spain, however, does corroborate exactly this, even though the dative is slightly more frequent.

For the verb *obligar* ‘to make someone do something’ (Table 5), all the regions of Latin America favour the accusative, while in Spain the dative is more frequent. However, similarly to *invitar* ‘to invite’, previous research on the subject seems to indicate that Spain, the Antilles, Mexico, and Central America hesitate more between two options (Fernández-Ordóñez 1999: 1326; NGLE 2009: §16.9o). The same as for the verb *invitar* ‘to invite’, Rivas (2013) concluded that *obligar* ‘to make someone do something’ generally prefers the accusative clitic. Comparing the verbs *obligar* ‘to make someone do something’ and *invitar* ‘to invite’ with *ayudar* ‘to help’, and *enseñar* ‘to teach’, he (Rivas 2013: 163) claims that *obligar* ‘to make someone do something’ and *invitar* ‘to invite’ «present their objects as passive entities that do not exert any control on the activity expressed by the matrix verb» and, because of this, disfavour the dative form. Again, this explanation has to be expanded to include the dialectal situation, especially in Spain, which Rivas (2013: 170) agrees with.

Table 5. Geographical distribution by areas of the accusative/dative alternation with *invitar + a*, *obligar + a* and *convencer + de*

Invitar a 'to invite'			Obligar a 'to make someone do something'			Convencer de 'to convince'		
Area	norm. freq. accusative <i>lo</i>	norm. freq. dative <i>le</i>	Area	norm. freq. accusative <i>lo</i>	norm. freq. dative <i>le</i>	Area	norm. freq. accusative <i>lo</i>	norm. freq. dative <i>le</i>
Chilean	23.41	1.98	Río de la Plata	25.45	2.68	Río de la Plata	8.76	0.71
Continental Caribbean	22.28	2.21	Chilean	25.00	2.43	Mexico & Central America	7.35	0.36
Río de la Plata	19.51	1.53	Mexico & Central America	23.87	4.19	Chilean	7.09	0.29
Mexico & Central America	19.09	3.07	Continental Caribbean	23.84	2.9	Continental Caribbean	6.85	0.36
The Antilles	16.12	2.75	Andean	20.73	6.31	The Antilles	6.13	0.35
Andean	15.78	3.64	The Antilles	20.52	5.24	Andean	6.09	0.48
Spain	7.81	8.38	Spain	10.07	16.3	Spain	4.31	3.35

The verb *convencer* 'to convince' (Table 5), like *invitar* 'to invite', shows a high degree of vacillation between the accusative and the dative in Spain, even though the accusative clitic seems to be somewhat more frequent. Of all the verbs that we investigated, *convencer* 'to convince' is the only one that exhibits a slight preference for the accusative in Spain. In Latin America, the general tendency is observed, and the accusative is the preferred option in all the areas, even though previous research on the subject seems to indicate that, again, the Antilles, Mexico, and Central America hesitate more between the two options (NGLE 2009: §16.9o).

Summarizing, even though the verb *ayudar* 'to help' in these constructions is slightly more frequent with the dative (Table 2), when we look into the geographical distribution, the image changes. The verb *ayudar* 'to help' occurs more often with the accusative marking in the areas of Río de

la Plata, the Antilles, Andean, and Chilean, while dative is preferred in the Continental Caribbean, Mexico and Central America, and Spain. On the other hand, the verb *enseñar* ‘to teach’ shows a clear preference for the dative both in Spain and in Latin America. The verbs *invitar* ‘to invite’ and *obligar* ‘to make someone do something’ prefer accusative in Latin America and dative in Spain, even though both forms are present in Spain, especially in the case of *invitar* ‘to invite’ that exhibits a high degree of vacillation between the two. The same can be said for *convencer* ‘to convince’, that is, surprisingly, more frequent with the accusative in Spain. The whole of Latin America prefers the accusative marking with this verb as well.

Our analysis shows that we have to be careful with any general conclusions and that it is crucial to take into account the dialectological differences between Spanish speaking countries (cf. García-Miguel 2015). Furthermore, and in line with Rivas (2013), it has been corroborated that these verbs in this construction differ between themselves, so any conclusions about them as a whole are likely to be incomplete.

Finally, the group of verbs analysed in this study points to a continuous dynamic relationship between the dative and accusative marking, because in these constructions dative is the one being replaced with the accusative. This goes against diachronic evidence gathered by Company Company (2001: 37) and her claim that «whenever a Dat is syntactically in a position to compete with an Acc, the Dat wins».

#### 4.4. Manual analysis of other variables

The second part of our research was the manual analysis of 15% of all examples by marking them for different variables. Each example was marked for five different variables, four of which (case, number, gender, and type of clause) are displayed in Table 6. The fifth variable was person, in order to exclude from the analysis the examples of previously mentioned *leísmo de cortesía*, that amounted to 7% of the total number of dative examples marked manually.

The example (19) contains a clitic *las* that is accusative, plural, and feminine and the verb of its subordinate clause is in the infinitive form, whereas the example (20) has a clitic *le* that is dative, singular, and masculine and the verb of its subordinate clause is in the finite form. Gender with dative clitics was marked by reading through each segment to find a referent and if one was not found, that was labelled by an X. It was also possible that the

referents were of both genders (marked as MF), as in (21).<sup>125</sup> A third type of clause was actually not a clause but a noun phrase (marked as NOM), as in (22).

Table 6. Variables marked manually and examples

Variables		Example
CASE	ACC	(19) <i>enseñarlas a colgarse de la barra</i> 'teach <b>them</b> to hang from the bar'
	DAT	(20) <i>...me costaría convencerle de que suba...</i> '...I would have a hard time convincing <b>him</b> to go up...'
NUMBER	SG	(20) <i>...me costaría convencerle de que suba...</i> '...I would have a hard time convincing <b>him</b> to go up...'
	PL	(19) <i>enseñarlas a colgarse de la barra</i> 'teach <b>them</b> to hang from the bar'
GENDER	M	(20) <i>El obispo es un buen amigo mío, me costaría convencerle de que suba hasta el monasterio.</i> ' <u>The bishop</u> is a good friend of mine, I would have a hard time convincing <b>him</b> to go up to the monastery.'
	F	(19) <i>enseñarlas a colgarse de la barra</i> 'teach <b>them</b> to hang from the bar'
	MF	(21) <i>poner límites a niñas y niños no sólo les ayuda a tener claro...</i> 'Setting limits for <u>girls and boys</u> not only helps <b>them</b> to be sure...'
TYPE OF CLAUSE	INF	(19) <i>enseñarlas a colgarse de la barra</i> 'teach them <b>to hang</b> from the bar'
	FIN	(20) <i>...me costaría convencerle de que suba...</i> '...I would have a hard time convincing him <b>to go</b> up...'
	NOM	(22) <i>lo obligó a eso el calor</i> 'the heat made him <b>do that</b> '

<sup>125</sup> It should be mentioned that in Spanish masculine plural forms can be used generically to refer to both genders, so, for example *los niños* can include girls as well as boys. In those cases, the example was marked as MF. However, due to the lack of wider context, this was often difficult to ascertain, leading to marking most of these masculine plurals as M, observing their morphological form. Having in mind that at least during the last two decades and the gender equality awareness the generic use of masculine has been dissuaded in favour of visibility of both genders, it would be interesting to find out the current situation in language use.

Looking into the type of complement clause, our analysis showed that Spanish manipulation verbs tend to come with the infinitive form, which is a conclusion that was already corroborated by other authors (Fernández-Ordóñez 1999: 1325–1326; Rivas 2013: 163). Finite clauses and the nominalized forms seem to be similar in regard to frequency in these constructions (Table 7).

Table 7. Results for manually marked variables: number, gender and type of clause

CLITIC TYPE: ACC				CLITIC TYPE: DAT			
2595				1798			
NUMBER	SG	1943	75%	NUMBER	SG	1207	67%
	PL	652	25%		PL	591	33%
GENDER	M	1615	62%	GENDER	M	1237	69%
	F	949	37%		F	270	15%
	MF	30	1%		MF	42	2%
	X	1	0%		X	249	14%
TYPE OF CLAUSE	INF	2065	80%	TYPE OF CLAUSE	INF	1630	91%
	FIN	259	10%		FIN	96	5%
	NOM	271	10%		NOM	72	4%

As for the variable of gender, as stated before, due to the nature of the manipulation verbs, the object referent in all our examples was a person. According to the literature (Fernández-Ordóñez 1999: 1319; NGLÉ 2009: §16.8c), we expected a higher number of masculine in comparison to feminine dative clitics, given that masculine *leísmo* tends to be significantly more frequent than feminine. This is corroborated in our data (Table 7), even though for 14% of dative examples, the gender of the referent could not be discerned from the text. Nevertheless, they were equally distributed throughout the corpus and were simply a result of insufficient available context. Therefore, there is no reason to believe that the gender distribution within that group would not be the same as for the rest of the dative examples. To investigate the relationship between the type of clitic and gender, we applied a chi-square test, taking into account only the examples where gender was clear from the context. The test has revealed that the two variables were related ( $\chi^2=164.994$ ,  $df=1$ ,  $p<0.0001$ ,  $\phi=0.201$ ), that the type of clitic is affected by the gender of the clitic's referent. In other words, the gender of the referent has a role in the choice of clitic, even though the

correlation is weak. We can conclude that the type of clitic can to a small degree be predicted by the gender, more specifically female gender, given that the observed frequency for female accusative clitics was higher than expected, contrary to the observed frequency for female dative clitics which was lower than expected (Table 8).

Table 8. Observed and expected frequencies for variables of gender and number across the type of clitic

	ACC	DAT
<b>M_exp</b>	1796.249	1055.751
<b>M_obs</b>	1615	1237
<b>F_exp</b>	767.7514	451.2486
<b>F_obs</b>	949	270
<b>SG_exp</b>	1860.744	1289.256
<b>SG_obs</b>	1943	1207
<b>PL_exp</b>	734.2556	508.7444
<b>PL_obs</b>	652	591

In regard to the variable of number, we expected singular *leísmo* to be more frequent than plural, which is also confirmed by our data (Table 7). In order to investigate a possible connection of the type of clitic and number, we have conducted a chi-square test here as well. The result ( $\chi^2=31.398$ ,  $df=1$ ,  $p<0.0001$ ,  $\varphi=0.084$ ) again showed a statistically significant correlation between the two variables, but with a very weak effect size. The same as with gender, the number of the referent influences the choice of clitic, albeit a very weak influence. The difference is that there was slightly more plural *leísmo* than expected and less singular (Table 8).

Therefore, in addition to concluding that in these constructions with manipulation verbs masculine *leísmo* is more frequent than feminine, as well as singular *leísmo* in comparison to plural, (which is in accordance with previously stated order of types of *leísmo* by frequency, see section 2.1.), we can also ascertain the correlation between the type of clitic and grammatical features of its referent, namely gender and number. However, the effect size is weak for the first and very weak for the second variable.

## 5. Conclusions

To sum up, it can be noted that our analysis confirms that manipulation verbs which govern a prepositional object exhibit alternation between accusative and dative clitics in modern Spanish. When it comes to the geographical distribution, generally speaking, the Peninsular Spanish favours the dative clitic, whereas the Latin American Spanish prefers the presence of an accusative clitic. In other words, the use of clitics is more conservative in the Peninsular Spanish than in the Latin American Spanish. However, it is crucial to add that these tendencies are not distributed equally across Spanish-speaking regions and vary from verb to verb.

This also brings us closer to understanding the grammatical relations of two core participants in Spanish: the direct and indirect object. We are inclined to agree with García-Miguel (2015) that there are no clear boundaries between them, and we believe that the results of our study, where the clitic case is subject to dialectal factors and there is even considerable variation across verbs in this relatively homogenous group, point to that.

As for the analysis of other variables, the most frequent type of clause introduced by the preposition tends to have the verb in the infinitive form. We can also conclude that the most frequent type of *leísmo* in our research is the one for human masculine singular object, that is to say, these constructions with the manipulation verbs follow the general tendency of distribution of types of *leísmo*. Furthermore, the gender and number of the referent seem to weakly affect the choice of clitic.

Future research should establish whether the tendencies we have observed in our analysis can be found in all Spanish varieties and account for more subtle differences between them. For example, according to some studies (s. section 2.2.), we could expect that, for those verbs that present a higher degree of vacillation in our data, some varieties of the Peninsular Spanish do favour one option over the other, but this could not be confirmed in our research since our corpus does not differentiate between various regions in Spain.

## Appendix

### Acknowledgments

We are grateful for the valuable comments given by two anonymous reviewers and the editors of this volume.

### Abbreviations

1	first person
2	second person
3	third person
acc	accusative
cl	clitic
dat	dative
det	determiner
exp	expected
f	feminine
fin	finite
indf	indefinite
inf	infinitive
m	masculine
MF	masculine/feminine
nom	nominative
obs	observed
pl	plural
pst	past
ptcp	participle
sg	singular
X	unknown referent

### Acronyms

NGLE *Nueva gramática de la lengua española*

## References

- Alarcos Llorach, E. (1999). *Gramática de la lengua española*. Madrid: Espasa Calpe.
- Aleza Izquierdo, M. (2009). El pronombre personal complemento: los clíticos. In: Hernández, C. (ed) *Estudios lingüísticos del español hablado en América. Volume 3.2*, 223–351. Madrid: Visor.
- Company Company, C. (2001). Multiple dative-marking grammaticalization: Spanish as a special kind of primary object language. *Studies in Language*, 25(1), 1–47.
- Company Company, C. (2003). Transitivity and grammaticalization of object. The struggle of direct and indirect object in Spanish. In: Fiorentino, G. (ed) *Romance objects. Transitivity in Romance languages*, 217–260. Berlin/New York: Mouton de Gruyter.
- DeMello, G. (2002). “Leísmo” in Contemporary Spanish American Educated Speech. *Linguistics*, 40(2), 261–283.
- Díaz Montesinos, F. (2017). Leísmo real y leísmo aparente. Sobre el uso de los pronombres átonos (le, les, la, las, lo, los) en Málaga. *Verba: Anuario galego de filoloxía*, 44, 409–480.
- Fernández-Ordóñez, I. (1993). Leísmo, laísmo y loísmo: estado de la cuestión. In: Fernández Soriano (ed) *Los pronombres átonos*, 63–96. Madrid: Taurus.
- Fernández-Ordóñez, I. (1999). Leísmo, laísmo y loísmo. In: Demonte Barreto, V. & Bosque, I. (eds) *Gramática descriptiva de la lengua española. Volume 1*, 1317–1398. Madrid: Espasa Calpe.
- Fernández Ramírez, S. (1964). Un proceso lingüístico en marcha. In: *Presente y futuro de la lengua española, Actas de la Asamblea de Filología del I Congreso de Instituciones Hispánicas. II*, 277–285. Madrid: Ediciones Cultura Hispánica.
- Fernández Ramírez, S. (1987). *Gramática española. 3.2. El pronombre*. Madrid: Arco Libros.
- García-Miguel, J. M. (2015). Variable coding and object alignment in Spanish: A corpus-based approach. *Folia Lingüística*, 49(1), 205–256.
- García González, F. (1981). /le (lu), la, lo (lu)/ en el Centro-Norte de la Península. *Verba*, 8, 347–535.
- Givón, T. (2001). *Syntax: an introduction. Volume I*. Amsterdam/Philadelphia: John Benjamins.
- Halpern, A. L. (2017). Clitics. In: Spencer, A. & Zwicky, A. (eds) *The Handbook of Morphology*, 101–122. Oxford: John Wiley & Sons.
- Klein-Andreu, F. (1993). Distintos sistemas de empleo de “le”, “la”, “lo” – perspectiva sincrónica, diacrónica y sociolingüística. In: Fernández Soriano, O. (ed) *Los pronombres átonos*, 337–354. Madrid: Taurus.
- Lapesa, R. (1968). Sobre los orígenes y evolución del leísmo, laísmo y loísmo. In: Baldinger, K. (ed) *Festschrift Walther von Wartburg*, 523–551. Tübingen: Max Niemeyer.
- Mayer, E. (2017). *Spanish clitics on the move: Variation in time and space*. Berlin, Boston: De Gruyter Mouton.
- Mikelenić, B. (2020). *Kontrastivna korpusna analiza prijedložne dopune u španjolskome i njezinih ekvivalenata u hrvatskome*. PhD thesis, University of Zagreb.
- Real Academia Española y Asociación de Academias de la Lengua Española (2009). *Nueva gramática de la lengua española*. Madrid: Espasa Calpe.
- Rivas, J. (2013). A Usage-Based Analysis of Morphosyntactic Variation in the Complementation of Manipulative Verbs. *Revista Internacional de Lingüística Iberoamericana*, 11(1/21), 155–171.
- Roegiest, E. (1990). La tipología sintáctica del objeto transitivo en español. *Verba*, 17, 239–248.
- Spencer, A. & Luís, A. R. (2012). *Clitics. An Introduction*. Cambridge University Press.
- Vinja, V. (1998). *Gramatika španjolskog jezika*. Zagreb: Školska knjiga.

# 9. Tests for distinguishing arguments and adjuncts and their application to Croatian locative phrases

MATEA BIRTIĆ, IVANA BRAČ

Keywords: *arguments; adjuncts; Croatian locative phrases; verbs of putting; tests for arguments and adjuncts*

In this paper, we have described the application of tests for distinguishing arguments and adjuncts to two types of Croatian locative phrases. The first type of phrase appears as a third complement with three-place verbs of putting and it is usually defined as a directional location or a goal. The second type of locative phrases appears with two-place verbs with a second complement bearing the semantic role of either a (stative) location or a (directional) location, i.e. goal. We have applied the following tests to two types of phrases: omission test, implication test, *do so* test, *this happened* test, replacement test, substitution test, paraphrase with a dependent clause, *wh*-island test, dialogue test, and iterativity test (Ágel 2000; Helbig 1992; Needham & Toivonen 2011; Przepiórkowski 2016a; Šojat 2008). We have concluded that the *wh*-island test and the iterativity test do not give substantial evidence for these types of phrases. The remaining eight tests are useful in Croatian with certain limitations: all examined phrases are confirmed as arguments. All locative phrases behave uniformly as arguments according to the implication test and as adjuncts according to the replacement test.

*Original scientific paper*

## 1. Introduction

In this paper, we will examine the application of tests for distinguishing arguments and adjuncts to two classes of locative phrases (locative phrases which appear with two-place and three-place verbs). We roughly call all examined phrases *locative phrases*, although we are aware that some express a location whereas other express a goal and despite the fact that all phrases are prepositional. Although the difference between arguments and adjuncts is tacitly present in most grammatical descriptions, their distinction is

not always clear (Needham & Toivonen 2011: 1).<sup>126</sup> Therefore, in several approaches, different procedures have been suggested for distinguishing between arguments and adjuncts, but none of them is universally applicable. For example, some tests that differentiate arguments from adjuncts in English (word order test, preposition stranding test, pseudocleft test or the VP preposing test) are not applicable to Croatian at all. On the other hand, some tests give very different results when applied to different languages.<sup>127</sup>

In this paper, the selection of tests will be applied to the locative phrases: the omission test, the implication test, the *do so* test, the *this happened* test, the replacement test, the substitution test, the paraphrase with a dependent clause, *wh*-island test, the dialogue test, and the iterativity test. The omission test, the implication test, the replacement test, and the paraphrase with a dependent clause originate in German dependency framework (Helbig 1992). The *do so* test and *this happened* test are used in both dependency and generative grammar (Lakoff & Ross 1967; Ackema 2015; Brown & Miller 1991; Helbig 1992). The *wh*-island test originates from generative grammar, and the dialogue test is designed in functional generative description (Sgall, Hajičová & Panevová 1986). By using tests from different theoretical approaches, this research is not aligned with only one specific theoretical approach, but instead, in general follows dependency grammars, but also incorporates some insights from generative grammar. Functional generative description from which we took the dialogue test is also a variety of dependency grammar. We think that the employment of two different theoretical approaches is justified by our aim of finding which tests function well for Croatian. Establishing the tests for distinguishing arguments and

---

<sup>126</sup>Needham & Toivonen (2011), and Toivonen (2016, 2021) research extensively the difference between arguments and adjuncts and tests for distinguishing them. Their work is mostly written in the theoretical framework of lexical functional grammar, but their conclusions are important in general.

<sup>127</sup>For example, the coordination of *wh*-words which refer to arguments and adjuncts behaves differently in English and Croatian. While in Croatian it is possible to adjoin the *wh*-word which refers to an adjunct with any *wh*-word which refers to an argument, in English adjunct *wh*-words can be conjoined only with an argument *wh*-word which refers to an optional argument. Although the test can be applied to Croatian, the results do not point to the argument vs. adjunct status of some phrases (Gračanin Yuksek 2007).

adjuncts, could, in our view, precede analysing some data in a certain theoretical framework.<sup>128</sup>

The results of the application of ten tests to the locative phrases were evaluated based on the authors' judgements and results were also checked with fifteen L1 speakers of Croatian.<sup>129</sup> The example sentences were constructed by the authors. However, their construction was inspired by corpus examples. Ten verbs divided into two syntactic classes (bivalent and trivalent verbs) were chosen to check the applicability of tests to Croatian data. Since for both syntactic groups of verbs it has been more or less stated that their examined complement is an argument, our goal was to see if this would be confirmed by standard syntactic tests (already applied in other languages), see for example Samardžija (1986), Šojat (2008), Palić (2011). Also, these two groups of verbs were selected since we wanted to see if the tests for distinguishing arguments and adjuncts function differently for bivalent and trivalent verbs.

After the introduction (section 1) and description of two classes of verbs whose complements will be examined (section 2), the tests for distinguishing arguments and adjuncts used for these two groups of locative phrases will be briefly presented, as well as their application to locative phrases with two groups of verbs (section 3). Through the application of these tests, their disadvantages and advantages will be emphasized. The results will be discussed in section 4 and the conclusion will be laid out in section 5.

## 2. Two syntactic classes of verbs with locative phrases

Levin (1993) defines verbs of putting as verbs that refer to putting some entity at some location, the putting can be done in some manner, or it can refer to moving something in some direction. The first type of phrases that are examined here appears with the verbs of putting. These verbs are three-place predicates whose meaning includes a participant (agent) who throws or puts some object (patient/theme) at some location (goal). Their

---

<sup>128</sup> See Buttinelli (*this volume*, Chapter 11) for a view of the distinction between arguments and adjuncts conceived as a continuum in the context of Cognitive Grammar.

<sup>129</sup> We have checked our judgements by an informal questionnaire that was given to the fifteen colleagues in the Institute of Croatian Language who are all L1 speakers of Croatian.

third argument is often characterized as a directional location (goal) and its status will be examined here. Complements<sup>130</sup> of the following verbs are investigated: 1. *baciti* ‘throw, cast’, 2. *složiti* ‘compose, put together, arrange, set’, 3. *staviti* ‘put, set, place, lay’, 4. *nastaniti* ‘inhabit, settle, colonize’, and 5. *trpati* ‘put, load, stuff’. All of them are non-stative verbs with an agentive subject. In (1–5) there are sentences with verbs of putting with the third (examined) part underlined:

- (1) *Brat je bacio kamen u vodu.*<sup>131</sup>  
 brother.NOM.SG AUX throw.PTCP.ACT.M.SG stone.ACC.SG into water.ACC.SG  
 ‘The brother threw a stone into the water.’
- (2) *Kuhar je složio janjetinu i krumpir na tanjur.*  
 chef.NOM.SG AUX arrange.PTCP.ACT.M.SG lamb.ACC.SG and potato.ACC.SG  
 on plate.ACC.SG  
 ‘The chef arranged the lamb and potatoes on a plate.’
- (3) *Jan je stavio lonac na štednjak.*  
 Jan.NOM.SG AUX put.PTCP.ACT.M.SG pot.ACC.SG on stove.ACC.SG  
 ‘Jan put the pot on the stove.’
- (4) *Država je nastanila ljude duž granice.*  
 state.NOM.SG AUX settle.PTCP.ACT.F.SG people.ACC.PL along border.GEN.SG  
 ‘The state has settled the people along the border.’
- (5) *Ana trpa papirić u džep.*  
 Ana.NOM.SG puts little paper.ACC.SG into pocket.ACC.SG  
 ‘Ana puts the paper into the pocket.’

<sup>130</sup> Although the term *complement* is sometimes interchangeable with the term *argument*, we use the term *complement* as a cover term for both arguments and adjuncts, so for any dependent of the verb (that is for both *kruh* ‘bread’ and *škola* ‘school’ in the sentence *Brat jede kruh u školi* ‘Brother eats bread in the school’). Various authors assign these terms to either syntax or semantics. For example, Koenig, Mauner & Bienvenue (2003: 68) suggest that a *complement* and a *modifier* belong to syntax and *argument* and *adjunct* to semantics. On the contrary, Ackema (2015: 246) suggests that a *modifier* is a semantic notion and an *adjunct* a syntactic notion. Ernst (2002: 7) considers an *adjunct* to be both a syntactic and semantic notion, and an argument only a syntactic notion (exactly opposite of Koenig, Mauner & Bienvenue 2003).

<sup>131</sup> We provide glosses for examples whose structure deviates from the structure of the English translation. Beside this, considering that we apply different tests to the same examples, we provide glosses for the first appearance of an example. Glosses follow the Leipzig Glossing Rules (Bickel *et al.* 2015).

A third complement with the Croatian verbs of putting is always a prepositional phrase if not an adverb (*Ana trpa papirić tamo* ‘Ana puts the paper there’) and it refers to a goal of an action. In English, some verbs of this class show locative alternation, and to a lesser extent in Croatian too (s. Zovko Dinković 2003). According to Zovko Dinković (2003), the prefixed counterparts of the verb *trpati* ‘put, load, stuff’, which is part of our investigation, *natrpati* (perfective) and *natrpavati* (imperfective), show the locative alternation (*Natrpao je džep papirićima* ‘He loaded the pocket with small papers’).

The second type of locative phrases appears with verbs *boraviti* ‘stay’, *dosegnuti* ‘reach’, *dospjeti* ‘arrive, get to, reach’, *nastaniti se* ‘settle in/down’, *ležati* ‘lie, be’. All five verbs are bivalent. Two of them are stative verbs (*boraviti* ‘stay’, and *ležati* ‘lie, be’), while verbs *dosegnuti* ‘reach’, *dospjeti* ‘arrive, get to, reach’, and *nastaniti se* ‘settle in/down’ are non-statives. Therefore, the second argument of the verbs *boraviti* ‘stay’ and *ležati* ‘lie, be’ is (stative) location, while the second argument of verbs *dosegnuti* ‘reach’, *dospjeti* ‘arrive, get to, reach’ and *nastaniti se* ‘settle in/down’ is directional location, i.e. goal. Examples are given in (6–10):

- (6) *Moj                    rođak                    boravi    u    Chicagu.*  
 my.M.NOM.SG    cousin.NOM.SG    stays    in    Chicago.LOC.SG  
 ‘My cousin is staying in Chicago.’
- (7) *Brat                    je    dosegnuo                    do najviše                    police.*  
 brother.NOM.SG    AUX    reach.PTCP.ACT.M.SG    to highest.F.GEN.SG    shelf.GEN.SG  
 ‘The brother reached the highest shelf.’
- (8) *Egipatske                    starine                    dospjele                    su    u    Italiju.*  
 Egyptian.F.NOM.PL    antiquities.NOM.PL    arrived                    AUX    in    Italy.ACC.SG  
 ‘Egyptian antiquities have arrived in Italy.’
- (9) *Župnik                    se    privremeno    nastanio                    u    Novim Selima*  
 pastor.NOM.SG    REFL    temporarily    settle.PTCP.ACT.M.SG    in    Nova(PL).LOC    Sela(PL).LOC  
*u                    privatnoj                    kući.*  
 in                    private.LOC.SG    house.NOM.SG  
 ‘The pastor temporarily settled in Nova Sela in a private house.’
- (10) *Dinara                    leži    na granici                    između Hrvatske                    i    Bosne*  
 Dinara.NOM.SG    lies    on border.LOC.SG    between Croatia.GEN.SG    and Bosnia.GEN.SG  
*i                    Hercegovine.*  
 and                    Herzegovina.GEN.SG  
 ‘Dinara lies on the border between Croatia and Bosnia and Herzegovina.’

Except for the verb *dosegnuti* ‘reach’ (7), which can also be transitive (*Dosegnuo je najvišu policu* ‘He reached the highest shelf’), all the remaining verbs of this class are intransitive in Croatian. On the other hand, the verbs with the first types of locative phrases (1-5) are transitive. Both types of prepositional phrases are very often cited in the analysis of dependency grammars as examples of phrases that were not recognized as arguments in traditional grammars. In dependency grammars, these types of complements are primarily placed in the special class of complements<sup>132</sup> (Engel 2009, Palić 2011). Palić (2011) calls the second complement of verbs *boraviti* ‘stay’, *ležati* ‘lie, be’, *dosegnuti* ‘reach’, *dospjeti* ‘arrive, get to, reach’, *nastaniti se* ‘settle in/down’ and the third complement of the verb *nastaniti* ‘inhabit, settle, colonize’ the (obligatory) *adverbial of place* (*obvezna priložna oznaka mjesta*). Belaj & Tanacković Faletar (2017: 233), within cognitive grammar, call them *obligatory adverbial complement of place* (*obvezna adverbijalna dopuna mjesta*). Palić (2011) does not include the third complement of verbs *baciti* ‘throw, cast’, *složiti* ‘compose, put together, arrange, set’, *trpati* ‘put, load, stuff’ and *staviti* ‘put, set, place, lay’ in his exhaustive list of adverbial complements, supposedly not considering them to be adverbial complements.

Engel (2009: 145–146), for example, would classify the prepositional phrase in (6) and (10) under the situational complement (*Situativergänzung*), while complements with other verbs presumably under directional complement (*Direktivergänzung*). In literature written in English, these complements are sometimes considered to have the semantic role of location. Some authors distinguish between semantic roles of location and spatial goal (Levin & Rapoport Hovav 2005) which is expressed in dependency grammar with placement of complements in different classes. For example, it is clear that verbs *boraviti* ‘stay’ and *ležati* ‘lie, be’ have prepositional complements whose semantic role is location, while verbs of putting, and verbs *dosegnuti* ‘reach’ and *dospjeti* ‘arrive, get to, reach’ have a prepositional complement with the semantic role of goal. However, it is not easy to decide whether the complement of the verb *nastaniti se* ‘settle in/down’ bears a semantic role of

<sup>132</sup>In dependency grammars, it is common to make a list of classes of complements that are found in the individual language (nominative complements (subject complements), genitive complements, accusative complements, prepositional complements, etc.). See such a list in Samardžija (1986) and Šojat (2008) for Croatian and Mrazović & Vukadinović (2009) for Serbian.

a location or of a goal. Although the verb *nastaniti se* ‘settle in/down’, in our opinion, describes a dynamic event, the prepositional phrase with this verb is marked with the locative case, which is very often linked to the semantic role of location.<sup>133</sup>

### 3. Tests for distinguishing arguments and adjuncts and their application to the prepositional locative phrases

#### 3.1. Omission test<sup>134</sup>

The omission test, also called the optionality test (Schütze 1995; Needham & Toivonen 2011), is a standard test for the separation of obligatory elements from non-obligatory elements in a sentence. If a syntactic phrase can be omitted and the sentence remains grammatical, the omitted syntactic phrase is either a non-obligatory argument or an adjunct. If the sentence becomes ungrammatical after a phrase is omitted, this syntactic phrase is an obligatory argument. While adjuncts can be always omitted, some arguments can be also omitted with certain verbs. The examples of dispensable arguments are, for example, objects of verbs *drink* and *eat* (*John likes to drink (tea)*). Williams (2015: 68) formulates that as follows: «So if removing a dependent results in *unacceptability*, this is good initial evidence that the omitted dependent is an argument. (...) If a dependent can acceptably be omitted, it need not be an adjunct...» So, the omission test does not say which elements in a sentence are arguments and which are adjuncts, but actually extracts obligatory arguments. The situation is even more complex if we assume that there are also obligatory adjuncts, as some authors do (Grimshaw & Vikner 1993; Goldberg & Ackerman 2001).<sup>135</sup> In dependency grammar, the omission test is applied to (simple) sentences to reach the so-called ‘syntactic minimum’.

<sup>133</sup> One of the reviewers questioned the statement that this phrase’s semantic role is goal.

<sup>134</sup> In German the omission test is called *Eliminierungstest* (Helbig 1992), *Weglassprobe*, *Reduktionstest* (Helbig & Schenkel 1969). Ágel (2000) also mentions all three terms.

<sup>135</sup> Grimshaw & Vikner (1993) mostly mention the obligatory adjuncts found in passives of accomplishment verbs (*The best tomatoes are grown \*(by organic farmers)*), but Goldberg & Ackerman (2001) also list some other examples. One of the reviewers asked how these authors justify the concept of obligatory adjuncts and how then they can make a difference between arguments and adjuncts. In fact, they do not explain their view of the difference between arguments and adjuncts.

The idea behind this procedure was that all elements which are not in the scope of the predicate's valency can be omitted from a sentence (Šojat 2008: 22).

The omission test reveals that only the third argument of verbs *baciti* 'throw' (11) and *složiti* 'arrange' (12) can be omitted. According to the omission test, they are either optional arguments or adjuncts. The arguments of all other verbs cannot be omitted, which indicates that investigated complements of the verbs *staviti* 'put, set, place, lay' (13), *nastaniti* 'inhabit, settle, colonize' (14), *trpati* 'put, load, stuff' (15), *boraviti* 'stay' (16), *dosegnuti* 'reach' (17), *dospjeti* 'arrive, get to, reach' (18), *nastaniti se* 'settle in/down' (19), and *ležati* 'lie, be' (20) are obligatory arguments.

- (11) *Brat je bacio kamen.*  
'The brother threw a stone.'
- (12) *Kuhar je složio janjetinu i krumpir.*  
'The chef arranged the lamb and potatoes.'
- (13) \**Jan je stavio lonac.*  
'\*Jan put the pot.'
- (14) \**Država je nastanila ljude.*  
'\*The state has settled people.'
- (15) \**Ana trpa papirić.*  
'\*Ana puts a paper.'
- (16) \**Moj rođak boravi.*  
'\*My cousin is staying.'
- (17) \**Brat je dosegnuo.*  
'\*The brother reached.'
- (18) \**Egipatske starine dospjele su.*  
'\*Egyptian antiquities have reached.'
- (19) \**Župnik se privremeno nastanio.*  
'\*The pastor temporarily settled.'
- (20) \**Dinara leži.*  
'\*Dinara lies.'

### 3.2. Implication test

The implication test<sup>136</sup> examines if the verb presupposes or implies an entity in its state or action. Therefore, with this test, the presence of semantic arguments is examined, not their actual syntactic realization. In dependency grammars (Schumacher *et al.* 1981: 145; Engel 2009), this test is only applied if the examined syntactic phrase can be omitted by the omission test. If a syntactic part cannot be omitted without the sentence becoming ungrammatical, the test is not applied. However, we have applied this test to all locative syntactic phrases.<sup>137</sup> If a syntactic element phrase can be omitted from the sentence, then one must see if there is a relation of implication between the sentence with omitted part and the sentence which contains it. So, if, for example, from the sentence *Somebody eats* always follows *Somebody eats something*, we are dealing with an implication relation.<sup>138</sup> An additional testing of an implication is the procedure in which the examined argument has to be denied and the sentence must bring a contradiction: *He ate, but he ate nothing*.

Helbig (1992: 87) points out that the implication test is similar to the dialogue test, but need not be equated with it. While for the dialogue test the important distinction is whether something is known to a speaker or not, this distinction is not relevant for the implication test. Helbig also criticized this test by saying that the sentence *Someone eats* also implies that someone eats *somewhere* and *sometimes*, in the same way as it implies that something is eaten.

Needham & Toivonen (2011: 3) mention the core participant test, which is also similar to the implication test. With this test, two events are compared, for example, the eating and sleeping event in order to see how many participants are necessary for the realization of an event; for sleeping one, and for eating two. Needham & Toivonen (2011: 3) further note: «the

---

<sup>136</sup>In the German literature, this test is called either *Implikationstest* (Schumacher *et al.* 1981; Helbig 1992) or *Folgerungstest* (Schumacher *et al.* 2004; Engel 2009).

<sup>137</sup>One of the reviewers asked how we have justified not strictly following the procedure of German dependency grammar. As we said at the beginning, our approach is the combination of two (even three) approaches whose aim is to see which tests are most suitable for the Croatian data. So, we do not follow any of the approaches strictly.

<sup>138</sup>Researchers working in German dependency grammar (for example Schumacher *et al.* 1981) often stress that the implication test must be performed in pair with the substitution test which examines verb specificity.

intuition of whether a participant is conceptually necessary is very basic, but is not always useful.» With that, they have in mind the English verbs such as *saddle*, which involve three entities but only have two participants.

According to the results of the application of the implication test, all third or second complements of examined verbs are arguments (21–31).

- (21) \**Brat je bacio kamen, ali nije ga bacio nikamo.*  
 ‘\*The brother threw a stone, but he threw it nowhere.’
- (22) \**Kuhar je složio janjetinu i krumpir, ali nije ih složio nikamo.*  
 ‘\*The chef arranged the lamb and potatoes, but he arranged them nowhere.’
- (23) \**Jan je stavio lonac, ali nije ga stavio nikamo.*  
 ‘\*Jan put the pot, but he put the pot nowhere.’
- (24) \**Država je nastanila ljude, ali nije ih nastanila nikamo.*  
 ‘\*The state has settled people, but it has settled them nowhere.’
- (25) \**Ana trpa papirić, ali ga ne trpa nigdje/nikamo.*  
 ‘\*Ana puts a paper, but she puts it nowhere.’
- (26) \**Moj rođak boravi, ali ne boravi nigdje.*  
 ‘\*My cousin is staying, but he is staying nowhere.’
- (27) \**Brat je dosegnuo, ali nije dosegnuo ni do kuda.*  
 ‘The brother reached, but he reached nowhere.’
- (28) \**Egipatske su starine dospjele, ali nisu dospjele nikamo.*  
 ‘\*Egyptian antiquities have reached, but they have reached nowhere.’
- (29) \**Župnik se privremeno nastanio, ali nije se nastanio nigdje.*  
 ‘\*The pastor temporarily settled, but he settled nowhere.’
- (30) \**Dinara leži, ali ne leži nigdje.*  
 ‘\*Dinara lies, but it lies nowhere.’

### 3.3. *Do so* test

The *do so* test also known as VP anaphora test (Needham & Toivonen 2011) is used for a structural (syntactic) distinction between verbal arguments and adjuncts. According to this test, *do so* may substitute for a verb phrase containing a non-stative verb and its arguments (31), while it cannot replace

only a verb if the verb is transitive (32) (e.g. Lakoff & Ross 1976; Ackema 2015: 260).<sup>139</sup> Consider (31) and (32) from Ackema (2015: 260):

(31) *John ate a banana yesterday, while Geraldine did so today.*

(32) \**John ate a banana, while Geraldine did so an apple.*

According to the test, elements that occur after *do so* are outside the verb phrase and they are adjuncts. Therefore, direct objects and indirect objects, directional adverbs and affected locatives are arguments, while time, place, manner, duration, frequency, instruments, and comitative are adjuncts. When applied to Croatian data, the verb phrases *učiniti/postupiti isto* ‘to do/act the same (so)’ are used. The first application of the *do so* test to Croatian was in Brač (2018), where *učiniti isto* was suggested as a correlate of *do so*. Although Lakoff & Ross (1967) point out that the test is not suitable for stative verbs, we have observed that the test is only applicable to agentive verbs (i.e. it cannot be applied to dynamic verbs without an agent). Therefore, we have not applied the test to the verbs *boraviti* ‘stay’, *ležati* ‘lie, be’, and *dospjeti* ‘arrive, get to, reach’.

(33) *Brat je bacio kamen u vodu,*  
 brother.NOM.SG AUX throw.PTCP.ACT.M.SG stone.ACC.SG into water.ACC.SG  
*a sestra je učinila isto u zrak.*  
 and sister AUX do.PTCP.ACT.F.SG so into air.ACC.SG  
 ‘The brother threw a stone into the water, and the sister did so into the air.’

(34) *??Kuhar je složio janjetinu i krumpir na*  
 chef.NOM.SG AUX arrange.PTCP.ACT.M.SG lamb.ACC.SG and potato.ACC.SG on  
*tanjur, a konobar je učinio isto na pladanj.*  
 plate.ACC.SG and waiter AUX do.PTCP.ACT.M.SG so on tray.ACC.SG  
 ‘??The chef arranged the lamb and potatoes on a plate, and the waiter did so on a tray.’

(35) \**Jan je stavio lonac na štednjak, a*  
 Jan.NOM.SG AUX put.PTCP.ACT.M.SG pot.ACC.SG on stove.ACC.SG and

<sup>139</sup> According to Schütze (1995), the same results can be obtained with pseudo-clefting (Vestergaard 1997): *What John did on Tuesday was meet Mary; \*What John did on the shelf was put the book.*

- Maja je učinila isto na radnu ploču.*  
 Maja AUX do.PTCP.ACT.F.SG so on worktop.ACC.SG  
 ‘\*Jan put the pot on the stove, and Maja did so on the worktop.’
- (36) \**Država je nastanila ljude u prazne kuće*  
 state.NOM.SG AUX settle.PTCP.ACT.F.SG people(PL).ACC in empty.F.ACC.PL houses.ACC.PL  
*na granici, a Crveni križ učinio je isto u hotele.*  
 on border.LOC.SG and Red Cross do.PTCP.ACT.M.SG AUX so in hotels.ACC.PL  
 ‘\*The state settled people in empty houses on the border, and the Red Cross did so in hotels.’
- (37) ??*Ana trpa papirić u džep, a Marija čini isto u torbu.*  
 Ana.NOM.SG puts paper.ACC.SG into pocket.ACC.SG and Marija does so into bag.ACC.SG  
 ‘??Ana puts a paper into a pocket, and Marija does so into a bag.’
- (38) ??*Moj brat dosegnuo je do najviše police, a ja sam učinio isto do druge police.*  
 my.M.NOM.SG brother.NOM.SG reach.PTCP.ACT.M.SG AUX to highest.F.GEN.SG shelf.GEN.SG and I AUX do.PTCP.ACT.M.SG so to second.F.GEN.SG shelf.GEN.SG  
 ‘??My brother was reaching for the highest shelf, and I did so for the second shelf.’
- (39) ??*Župnik se privremeno nastanio u Novim Selima, a bivši župnik učinio je isto u susjednom selu.*  
 pastor.NOM.SG REFL temporarily settle.PTCP.ACT.M.SG in Nova(PL).LOC Sela(PL).LOC and former.M.NOM.SG pastor.NOM.SG do.PTCP.ACT.M.SG aux so in next.N.LOC.SG village.LOC.SG  
 ‘??The pastor temporarily settled in Nova Sela, and the former pastor did so in the next village.’

If we look at these results, it seems that only the verb *baciti* ‘throw, cast’ (33) can be replaced with the *učiniti isto* phrase and its third complement can be repeated in the second clause, which would indicate that the third complement is an adjunct. The *do so* replacement of the verb *dosegnuti* and its second complement is also marginally acceptable.

There are several problems with this test.<sup>140</sup> First, as it was already said the *do so* test can only be applied to non-stative verbs, moreover to agentive verbs.

<sup>140</sup> Many authors (e.g. Miller 1992; Przepiórkowski 1999) claim that the *do so* test cannot be used as a test for a structural distinction between arguments and adjuncts. On the other hand, Haspelmath (2014) mentions a variant of *do so* test as the only test which can be, perhaps, applied cross-linguistically.

Second, the Croatian verbs *činiti* or *postupati* are not equivalent to the English verb *to do* that is both the main verb and auxiliary. *Činiti* and *postupati* are the main verbs in Croatian with their own argument structure which is different for each of those two verbs. *Činiti* is a transitive verb with an accusative and dative complement whereas the verb *postupati* and its perfective form *postupiti* are intransitive verbs that can have a prepositional complement. Therefore, if the distributional properties of an examined verb match the distributional properties of the pro-verb (*činiti* and *postupati*), the acceptance of the complement with the pro-verb in the next clause will be higher. If the distributional properties of an examined verb and pro-verb do not match (*pružati* ‘give’ and *postupati* ‘do’), the sentence is unacceptable. So, if we use the verb *činiti* when testing the complement status of the dative noun phrase in the sentence in (40a) the result is that the dative noun phrase *odraslima* ‘adults’ is an argument, but if the replacement is made with the verb *postupati*, it turns out that the same dative complement is an adjunct (40b) (examples taken from Gračanin-Yukseš 2021). Obviously, both cannot be true.<sup>141</sup>

- (40) a. *Djeca pružaju utjehu odraslima, a odrasli to*  
 children.NOM give comfort.ACC.SG adults.DAT.PL and adults it  
*čine djeci.*  
 do.PRS.3PL children.DAT  
 ‘Children give comfort to adults, and adults do so to children.’
- b. \**Djeca pružaju utjehu odraslima, a odrasli tako*  
 children.NOM give.PRS.3PL comfort.ACC.SG adults.DAT.PL and adults so  
*postupaju djeci.*  
 do.PRS.3PL children.DAT  
 ‘Children give comfort to adults, and adults do so to children.’

<sup>141</sup> Miller (1992: 96; Przepiórkowski 1999: 290) noticed that: «the relevant factor for acceptability of a PP complement after *do so*, *do it*, *do that*, is not whether or not the corresponding complement of the antecedent verbs is within the VP [V'] of the antecedent, but whether or not the PP complement is acceptable as a complement for main verb *do* with a thematic role compatible with that which corresponding complement of the antecedent verb has with respect to the antecedent verb.»

### 3.4. *This happened* test

In difference to the *do so* test, which is related to the syntactic structure, *this happened* test is related to the semantic structure, i.e. it reveals if a (prepositional) phrase bears a semantic role (Hoffmann 2007). According to *this happened* test, if a sentence can be paraphrased by two sentences, one contains the “nuclear predication”<sup>142</sup> and other the adverbial, i.e. adjunct (41) (Brown & Miller 1991: 90):

- (41) a. *John stood on the table. This happened in the bathroom.*  
 b. *\*John stood. This happened on the table.*

The prepositional phrases with the verbs analysed in this paper are arguments according to this test (42–49). The native speakers of Croatian (that we have tested) have only marginally accepted the placement of the third argument of the verb *dosegnuti* ‘reach’ (47) in the second sentence. This test can be applied only to non-stative verbs too, but unlike *do so* test, the verbs to which the test applies do not need to be agentive. Therefore, we have not applied the test to the verbs *boraviti* ‘stay’ and *ležati* ‘lie, be’.

- (42) *\*Brat je bacio kamen. To se dogodilo u vodu.*  
 brother.NOM.SG AUX throw.PTCP.ACT.M.SG stone.ACC.SG this REFL  
 happen.PTCP.ACT.N.SG into water.ACC.SG  
 ‘\*The brother threw a stone. This happened into the water.’

- (43) *\*Kuhar je složio janjetinu i krumpir. To se dogodilo na tanjur.*  
 chef.NOM.SG AUX arrange.PTCP.ACT.M.SG lamb.ACC.SG and potato.ACC.SG that  
 REFL happen.PTCP.ACT.N.SG on plate.ACC.SG  
 ‘\*The chef arranged the lamb and potatoes. This happened on a plate.’

- (44) *\*Jan je stavio lonac. To se dogodilo na štednjak.*  
 Jan.NOM.SG AUX put.PTCP.ACT.M.SG pot.ACC.SG this REFL happen.PTCP.ACT.N.SG  
 on stove.ACC.SG  
 ‘\*Jan put a pot. This happened on the stove.’

<sup>142</sup>Nuclear predication comprises a verb and its arguments.

- (45) \**Država je nastanila ljude. To se dogodilo*  
state.NOM.SG AUX settle.PTCP.ACT.F.SG people.ACC.PL this REFL happen.PTCP.ACT.N.SG  
*duž granice.*  
along border.GEN.SG  
‘\*The state has settled people. This happened along the border.’
- (46) \**Ana trpa papir-ić. To se događa u džep.*  
Ana.NOM.SG puts paper-DIM.ACC.SG this REFL happens into pocket.ACC.SG  
‘Ana puts a paper. This happens into a pocket.’
- (47) ??*Brat je dosegnuo. To se događa do najviše*  
brother.NOM.SG AUX reach.PTCP.ACT.M.SG this REFL happens to highest.F.GEN.SG  
*police.*  
shelf.GEN.SG  
‘??The brother reached. This happened for the highest shelf.’
- (48) \**Egipatske su starine dospjele. To se dogodilo u*  
Egyptian.F.NOM.PL AUX antiquities reached this REFL happen.PTCP.ACT.N.SG in  
*Italiju.*  
Italy.ACC.SG  
‘\*Egyptian antiquities have reached. This happened Italy.’
- (49) \**Župnik se privremeno nastanio. To se dogodilo*  
pastor.NOM.SG REFL temporarily settle.PTCP.ACT.M.SG this REFL happen.PTCP.ACT.N.SG  
*u Novim Selima.*  
in Nova(PL).LOC Sela(PL).LOC  
‘\*The pastor temporarily settled. This happened in Nova Sela.’<sup>143</sup>

In German dependency grammars (Helbig 1992: 80), *this happened* test is collapsed with a variant of the *do so* test<sup>144</sup> and is either called *Anschlusstest* or *Geschehen* test (Ágel 2000)<sup>145</sup>, with *tun*, *machen* ‘to do’ being proforms for action verbs and *geschehen* ‘happen’ a proform for event verbs. A variant of the *this*

<sup>143</sup> One of the reviewers thinks that *this happened* test can be applied only to the verbs which have prepositional phrase with the locative case on the noun as complements. In such cases, ungrammaticality would not follow from distributional incompatibility.

<sup>144</sup> According to Helbig (1992: 80) «the A (arguments) can – in contrast to E (adjuncts) – usually be removed from the sentence and repeated in the second clause with a pro-verb (e.g. *tun* ‘do’, *machen* ‘do, work’, *geschehen* ‘happen’). («Die A können – im Unterschied zu den E – in der Regel aus dem Satz herausgelöst und in einem zweiten Satz mit einem ‘Proverb’ (z. B. *tun*, *machen*, *geschehen*) angeschlossen werden...»).

<sup>145</sup> Ágel (2000: 176) thinks that *Geschehen* test is the central test to define argumenthood. The test is established on breaking up of complete predication and shifting the argument in the next clause.

*happened* test is to separate the tested constituent and place it in a dependent clause that begins with *which happened*. Osborne (2019: 18) calls it relativization diagnostic. For example, *Brat je bacio kamen, što se dogodilo u vodu* ‘The brother threw a stone, which happened into the water’.

### 3.5. Replacement test

By the replacement test<sup>146</sup> the examined phrase will be replaced with another syntactic phrase in order to prove or disprove that the examined phrase is selected by the verb and its morphological form is determined by that verb (in this case the same preposition which requires the same case form of the noun). This test is supposed to distinguish between arguments and adjuncts and it operates on the syntactic level. According to Ágel (2000: 180), the replacement test is used to test one of the valency relations (the so-called FOSP-relation; short for formal specificity). Ágel (2000), following Jacobs (1994), thinks that valency is not one relation, but rather consists of several relations, and formal specificity is one of three ‘classical’ valency relations.<sup>147</sup>

The locative phrases, examined in this research, are adjuncts according to this test since they can be very often replaced by some other type of syntactic phrase with different morphological form (50–59).

- (50) *Brat je bacio kamen u vodu /*  
 brother.NOM.SG AUX throw.PTCP.ACT.M.SG stone into water.ACC.SG  
*na krov / preko kuće.*  
 onto roof.ACC.SG over house.GEN.SG  
 ‘The brother threw a stone into the water / onto the roof / over the house.’

<sup>146</sup> The replacement test is called *Ersatzprobe* or *Kommutationstest* in German literature (Ágel 2000: 180).

<sup>147</sup> Ágel (2000:172) says that valency consists of five relations and each valency relation corresponds to one of the tests for argument/adjunct distinction (1. obligatoriness/necessity (NOT), 2. argumenthood/logic valency (ARG), 3. formal specificity (government), FOSP, 4. semantic valency (selection), INSP, 5. subcategorization/subclass specificity (SUBKLASS). According to Ágel (2000: 172) obligatoriness (NOT); formal specificity (FOSP) and SUBKLASS (subcategorization or subclass specificity) are classic valency relations that have a grammatical, lexicographic, and didactic relevance.

- (51) *Kuhar je složio janjetinu i krumpir na tanjur / u lonac.*  
 chef.NOM.SG AUX arrange.PTCP.ACT.M.SG lamb.ACC.SG and potato.ACC.SG on plate.ACC.SG in pot.ACC.SG  
 ‘The chef put the lamb and potatoes on a plate / in a pot.’
- (52) *Jan je stavio lonac na štednjak / u ormarić / na stol / ispred djevojčice.*  
 Jan.NOM.SG AUX put.PTCP.ACT.M.SG pot.ACC.SG on stove.ACC.SG in cabinet.ACC.SG on table.ACC.SG in\_front\_of girl.GEN.SG  
 ‘Jan put the pot on the stove / in the cabinet / on the table / in front of the girl.’
- (53) *Država je nastanila ljude duž granice / na nove teritorije u napuštene kuće.*  
 state.NOM.SG AUX settle.PTCP.ACT.F.SG people.ACC.PL along border.GEN.SG on new.M.ACC.PL territories.ACC.PL in abandoned.F.ACC.PL houses.ACC.PL  
 ‘The state settled people along the border / in new territories / in abandoned houses.’
- (54) *Ana trpa papir-ić u džep / među knjige.*  
 Ana.NOM.SG puts paper-DIM.ACC.SG into pocket between book.ACC.PL  
 ‘Ana puts a piece of paper into the pocket / among the books.’
- (55) *Moj rođak boravi u Chicagu / na selu / među domorocima.*  
 my.M.NOM.SG cousin.NOM.SG stays in Chicago.LOC.SG on village.LOC.SG among native.INST.PL  
 ‘My cousin is staying in Chicago / in the countryside / among the natives.’
- (56) *Egipatske starine dospjele su u Italiju / na brod / do obale.*  
 Egyptian.F.NOM.PL antiquities reached AUX in Italy.ACC.SG on ship.ACC.SG to coast.GEN.SG  
 ‘Egyptian antiquities reached Italy / a ship / the coast.’
- (57) *Brat je dosegnuo do najviše police / najvišu policu.*  
 brother.NOM.SG AUX reach.PTCP.ACT.M.SG to highest.F.GEN.SG shelf.GEN.SG highest.F.ACC.SG shelf.ACC.SG  
 ‘The brother reached for the highest shelf / highest shelf.’
- (58) *Župnik se privremeno nastanio u Novim Selima u privatnoj kući / na osamljenom imanju.*  
 pastor.NOM.SG REFL temporarily settle.PTCP.ACT.M.SG in Nova(PL).LOC Sela(PL).LOC in private.LOC.SG house.NOM.SG on dilapidated.N.LOC.SG estate.LOC.SG  
 ‘The pastor has temporarily settled in Nova Sela in a private house / on a dilapidated estate.’

- (59) *Dinara leži na granici između Hrvatske i  
Dinara.NOM.SG lies on border.LOC.SG between Croatia.GEN.SG and  
Bosne i Hercegovine / među dvjema drugim planinama /  
Bosnia.GEN.SG and Herzegovina.GEN.SG between two.F.INST other.INST.PL mountains.INST.PL  
iznad rijeke  
above river.GEN.SG*  
'Dinara lies on the border between Croatia and Bosnia and Herzegovina / between two  
other mountains / above the river.'

Obviously, the prepositional phrases with locative or directional meanings are not closely related to a verb, i.e. they are not selected by these verbs and their form is not determined by the verb. Similar to the replacement test is the *fixed preposition test* according to which «if the verb asks for a specific preposition, the PP is an argument» (Needham & Toivonen 2011: 4).

### 3.6. Substitution test

With the substitution test, the verb next to the examined phrase is substituted by another verb. If a sentence remains grammatical after the verb is replaced by another verb, then the examined phrase is an adjunct (Ágel 2000; Šojat 2008). If a substitution yields an ungrammatical sentence, the questioned phrase is an argument. According to Ágel (2000: 187), the substitution test examines verb specificity. Verb specificity is also one of the central valency notions according to Ágel (2000: 187). Dependency grammar claims that arguments are specific for the subclass of verbs. Therefore, arguments (complements, in the dependency grammar terminology) divide all verbs from an individual language into specific classes (German *Subklassenspezifik*), whereas adjuncts do not divide the verbs into specific classes. Dependency grammars argue that arguments are specific to a class of verbs, not to each verb (Engel 2009: 130; Šojat 2008: 31; Ágel 2000: 187). Underlying the verb specificity is also the idea that adjuncts are compatible with almost all verbs. However, this assumption has been put under scrutiny lately. Among other authors, Przepiórkowski (2016a: 565) has also questioned the assumption that adjuncts can be attached to any verb mostly without restrictions, stating that «...very few adjunct types may really depend on every verb». One has to keep in mind that some authors working in the dependency grammar framework distinguish between the linguistic compatibility of verbs and accompanying phrases (linguistic knowledge, *Sprachwissen*) and the compatibility of verbs and accompanying phrases which relies on world

knowledge (*Weltwissen*, Ágel 2000). It is completely clear that not every adjunct can be attached to every verb, but the question is whether this incompatibility is based on linguistic or non-linguistic factors.<sup>148</sup>

If we apply this test to locative phrases, it seems that most verbs can be replaced only by similar verbs (from the same semantic class) when surrounded with examined phrases, which speaks for treating these phrases as arguments. It seems that the verbs *nastaniti* ‘inhabit, settle, colonize’ from the example (63) and *boraviti* ‘stay’ from the example (65) can be replaced by some verbs outside its semantic class (*Država je zatvorila ljude duž granice* ‘The state has imprisoned people along the border’, *Moj rođak jede u Chicagu* ‘My cousin eats in Chicago’) which speaks for treating the phrase *duž granice* ‘along the border’ and *u Chicagu* ‘in Chicago’ as adjuncts.

However, the situation is somewhat complicated with verbs that have three complements. If the examined complement is the last in the string, the substituted verb is usually next to the second complement, not the third. Therefore, it is not clear to which extent the acceptability of the new sentence is influenced by the compatibility of the verb and its object (second complement), and to which extent by the compatibility of the verb and the third complement.

- (60) *Brat je bacio / gurnuo / izbacio / zavitlao / \*razveselio se / \*pojeo kamen u vodu.*  
‘The brother threw / pushed / ejected / swirled / \*cheered / \*ate a stone into the water.’
- (61) *Kuhar je složio / stavio / namjestio / \*pojeo / \*sjećao se / \*izgradio janjetinu i krumpir na tanjur.*  
‘The chef arranged / put / set / \*ate / \*remembered / \*built the lamb and potatoes on the plate.’
- (62) *Jan je stavio / složio / \*pojeo / \*sjećao se lonac na štednjak.*  
‘Jan put / set / \*ate / \*remembered the pot on the stove.’
- (63) *Država je nastanila / naselila / globila / zatvorila / \*sjećala se ljude duž granice.*  
‘The state settled / inhabited / fined / imprisoned / \*remembered the people along the border.’
- (64) *Ana trpa / stavlja / \*jede / \*sjeća se papirić u džep.*  
‘Ana loads / puts / \*eats / \*remembers the paper into the pocket.’

<sup>148</sup> For example, the adverb *rado* ‘gladly’ cannot be attached to verbs that do not express voluntary activity or the adverb *uvijek* ‘always’ to verbs that express an activity that can happen only once (s. Mrazović & Vukadinović 2009: 577).

- (65) *Moj rođak boravi / jede / sjeća se u Chicagu.*  
‘My cousin is staying / eating / remembering in Chicago.’
- (66) *Brat je dosegnuo / \*jede / \*sjeća se do najviše police.*  
‘My brother reached / \*ate / \*remembered to the highest shelf.’
- (67) *Egipatske starine dospjele su / \*nalaze se / \*izgubile su se u Italiju.*  
‘Egyptian antiquities have reached / are located / \*got lost (in) Italy.’
- (68) *Župnik se privremeno nastanio / naselio / \*je pojeo / \*se sjeća u Novim Selima u privatnoj kući.*  
‘The pastor temporarily settled / inhabited / \*ate / \*remembered in Nova Sela in a private house.’
- (69) *Dinara leži / prostire se / izbija / \*zaposjeda / \*zauzima na granici između Hrvatske i Bosne i Hercegovine.*  
‘Dinara lies on / spreads / breaks out / possesses / takes the border between Croatia and Bosnia and Herzegovina.’

### 3.7. Paraphrase with dependent clause

According to dependency grammars, adjuncts can be replaced by separate sentences, as they are truly reduced sentences. The adjunct *in der Schule* ‘at school’ from the sentence (70a) can be replaced with a subordinate temporal clause (*als er in der Schule war* ‘while he was at school’) (70b), whereas an argument *in Dresden* from the sentence (71a) ‘in Dresden’ cannot (71b), (Helbig 1992: 78; Šojat 2008: 121).

- (70) a. *Er ass sein Brot in der Schule.*  
‘He ate a sandwich at school.’  
b. *Er ass sein Brot, als er in der Schule war.*  
‘He ate a sandwich while he was at school.’
- (71) a. *Mein Freund wohnte in Dresden.*  
‘My friend lived in Dresden.’  
b. *\*Mein Freund wohnte, als er in Dresden war.*  
‘\*My friend lived while he was in Dresden.’

As Helbig (1992: 78) further explains, if the test is applied to obligatory arguments, the result is an ungrammatical sentence, but if the test is applied to optional arguments, it results in sentences with different meanings, as shown in (72a) and (72b).

- (72) a. *Ušao je u tramvaj.*  
 enter.PTCP.ACT.M.SG AUX into tram.ACC.SG  
 ‘He entered the tram.’
- b. *Ušao je dok je tramvaj bio tamo.*  
 enter.PTCP.ACT.M.SG AUX while AUX tram been.M.SG there  
 ‘He entered, while the tram was there.’

Not all adjuncts can be paraphrased by the same type of adverbial dependent clauses. Subject-oriented adverbs can be paraphrased by temporal sentences with copula (72c), object-oriented adverbs can be paraphrased by independent main clauses, datives are again paraphrased in a different way. Helbig (1992: 79) cites Heringer (1970) and says that this paraphrase supports the claim that adjuncts are predicates of the rest of the sentence.

Šojat (2008) points out the shortcomings of this test by quoting Engel & Schumacher (1978) and Vater (1972) who say that not every adjunct can be paraphrased by a sentence. According to Šojat (2008: 122) and Vater (1978: 29), some arguments can be paraphrased too. Consider the following example taken from Šojat (2008: 122):

- c. *Teta Frida je preživjela rusku okupaciju.* → *Teta Frida je preživjela kada/dok su Rusi okupirali Berlin.*  
 ‘Aunt Frida survived the Russian occupation. → Aunt Frida survived when/while the Russians occupied Berlin.’

In sentences with *baciti* ‘throw, cast’ (73), *složiti* ‘compose, put together, set, arrange’ (74), and *nastaniti* ‘compose, put together, arrange, set’ (76), the paraphrase yields a sentence with different or peculiar meaning, and in all other cases, the result is an ungrammatical sentence (75, 77–82). This would suggest that the phrases next to *baciti* ‘throw, cast’, *složiti* ‘compose, put together, arrange, set’, and *nastaniti* ‘inhabit, settle, colonize’ are optional arguments, and all other phrases are obligatory arguments.

- (73) a. *Brat je bacio kamen u vodu.*  
 ‘The brother threw a stone into the water.’
- b. #*Brat je bacio kamen dok je bio u vodi.*<sup>149</sup>  
 ‘#The brother threw a stone while it was in the water.’

<sup>149</sup> One of the reviewers drew our attention to the fact that the nouns *brat* ‘brother’ and *kamen* ‘stone’ (subject and the object of the main clause) are of the same number and gender. Therefore, the phrase *dok je bio u vodi* ‘while it was in the water’ can refer to *kamen* ‘stone’ as well as to *brat* ‘brother’, but the intended meaning was the one in which *dok je bio u vodi* ‘while it was in the water’ refers to *kamen* ‘stone’.

- (74) a. *Kuhar je složio janjetinu i krumpir na tanjur.*  
 ‘The chef arranged the lamb and potatoes on a plate.’  
 b. *#Kuhar je složio janjetinu i krumpir dok su bili na tanjuru.*  
 ‘#The chef arranged the lamb and potatoes while they were on the plate.’
- (75) a. *Jan je stavio lonac na štednjak.*  
 ‘Jan put the pot on the stove.’  
 b. *\*Jan je stavio lonac dok je bio na štednjaku.*  
 ‘Jan put the pot while it was on the stove.’
- (76) a. *Država je nastanila ljude duž granice.*  
 ‘The state settled people along the border.’  
 b. *#Država je nastanila ljude kad su bili duž granice.*  
 ‘#The state settled people while they were along the border.’
- (77) a. *Ana trpa papirić u džep.*  
 ‘Ana put the paper into a pocket.’  
 b. *\*Ana trpa papirić dok je bio u džepu.*  
 ‘\*Ana put the paper while it was in a pocket.’
- (78) a. *Moj rođak boravi u Chicagu.*  
 ‘My cousin is staying in Chicago.’  
 b. *\*Moj rođak boravi dok je u Chicagu.*  
 ‘\*My cousin is staying while he is in Chicago.’
- (79) a. *Brat je dosegnuo do najviše police.*  
 ‘The brother reached to the highest shelf.’  
 b. *\*Brat je dosegnuo dok je bio do najviše police.*  
 ‘\*The brother reached while he was to the highest shelf.’
- (80) a. *Egipatske starine dospjele su u Italiju.*  
 ‘Egyptian antiquities reached Italy.’  
 b. *\*Egipatske starine dospjele su dok su bile u Italiji.*  
 ‘\*Egyptian antiquities reached while they were in Italy.’
- (81) a. *Župnik se privremeno nastanio u Novim Selima u privatnoj kući.*  
 ‘The pastor temporarily settled in Nova Sela in a private house.’  
 b. *\*Župnik se nastanio dok je bio u Novim Selima u privatnoj kući.*  
 ‘\*The pastor settled while he was in Nova Sela in a private house.’
- (82) a. *Dinara leži na granici između Hrvatske i Bosne i Hercegovine.*  
 ‘Dinara lies on the border between Croatia and Bosnia and Herzegovina.’  
 b. *\*Dok je na granici Hrvatske i Bosne i Hercegovine, Dinara leži.*  
 ‘\*Dinara lies while it is on the border between Croatia and Bosnia and Herzegovina.’

### 3.8. Extraction out of *wh*-islands

In line with observations made within generative grammar, there are specific parts of a sentence that do not allow syntactic material to be extracted or moved out of them. These parts have been called islands since Ross (1967). The islands can be strong (sequences from which nothing can be extracted) and weak (sequences from which only some elements can be extracted). Weak islands are of some importance for the argument/adjunct distinction; namely, it is commonly assumed that only arguments can be extracted from weak islands, but not adjuncts (Chomsky 1986; Miliorini 2019). Weak islands include *wh*-islands, negative islands, and factitive islands. We have tried to examine the argument/adjunct status of our locative phrases by inserting them into interrogative sentences starting with question particle *li* (presumably *wh*-islands) and then extracting them to the beginning of the sentence. There are two main problems with this potential test for distinguishing arguments and adjuncts.

Firstly, many authors doubt that the possibility of extracting some elements and the impossibility of extracting others from weak islands is connected with argument/adjunct distinction (for explanations of extraction possibilities not related to argument/adjunct distinction s. Miliorini 2019). And secondly, it is not clear if Croatian *wh*-clauses starting with the question particle *li* are *wh*-islands. Here are some examples (83), but we hesitate to use this test as a reliable tool in our further investigation since we are not sure about what counts as a weak island in Croatian.<sup>150</sup>

- (83) a. *Brat je bacio kamen u vodu.*  
 ‘The brother threw a stone into the water.’  
 b. *Marija se pita je li brat bacio kamen u vodu.*  
 ‘Marija is wondering if the brother threw a stone into the water.’  
 c. *??Kamo se Marija pita je li brat bacio kamen?*  
 ‘Where to is Marija wondering if the brother threw a stone?’  
 d. *??Gdje se Marija pita je li brat bacio kamen?*  
 ‘Where is Marija wondering if the brother threw a stone?’

<sup>150</sup> One of the reviewers of the book asked for the reason why we included the *wh*-island test in our argumentation given that it is not clear how to use it properly and it does not give any reliable results. It should be noted that we tried to apply it to Croatian, but we showed that it is not applicable in its original form to Croatian as a tool for distinguishing between arguments and adjuncts.

e. <sup>??</sup>*Kada se Marija pita je li brat bacio kamen u vodu?*

‘When is Marija wondering if the brother threw a stone into the water?’

Under (83a) there is an example sentence with the locative phrase we are examining. The example (83b) contains the examined clause included in the interrogative sentence with *li*, and in (83c) and (83d) examined phrase is moved to the beginning of the sentence in the form of *wh*-words. Both sentences with *kamo* ‘where to’ and *gdje* ‘where’ were examined since most speakers do not differentiate between *kamo* and *gdje*. The example (83e) is a sentence that serves as a checking sentence in which an adjunct is extracted. Although Croatian speakers differ to a great extent in whether they accept the sentences similar to (83c) and (83d),<sup>151</sup> it seems that the extraction of temporal adjuncts as in (83e) is worse than the extraction of locative phrases, which has been confirmed by native speakers of Croatian (83c, 83d). The distinction speaks in favour of locative phrases being arguments.

### 3.9. Dialogue test

The dialogue test was conceived within the functional generative description (Sgall, Hajičová & Panevová 1986). According to the test, we can check if some participants of an event are semantic arguments by asking a question about them. If we are dealing with an argument, it need not be present in a surface structure, but information about it is always present in the speaker’s mind.

If we have an utterance under (84a) and we ask an utterer to answer the question under (84b), their answer cannot be (84c) ‘I don’t know’ since the place where someone has arrived is a semantic argument of the verb *arrive*.<sup>152</sup>

- (84) a. *Charles arrived by car.*<sup>153</sup>  
 b. *Where did he arrive?*  
 c. # *I don’t know.*

<sup>151</sup> For many native speakers of Croatian these sentences are more or less unacceptable, but for one author of this paper and for one of the reviewers sentences are grammatical. Therefore, we marked the sentences with two question marks.

<sup>152</sup> This test implies that a verb is presented together with its argument structure in speakers’ mind.

<sup>153</sup> The dialogue is adapted from Sgall, Hajičová & Panevová (1986: 109).

The test accounts for the fact that very often some complements are not expressed but are nevertheless semantic arguments of the verb. Sgall, Hajičová & Panevová (1986: 109) state that the procedure is most successful with adverbial modification such as *arrive somewhere*, *behave in a certain way*. Helbig (1992: 86) concludes that this test is relevant to make the primary distinction between arguments and adjuncts, not the distinction between obligatory and optional arguments. Przepiórkowski (2016b) confirms that the dialogue test identifies semantically obligatory arguments which can be omitted when contextually provided, so he concludes that the test is specifically designed for syntactically optional expressions.

If we apply this test to locative phrases researched in this paper, it seems that the answer to the question about the knowledge of these arguments cannot be *I don't know*. In the case of locative phrases with the verbs *baciti* 'throw, cast' (85), *složiti* 'compose, put together, arrange, set' (86), *staviti* 'put, set, place, lay' (87), and *trpati* 'put, load, stuff' the answer *I don't know* is somewhat more acceptable than with other verbs. According to the answers, all phrases are arguments, but the third argument of the verbs of putting can stay less known to the speaker. Further research is needed to explain this fact.

- (85) a. *Brat je bacio kamen.*  
 'The brother threw a stone.'  
 b. *Kamo ga je bacio?*  
 'Where did he throw it?'  
 c. *?Ne znam.*  
 'I don't know.'
- (86) a. *Kuhar je složio janjetinu i krumpir.*  
 'The chef arranged the lamb and potatoes.'  
 b. *Kamo/Gdje ih je složio?*  
 'Where did he arrange them?'  
 c. *?Ne znam.*  
 'I don't know.'
- (87) a. *Jan je stavio lonac.*  
 'Jan put the pot.'  
 b. *Kamo ga je stavio?*  
 'Where did he put it?'  
 c. *?Ne znam.*  
 'I don't know.'

### 3.10. Iterativity test

The iterativity test is employed in many different theories – functional generative description, LFG, HPSG, and generative grammar (Przepiórkowski 2016a). It is often said that adjuncts can be iterated freely, but arguments cannot (Przepiórkowski 2016a, 2016b, Williams 2015).<sup>154</sup> However, things are not so simple – adjuncts cannot be freely added to some structures without any restrictions (89):

(88) *They came home yesterday.*

(89) \**They came home yesterday today.*

(90) *They came home yesterday around noon.*

After a closer look at examples, linguists reached a conclusion that adjuncts can be iterated only if they refer to the same entity with a different size of granularity (90) (Przepiórkowski 2016a, 2016b). Goldberg (2002: 334) has formulated that in the following way: «Adjuncts can freely be added as long as they do not imply a semantic contradiction; in particular, they must be construed to have concentric semantic scope such that one more narrowly specifies another.»

Williams (2015) has a broader look at the doubling of adjuncts. He says that although adjuncts cannot change the syntactic category, their doubling can result in the unacceptability of a sentence or different distribution of the string which is due to other reasons (phonological, semantic, and pragmatic factors). He also points out that actually no semantic relation can be doubled without coordination regardless of their argument or adjunct status (Williams 2015; Przepiórkowski 2016a) which can be connected with the long-standing observation that only one semantic role of each type may occur in a single clause.<sup>155</sup> After establishing this, it is difficult to prove that

<sup>154</sup> Williams (2015: 69): «Adjuncts may be iterated freely without any effect on syntactic well-formedness.»

<sup>155</sup> Williams's (2015: 163) explanation boils down to *Role Iteration Generalization* (RIG) which says that: «Generally or always, two distinct dependents do not bind the exact same type of semantic relation.» It means that no pair of dependents (arguments and adjuncts) can share the same relation to the same event. RIG is expressing Role Exhaustion: When a dependent is assigned a relation to some (group of) event(s), it identifies *all* and *only* the individuals in that relation to that (those) events.

there is any difference in the iterativity of the two categories. Sometimes instruments and undoubted complements can be iterated too.<sup>156</sup>

According to this, it is easy to create an iteration of examined locative phrases especially if the extension of the location is different. Sometimes we can even add the second phrase without means of coordination. Some of the examples are given in (91–96).

- (91) *Brat je bacio kamen u vodu, tik do druge obale.*  
‘The brother threw a stone into the water, right next to the shore.’
- (92) *Kuhar je složio janjetinu i krumpir na tanjur na stol.*  
‘The chef arranged the lamb and potatoes on a plate on the table.’
- (93) *Država je nastanila ljude u napuštene kuće duž granice.*  
‘The state has settled the people in abandoned houses along the border.’
- (94) *Moj rođak boravi u kući u Chicagu.*  
‘My cousin is staying in a house in Chicago.’
- (95) *Brat je dosegnuo do najviše police iznad onih knjiga.*  
‘The brother reached to the highest shelf above those books.’
- (96) *Župnik se privremeno nastanio u Novim Selima u privatnoj kući.*<sup>157</sup>  
‘The pastor temporarily settled in Nova Sela in a private house.’

With this in mind, it seems that iteration is difficult to pursue as a test for distinguishing arguments and adjuncts, and the results of this test were not considered as important for defining the argument/adjunct status.<sup>158</sup>

<sup>156</sup> Przepiórkowski (2016a: 564) takes over an example of iteration of complements from Zaenan & Crouch (2009): *I count on you, on your kindness*. In this example, presumably two complements (*you, your kindness*) are pointing to the same.

<sup>157</sup> This example has not been changed, the verb *nastaniti* ‘settle’ is accompanied by two prepositional (locative) phrases from the beginning. The same is true for the example with the verb *nastaniti* ‘settle, colonize’ too.

<sup>158</sup> In Suozzi *et al.* (*this volume*, Chapter 10) the authors use the iterativity test in a slightly different manner. In the repetition of instruments they use two similar prepositional phrases with the same preposition, but they never claim that the same preposition should be used, which was pointed out to us by one of the reviewers of the book. As that may be true for instruments, we are not convinced that the test has to be applied in the same way with other (adverbial) phrases: *\*Išao sam u školu / na fakultet*. ‘I went to school / to the university’.

## 4. Discussion

In this paper, we presented ten tests, whose general overview is given in Table 1. The table also shows how a positive or a negative test value correlates with an examined phrase being an argument or adjunct.

Table 1. A general overview of the presented tests

TEST	+	-
1. Omission test	obligatory argument	optional argument adjunct
2. Implication test	argument	adjunct
3. <i>Do so</i> test	adjunct	argument
4. <i>This happened</i> test	adjunct	argument
5. Replacement test	adjunct	argument
6. Substitution test (replacement of a verb)	adjunct	argument
7. Paraphrase with a dependent clause	adjunct	obligatory argument
8. Extraction out of <i>wh</i> -islands	object	adjunct (subject)
9. Dialogue test	adjunct	argument
10. Iterativity test	adjunct	argument

In Table 2, we show the results of the application of eight tests to two groups of complements, trivalent and bivalent verbs with the locative phrase (location or goal). At first, we chose ten tests (shown in Table 2). The criteria for choosing tests were their applicability to Croatian, and applying tests from different theoretical approaches to see if they give similar results. According to the first criterion, some of the tests were excluded immediately. For example, the word order test, preposition stranding test, pseudocleft test or the VP preposing test are not applicable to the Croatian language due to its syntactic and morphological structure (for an extensive list of tests s. Needham & Toivonen 2011; Toivonen 2016). In the course of the investigation, we also excluded the extraction of *wh*-islands and the iterativity test. The extraction of *wh*-islands test is difficult to apply to Croatian, and we were not sure if we have applied it correctly, and for the iterativity test we think that it would not give us substantial evidence. As for the second criterion, the first seven tests are intensively used in dependency grammar, but some of them, for example, omission test and replacement test are used

universally. Testing by repeating a syntactic phrase (iterativity test) is also used in all linguistic approaches. The *do so* test is also used in generative grammar. The eighth test (extraction out of *wh*-islands) is conceived of in generative grammar, and the dialogue test in functional generative description.

Table 2. An overview of the applied tests and their results

Tests	Omission test	Implication Test /Core Participant test	<i>Do so</i> test	<i>This happened</i> test	Replacement test	Substitution test	Paraphrase with dependent clause	Dialogue test	Result
Verbs	- adj - opt. arg + obl. arg	+ arg	+ adj	+ adj	+ adj	+ adj	+ adj	+ adj - arg	
<i>baciti</i>	adj opt. arg.	+ arg	adj	arg	adj	arg	opt. arg	+/- arg	optional argument
<i>nastaniti</i>	obl. arg	arg	arg	arg	adj	arg	arg	- arg	argument
<i>složiti</i>	adj opt. arg	arg	arg	arg	adj	arg	opt. arg	+/- arg	optional argument
<i>staviti</i>	obl. arg	arg	arg	arg	adj	arg	arg	+/- arg	argument
<i>trpati</i>	obl. arg	arg	arg	arg	adj	arg	arg	+/- arg	argument
<i>boraviti</i>	obl. arg	arg	inapplicable	inapplicable	adj	arg	arg	arg	argument
<i>dosegnuti</i>	obl. arg	arg	+/-	adj	adj	arg	arg	arg	argument
<i>dospjeti</i>	obl. arg	arg	inapplicable	- arg	adj	arg	arg	arg	argument
<i>ležati</i>	obl. arg	arg	inapplicable	inapplicable	adj	arg	arg	arg	argument
<i>nastaniti se</i>	obl. arg	arg	arg	arg	adj	arg	opt. arg	arg	argument

A brief view of our results reveals that all examined phrases are arguments. The complements of all verbs behave uniformly on two tests. They are all arguments according to the implication test, and adjuncts according to the replacement test. The results of the other six tests differ depending on the verb. However, in this paper, we applied a scalar approach, by which we decide on the status of a phrase based on the predominance of tests that speaks either for argument or adjunct status. The omission test differentiates between obligatory arguments on one side and optional arguments and adjuncts on the other side. Hence, by this test, we can extract only obligatory arguments. All other tests extract both obligatory and optional arguments, but some of them differentiate between these two categories. The results of the paraphrase with a dependent clause indicate which argument is obligatory and which optional (ungrammatical sentence points to obligatory

arguments, and semantically peculiar sentences to optional arguments) according to Helbig (1992). The implication test and the dialogue test are specifically designed to reveal optional arguments.

According to the omission test, complements of all verbs except for verbs *baciti* 'throw, cast' and *složiti* 'compose, put together' have obligatory arguments, i. e. only these verbs are negative on the omission test, which means that their complements are either optional arguments or adjuncts. If we strictly followed German dependency grammars, we would not need to apply the implication test to complements of all verbs, only to complements of verbs *baciti* and *složiti*. However, we have applied the test to complements of all verbs and, as expected, all locative phrases with examined verbs are arguments (including the third arguments of verbs *baciti* and *složiti* which have turned out to be optional arguments – negative on the omission test, positive on the implication test). The third complement of the verb *baciti* 'throw, cast' is also positive on the *do so test* (the only complement positive to this test among all analysed complements), which indicates that it is an adjunct (33). Also, it is difficult to decide if the complement of the verb *dosegnuti* 'reach' can pass this test. The verb *dosegnuti* 'reach' is not completely ungrammatical when embedded in *do so* replacement, so we have put +/- in the corresponding box. As said earlier, the test should not be applied to verbs *ležati* 'lie, be', *boraviti* 'stay', and *dospjeti* 'arrive, get to, reach' since the first two verbs are stative, and the third is non-agentive. There are other shortcomings of this test mentioned in section 3; nevertheless, we have applied this test to these complements. The *this happened* test is not applicable to *boraviti* 'stay, be' and *ležati* 'lie, be' either, and the verb *dosegnuti* 'reach' is positive to this test, which means that it is an adjunct according to this test. All other verbs are arguments according to this test (*baciti* 'throw, cast', *nastaniti* 'inhabit, settle, colonize', *složiti* 'arrange', *staviti* 'put, set, place, lay', *trpati* 'put, load, stuff', *nastaniti se* 'settle in/down'). Since all verbs examined in this paper accept other morphological forms instead of examined prepositional locative phrase as its second or third complement, the morphological form of a prepositional locative complement is not 'dictated' by the verb, i. e. phrases are positive on the replacement test which indicates adjunct status. The application of the substitution test confirms the argument status of all examined phrases since verbs can be mostly replaced by verbs of the same semantic class except for some uncertainty in the application of this test to the construction with the verb

*nastaniti se* ‘settle in/down’.<sup>159</sup> As it is said earlier, the substitution test is clearer with bivalent verbs since the substitution of a verb occurs next to the examined phrase. According to the paraphrase with dependent clause, complements of verbs *baciti* ‘throw, cast’, *složiti* ‘compose, put together’, and *nastaniti se* ‘settle in/down’ are optional arguments, and the complements of verbs *nastaniti* ‘inhabit, settle, colonize’, *staviti* ‘put, set, place, lay’, *trpati* ‘put, load, stuff’, *dosegnuti* ‘reach’, and *dospjeti* ‘arrive, get to, reach’ are obligatory arguments. Dialogue test filters arguments from non-arguments and the answer to the question cannot be *I don’t know*. It seems that with verbs *baciti* ‘throw, cast’, *složiti* ‘compose, put together’, *staviti* ‘put, set, place, lay’, and *trpati* ‘put, load, stuff’ the answer *I don’t know* is easier to accept than with the other verbs, which we have marked in the table by +/-arg. That is probably connected to the fact that these verbs are also bivalent in some contexts.

As already pointed out, the biggest exception among all examples is a prepositional complement of the verb *baciti* ‘throw, cast’. This complement is an optional argument according to the omission test (if we connect this test with the implication test), and according to the paraphrase with a dependent clause. The complement is an adjunct according to the *do so* test and the replacement test. The result of the dialogue test is ambiguous. Although the results of two tests point to adjunct status, still more tests point to argument status, and precisely to the status of an optional argument. The complement of the verb *složiti* also turns out to be an optional argument according to two tests (the omission test and the paraphrase with the dependent clause), so we decided to define it as an optional argument (the result of the dialogue test is also ambiguous). The complement of the verb *dosegnuti* ‘reach’ is an argument according to five tests and it is an adjunct on two tests, and as we said, it is neither negative nor positive on the *do so* test. It seems that *do so* replacement is easier to imagine with trivalent verbs with the locative complement than with bivalent verbs with the locative complement. Some complements of trivalent verbs are also ambivalent on the dialogue test which speaks for their status of optional arguments (*nastaniti se* ‘settle in/down’ is also ambivalent on the dialogue test).

<sup>159</sup> As it is pointed out in section 3.6, it is imaginable that *nastaniti* ‘inhabit, settle, colonize’ in the construction *nastaniti ljude duž granice* ‘to settle the people along the border’ can be replaced with a verb from other semantic class *zatvoriti* ‘imprison’ (*Država je zatvorila ljude duž granice* ‘The state imprisoned the people along the border’).

## 5. Conclusion

Finally, based on these tests we can conclude with certainty that the third complement of verbs *baciti* ‘throw, cast’ and *složiti* ‘compose, put together, arrange, set’ is an optional argument. For complements of other verbs, we cannot claim with certainty that they are either obligatory or optional arguments: they are obligatory arguments according to the omission test, but optional arguments according to some other tests. The second argument of bivalent verbs, cited here, is considered obligatory by linguists working both on Croatian and German (Helbig 1992; Samardžija 1986). Exactly non-optionalness of complements of verbs *boraviti* ‘stay’, *ležati* ‘lie, be’, *dosegnuti* ‘reach’, *dospjeti* ‘arrive, get to, reach’, and *nastaniti se* ‘settle down/in’ has led the linguists working in dependency grammar to add this group of phrases into arguments (Helbig 1992; Engel 1999). The research has also shown that the iterativity test and extraction out of *wh*-islands are very difficult to apply to the complements of Croatian verbs, and that *do so* test and *this happened* test can be applied to a limited number of verbs.

In the future, if we would be able to apply these tests to complements of other semantic classes of verbs (or all semantic classes of Croatian verbs), it will be even clearer which tests would be most suitable for testing argument vs. adjunct status of complements in Croatian.

## Appendix

### Acknowledgments

This research has been fully supported by the Croatian Science Foundation under the projects *Syntactic and Semantic Analysis of Arguments and Adjuncts in Croatian – SARGADA* (IP-2019-04-7896) and *Semantic-Syntactic Classification of Croatian Verbs – SEMTACTIC* (IP-2022-10-8074). The results obtained in this research are also applied in building the SARGADA repository.

We are grateful to the two reviewers of the paper as well as the reviewers of the book for numerous useful remarks and precious advice.

## Abbreviations

#	semantic anomaly
*	strongly ungrammatical
??	ungrammatical, perhaps grammatical for some speakers
?	odd
ACC	accusative
ACT	active
adj	adjunct
arg	argument
AUX	auxiliary
DAT	dative
DIM	diminutive
F	feminine
GEN	genitive
INST	instrumental
LOC	locative
M	masculine
N	neuter
obl.	obligatory
opt.	optional
PL	plural
PRS	present
PTCP	participle
REFL	reflexive
SG	singular

## References

- Ackema, P. (2015). Arguments and Adjuncts. In: Kiss, T. & Alexiadou, A. (eds) *Syntax – Theory and Analysis: An International Handbook*, 246–273. Berlin: De Gruyter Mouton.
- Ågel, V. (2000). *Valenztheorie*. Tübingen: Gunther Narr Verlag.
- Belaj, B. & Tanacković Faletar, G. (2017). *Kognitivna gramatika. Knjiga druga*. Zagreb: Disput.
- Bickel, B., Comrie, B. & Haspelmath, M. (2015). *Leipzig Glossing Rules*. Available online: <https://www.eva.mpg.de/lingua/resources/glossing-rules.php>, Accessed on 2023-12-06).
- Brown, K. & Miller, J. (1991). *Syntax: A Linguistic Introduction to Sentence Structure*. London: Routledge.
- Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.
- Engel, U. (2009). *Syntax der deutschen Gegenwartssprache. 4., völlig neu bearbeitete Auflage*. Berlin: Erich Schmidt Verlag.

- Engel, U. & Schumacher, H. (1978). *Kleines Valenzlexikon deutscher Verben*. Tübingen: Gunther Narr.
- Goldberg, A. (2002). Surface generalizations: An alternative to alternations. *Cognitive Linguistics*, 13(4), 327–356.
- Goldberg, A. & Ackerman, F. (2001). The Pragmatics of Obligatory Adjuncts. *Language*, 77(4), 798–814.
- Grimshaw, J. & Vikner, S. (1993). Obligatory Adjuncts and the Structure of Events. In: Reuland, E. & Abraham, W. (eds) *Knowledge and Language. Volume II: Lexical and Conceptual Structure*, 143–155. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Gračanin-Yukseš, M. (2007). *About sharing*. PhD thesis, Massachusetts Institute of Technology.
- Gračanin-Yukseš, M. (2021). Application of complement/adjunct tests to Croatian verb phrases. Workshop presentation, Workshop held as a part of the project Syntactic and Semantic analysis of arguments and adjuncts in Croatian on 8<sup>th</sup> November 2021.
- Haspelmath, M. (2014.) Arguments and Adjuncts as Language-Particular Syntactic Categories and as Comparative Concepts. *Linguistic Discovery*, 12(2), 3–11.
- Helbig, G. (1992). *Probleme der Valenz- und Kasuslehre*. Tübingen: Niemeyer.
- Helbig, G. & Schenkel, W. (1969). *Wörterbuch zur Valenz und Distribution deutscher Verben*. Leipzig: VEB Bibliographisches Institut.
- Heringer, H.-J. (1970). *Deutsche Syntax*. Berlin: De Gruyter Mouton
- Hoffmann, T. (2007). Complements versus Adjuncts? A Construction Grammar account of English prepositional phrases. *Occasional Papers in Language and Linguistics*, 3, 92–119.
- Jacobs, J. (1994). *Kontravalenz*. Trier: Wissenschaftlicher Verlag.
- Koenig, J.-P, Mauner, G. & Bienvenue, B. (2003). Arguments for adjuncts. *Cognition*, 89, 67–103.
- Lakoff, G. & Ross, J. R. (1976). Why you can't do so into the sink. In: McCawley, J. D. (ed) *Syntax and Semantics. Volume 7: Notes from the linguistic underground*, 101–131. New York: Academic Press.
- Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.
- Levin, B. & Rappaport Hovav, M. (2005). *Argument realization*. Cambridge: Cambridge University Press.
- Miliorini, R. (2019). Extraction from weak islands: alternatives to the argument/adjunct distinction. *ReVel, edição especial*, 17(16), 37–58.
- Needham, S. & Toivonen, I. (2011). *Derived Arguments*. (Available online: [web.stanford.edu/group/cslicpublications/cslicpublications/LFG/16/papers/lfg11needhamtoivonen.pdf](http://web.stanford.edu/group/cslicpublications/cslicpublications/LFG/16/papers/lfg11needhamtoivonen.pdf), Accessed on 2022-11-14.)
- Osborne, T. (2019). *A Dependency Grammar of English. An introduction and beyond*. Amsterdam – Philadelphia: John Benjamins.
- Palić, I. (2011). O glagolima koji vežu obvezatne adverbijalne dopune u bosanskoj jeziku. *Suvremena lingvistika*, 72, 201–217.
- Schumacher, H., Bourstin, P., Frosch, H., Ballweg-Schramm, A., Ballweg, J., Kubczak, J. (1981). *Verbvalenz. Konzeption eines Wörterbuchs deutscher Verben. Zu Theorie und Praxis einer semantisch orientierten Valenzlexikographie*. Tübingen: Gunter Narr Verlag.

- Przepiórkowski, A. (1999). *Case Assignment and the Complement/Adjunct Dichotomy. A Non-Configurational Constraint-Based Approach*. PhD thesis, University of Tübingen.
- Przepiórkowski, A. (2016a). How Not to Distinguish Arguments from Adjuncts in LFG. In: Arnold, D., Butt, M., Crysmann, B., Holloway King, T. & Muller, S. (eds) *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 560–580. Warsaw: Polish Academy of Sciences.
- Przepiórkowski, A. (2016b). Against the Argument-Adjunct Distinction in Functional Generative Description. *The Prague Bulletin of Mathematical Linguistics*, 106, 5–20.
- Reuland, E. & Abraham, W. (1993). *Knowledge and Language. Volume II: Lexical and Conceptual Structure*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Ross, J. R. (1967). *Constraints on variables in syntax*. PhD thesis, Massachusetts Institute of Technology.
- Samardžija, M. (1986). *Valentnost glagola u suvremenom hrvatskom književnom jeziku*. PhD thesis, University of Zagreb.
- Schumacher, H., Kubczak, R., Schmidt, J. & de Ruyter, V. (2004). *VALBU – Valenzwörterbuch deutscher Verben*. Tübingen: Gunter Narr Verlag.
- Schütze, C. T. (1995). PP Attachment and Argumenthood. In: Schütze, C. T., Ganger, J. & Broihier, K. (eds) *Papers on Language Processing and Acquisition. MIT Working Papers in Linguistics*, 26, 95–151. Cambridge (MA): MIT Press.
- Sgall, P., Hajičová, E. & Panevová, J. (1986). *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. Dordrecht/Boston/Lancaster/Tokyo: D. Reidel Publishing Company.
- Šojat, K. (2008). *Sintaktički i semantički opis glagolskih valencija u hrvatskom*. PhD thesis, University of Zagreb.
- Toivonen, I. (2016). Revisiting the argument-adjunct distinction. Hand-out of a talk given at Stanford University, October 27, 2016.
- Vater, H. (1978). On the Possibility of Distinguishing between Complements and Adjuncts. In: Abraham, W. (ed) *Valence, Semantic Case and Grammatical Relations*, 21–46. Amsterdam: John Benjamins.
- Verspoor, C. M. (1997). *Contextually-Dependent Lexical Semantics*. PhD thesis, University of Edinburgh.
- Vestergaard, T. (1977). *Prepositional phrases and prepositional verbs: A study in grammatical function*. The Hague: Mouton.
- Williams, A. (2015). *Arguments in Syntax and Semantics*. Cambridge: Cambridge University Press.
- Zovko Dinković, I. (2003). Lokativna alternacija u engleskom i hrvatskom. *Studia Romanica et Anglica Zagrabiensia*, 47/48, 585–596.

# 10. On the argument status of Instruments in Italian

ALICE SUOZZI, ANNA CARDINALETTI, GIANLUCA E. LEBANI

Keywords: *instrumental role; Italian; argumenthood; semantic recoverability; optionality*

This study aims at discussing the semantic and syntactic status of Instruments in Italian and to account for their syntactic optionality. Our claim is that Instruments are not adjuncts; instead, semantically they are arguments/secondary participants and, syntactically, secondary complements. Their optionality is thus to be treated as an instance of argument omission, ruled by semantic recoverability: the more an Instrument is recoverable from the verb the more it is syntactically dropped. Our proposal, inspired by Pustejovsky's (1995) classification of arguments, is supported by the analysis of three corpora of spoken Italian.

*Original scientific paper*

## 1. Introduction

Instruments have been extensively studied over the years (s. Lakoff 1968; Rappaport-Hovav & Levin 1988; Dowty 1991; Koenig *et al.* 2003; Rissman 2013; Rissman & Rawlins 2017, among others), but no consensus has been reached on their status. They have been alternately considered as adjuncts (Dowty 1982; Rappaport-Hovav & Levin 1988; Jackendoff 1990; Bresnan 1982; Rissman & Rawlins 2017), arguments (Abney 1989; Schütze 1995), or secondary complements (Speas 1990; Ono 1992; Brunson 1992; Cardinaletti 2015; 2016; 2019).

The main issue with the classification of Instruments is the mismatch between their semantic and syntactic behaviour: while Instruments are in many cases semantically required by the verb, they are always syntactically optional. In other words, if Instruments are semantic arguments but syntactic adjuncts, this «would be a case where the semantic structure of the verb does not determine syntactic argument behavior.» (Rissman 2013: 16).

In this study, we address the issue of the argument status of Instruments along with their optionality, showing that their syntactic omission – far from being a proof of them being adjuncts – is in fact to be treated as an instance of argument omission. Applying Pustejovsky's (1995) tripartition of arguments to Instruments in Italian, we claim that different types of Instruments exist, depending on their semantic recoverability and their incorporation into the verb meaning, and that these two factors rule the pattern of their production/omission. Finally, we present the results of a corpus analysis that confirms our predictions about the omission of Instruments.

The paper is organized as follows: in section (2), a semantic and syntactic definition of Instruments is provided; section (3) is devoted to a review of the main semantic (3.2.1) and syntactic (3.2.2) diagnostics for argumenthood and to their application to Instruments; our proposal is illustrated in section (4). In section (5), the corpus analysis is described. In section (6), we draw some conclusions.

## 2. Instruments in Italian

### 2.1. A semantic definition

Since the first attempts at the definition of Instruments, their causal nature was stressed. Fillmore (1968: 24) characterized Instruments as «causally involved in the action or state identified by the verb». They have successively been defined “causal intermediaries” (Talmy 1976; Jackendoff 1990; Goldberg 2002), i.e. intermediate entities in the causal chain between the Agent and the Patient (Croft 1991). In a similar vein, Talmy (2000) defined them as both the acted-upon entity of one part of the causal chain and the “cause” of another part. According to Dowty (1991), they share some features with the proto-role Agent as well as with the proto-role Patient. Brunson (1993) made an even stronger claim: Instruments cannot be described as an independent thematic role. Instead, together with Agents, Instruments are “disconstituents” of a single Cause role, i.e. they are different syntactic constituents that realize a single – broader – thematic role.

However, since at least Marantz (1984), it has been observed that Instruments do not always play a causal role in the event by which they are implied; in fact, different interpretations of Instruments (Dowty 1994) or Instrumental subroles (Marantz 1984; Ono 1992; Schütze 1995; Schlesinger

1995; Koenig *et al.* 2003, 2007) have been identified, depending on the role the Instrument plays in the event.<sup>160</sup>

In order to address this issue, we propose the following definition of Instruments which – as will be clear in section 3.2.1 – best captures the different types of Instruments, particularly the non-intermediary ones.

- (1) The Instrument is the participant which, being somehow manipulated by the Agent, contributes to the realization of the event by which it is implied.

## 2.2. The syntax of Instruments

Instruments are mapped onto syntax via two linking rules (Lakoff 1968; Koenig *et al.* 2003, 2007). The first rule links the Instrument to the complement of the preposition ‘with’, *con* in Italian, as in (2a), while the second links it to the complement of the active form of the verb ‘to use’, *usare* in Italian, as in (2b).<sup>161</sup>

- (2) a. X V Y [con [Z]]  
       ‘X V-s Y [with [Z]]’  
       b. X usa [Z] per V Y  
       ‘X uses [Z] to V Y’

Thus, a DP is syntactically realizing an Instrument if it can occur as Z in both (2a) and (2b).

The possibility for a DP to occur in both the structures, and particularly in (2b), has been used as a diagnostic for identifying instruments (Lakoff 1968; Koenig *et al.* 2003, 2007), since it rules out non-instrumental meanings of the PP headed by *con* ‘with’ (henceforth, *con*-PP), e.g. comitative or modal meanings, as in (3):

- (3) a. *Marco ha potato la siepe {con le cesoie/con precisione/con Sara}*  
       ‘Marco trimmed the hedge {with shears/with precision/with Sara}’  
       b. *Marco ha usato {le cesoie/\*la precisione/\*Sara} per potare la siepe*  
       ‘Marco used {the shears/\*the precision/\*Sara} to trim the hedge’

<sup>160</sup>The different types of Instruments are described in detail in section 3.2.1 since their existence makes a strong case for considering Instruments arguments.

<sup>161</sup>According to Lakoff (1968) and Koenig *et al.* (2003, 2007), the structure in (2b) is a paraphrase of the structure in (2a) but see Rissman (2013) and Rissman & Rawlins (2017) for the differences in meaning between the two.

Italian exhibits a third strategy for the syntactic realization of Instruments, i.e. the clitic pronoun *ci* ‘with it’, as in (4).

- (4) (*Con il coltello*<sub>i</sub>) Gianni *ci*<sub>i</sub>=*ha* tagliato il pane  
 (With the knife<sub>i</sub>) Gianni with.it<sub>i</sub>=has cut the bread  
 ‘Gianni cut the bread with it’

Our work focuses on the structures in (2a) and (4), i.e. on the instrumental PP (henceforth, inst-PP) and on the instrumental clitic (henceforth, inst-cl), for two main reasons: first, cross-linguistically the PP is a more prototypical realization of Instruments than the complement of ‘to use’;<sup>162</sup> second, the inst-PP provides a better insight on the relationship between the verb meaning and the selection of (a class of) Instruments than the verb ‘to use’ does. See the sentences in (5), which replicate Rissman & Rawlins’ (2017) observations for English:

- (5) a. *Gianni ha usato il coltello per tagliare il pane*  
 ‘Marco used the knife to cut the bread’  
 b. *Marco ha usato il coltello*  
 ‘Marco used the knife’

In (5a) the instrumental DP *il coltello* ‘the knife’ is selected by the embedded verb *tagliare* ‘to cut’, in spite of syntactically being the internal complement of the verb *usare* ‘to use’; this is further proved by the fact that (5b) is interpretable only as long as the meaning of the infinitival clause is available in the (extra-)linguistic context.

The choice of also focusing on the inst-cl naturally follows, since the antecedent of the clitic pronoun is the PP and not the *use*-complement DP.

However, following Lakoff (1968) and Koenig *et al.* (2003, 2007) we used the possibility of occurring in both (2a) and (2b) as a diagnostics to rule out non-instrumental meanings, since the clitic pronoun *ci* and the *con*-PP are polysemic. Throughout the work, we use ‘inst-PP’ and ‘inst-cl’ to refer to *con*-PPs and the clitic pronoun *ci* when they are used in their instrumental function.

<sup>162</sup>Languages realize Instruments via either adpositional or case marking (Haspelmath 2014; Van Hooste 2018); an additional strategy is Noun Incorporation (Baker 1988; Johns 2017), which is always displayed by languages that possess either of the two strategies above.

In Italian, Instruments are always optional, i.e. they can be omitted without making the sentence ungrammatical, as in (6):

- (6) a. *Luca sta giocando con la palla*  
 ‘Luca is playing with the ball’  
 b. *Luca ci=sta                    giocando*  
 Luca with.it=is            playing  
 ‘Luca is playing with it’  
 c. *Luca sta giocando*  
 ‘Luca is playing’

### 3. Instruments: arguments or adjuncts?

#### 3.1. The argument/adjunct distinction

The distinction between arguments and adjuncts is central to many theoretical frameworks (among others, Bresnan 1982; Chomsky 1981; Foley & Van Valin 1984; Pollard & Sag 1987), where the information about arguments is considered to be encoded in the semantic representation of verbs guiding the construction of sentences, while the information about adjuncts is not (Koenig *et al.* 2003).

The argument/adjunct distinction has traditionally been claimed to be categorical and dichotomic. However, various participants such as Locations, Passive Agents, and Instruments (Haspelmath 2014) do not neatly pattern as arguments nor as adjuncts and are thus difficult to classify. Moreover, the diagnostics for argumenthood are believed not to be always conclusive (Schütze 1995; Miller 1997; Koenig *et al.* 2003; Birtić & Brač *this volume*).

Because of these uncertainties, it has been proposed that the argument/adjunct distinction is non-categorical, but scalar (Langacker 1987; Schütze 1995; Croft 2001; Rissman 2013); that a third intermediate class exists, the one of quasi-arguments, secondary complements (DeAdmond & Hedberg 1998; Cardinaletti 2015; 2016; 2019) or “theta-marked adjuncts” (Speas 1990); or that each participant is represented as both an argument and an adjunct (Dowty 2003).

### 3.2. Instruments and the argument-adjunct classification

In the next sections, through the application of the main diagnostics for argumenthood to Italian Instruments, it will be shown that their behaviour is better understood by combining the hypothesis of a non-categorical argument-adjunct distinction with the hypothesis of a third class, the one of secondary complements.

#### 3.2.1. Semantic criteria

The main semantic diagnostics for argumenthood follow from the definitions of arguments provided over the years (e.g. Dowty 1982; Marantz 1984; Pollard & Sag 1987; Grimshaw 1990; Jackendoff 1990):

##### *I. Semantic obligatoriness*

Arguments are obligatorily entailed by the verb meaning, while adjuncts are not (Dowty 1982; Marantz 1984; Grimshaw 1990). Koenig *et al.* (2003) call it the Semantic Obligatoriness Criterion (SOC)<sup>163</sup>.

Koenig and colleagues (Koenig *et al.* 2003; Koenig *et al.* 2007) conducted a comprehensive survey on English verbs: they asked trained participants to answer two questions, given in (7):

- (7) a. Does one of the verb's senses describe situations in which one participant must or can use another participant to perform an action?  
Yes, it must/Yes, it can.
- b. Does one of the verb's senses describe situations in which one participant must or can perform an action described by the verb *with* something?  
Yes, it must/Yes, it can.

The first question is focused on the linking rule that associates the Instrument with the complement of the verb *to use*, the second question on the rule that associates it with the complement of the preposition *with*. The two questions target the positions Instruments can occur in, as mentioned in

<sup>163</sup> «**Semantic Obligatoriness Criterion** (SOC): If *r* is an argument participant role of predicate *P*, then any situation that *P* felicitously describes includes the referent of the filler of *r*.» (Koenig *et al.* 2003: 72)

(2). If the answers to both questions were *Yes, it must*, then the Instrument is semantically required by the verb; otherwise, the Instrument is semantically optional (allowed but not required by the verb).

Their study showed that Instruments «come in two categories. They are sometimes semantically required by the meaning of the verb; and they are sometimes not semantically required, but merely permitted» (Koenig *et al.* 2007: 177). In the former case, they are semantic arguments of the verb, in the latter they are not; in other words, only in the first case they are part of the verb meaning. The two types of verbs have different semantic representations (Koenig & Davis 2006). With verbs that semantically require an Instrument, henceforth verbs [+I], the latter is part of the semantic representation, and the preposition *with* only redundantly marks a portion of meaning already included in the verb's one. On the contrary, with verbs that semantically allow – but not require – the Instrument, henceforth verbs [±I], the instrumental component of meaning is added to the verbal semantics by the preposition.<sup>164</sup>

## II. Restricted range of verbs

As discussed by Marantz (1984), arguments can only occur with a restricted range of verbs – Semantic Specificity Criterion (Koenig *et al.* 2003) –<sup>165</sup> while “pure” adjuncts can modify virtually every verb.<sup>166</sup> This criterion is strictly linked to semantic obligatoriness: both, considered alone, are necessary

<sup>164</sup> In an unpublished work we replicated the survey of Koenig *et al.* (2003). We selected 752 verbs that entail an Instrument and asked eight trained participants – MA students and PhD candidates in Linguistics – to answer two questions for each verb. The questions are the Italian translation of the ones used by Koenig and colleagues. Overall, the collected judgements showed that the distinction individuated for Instruments in English holds for Italian as well, confirming the intuition that the contrast between two variants (one semantically obligatory and the other semantically optional, both syntactically optional) of the same role type is endemic to the Instrumental role.

<sup>165</sup> «**Semantic Specificity Criterion** (SSC): If *r* is an argument participant role of predicate *P* denoted by verb *V*, then *r* is specific to *V* and a restricted class of verbs/events.» (Koenig *et al.* 2003: 73)

<sup>166</sup> Not all adjuncts behave alike with respect to this criterion: some adjuncts display restrictions on the verb they can occur with. This is one of the reasons that led to the theorization of more than one class of adjuncts, e.g. event-internal vs. event-external, or theta-marked vs. pure adjuncts (Ernst 2001; Speas 1990).

but not sufficient conditions.<sup>167</sup> A participant must meet both conditions, semantic obligatoriness and the possibility of occurring with a restricted range of verbal heads in order to be classified as an argument.

Participants that pattern as semantic arguments are implied by a narrower range of verbs, as *del pericolo* ‘of the danger’ in (8a), while participants that pattern as adjuncts, not being implied in the semantic representation of the verb, can modify every event, as *nel pomeriggio* ‘in the afternoon’ in (8b) shows (the examples are adapted from Schütze 1995: 102).

- (8) a. *John {ha informato il suo amico/\*ha visto il suo amico/\*ha colpito il suo amico/\*ha ammirato il suo amico/\*ha sorpreso il suo amico} del pericolo.*  
 ‘John {informed his friend /\*saw his friend /\*hit his friend /\*admired his friend /\*surprised his friend} of the danger.’
- b. *John {è morto/ha starnutito/è esploso/si è rotto il braccio/ha visto Fred/ha riso di Bill} nel pomeriggio*  
 ‘John {died/sneezed/exploded/broke his arm/saw Fred/laughed at Bill} in the afternoon’

Restrictions on the distribution of instruments exist, i.e. instruments tend to co-occur with a narrow range of verbs (Brunson 1993; Levin 1993).

Instruments are only entailed by [+ dynamic] verbs that assign a causal role (Brunson 1993), e.g. *spostare* ‘to move’ (9a, 9b), while many other classes of verbs, like non-causal verbs, e.g. *sentire* ‘to hear’ (9c, 9d), stative verbs, e.g. *abitare* ‘to live’ (9e, 9f), psychological verbs, e.g. *pensare* ‘to think’ (9g, 9h), and verbs of change of state, e.g. *cadere* ‘to fall’ (9i, 9j), do not entail instruments.<sup>168</sup>

- (9) a. *Lara ha spostato la palla [con un bastone]*  
 ‘Lara moved the ball [with a stick]’
- b. *Lara ha usato [un bastone] per spostare la palla*  
 ‘Lara used [a stick] to move the ball’
- c. *#Lara ha sentito un rumore [con le cuffie]*  
 ‘#Lara heard a noise [with her headphones]’

<sup>167</sup> As Bresnan (1982) pointed out, locative and temporal participants are always entailed by events and should be considered arguments. However, they can modify every verb, i.e. they are not semantically specific. Considering both SOC and SSC prevents us from considering locative and temporal participants as arguments, when they are not.

<sup>168</sup> In the examples (9c-j), ‘#’ is used instead of ‘\*’ to underline a semantic more than syntactic anomaly, since we are talking about primarily semantic constraints. The symbol ‘\*’ marks as usual an ungrammatical sentence.

- d. *#Lara ha usato [le cuffie] per sentire un rumore*  
 ‘#Lara used [her headphones] to hear a noise’
- e. *#Lara abita a Roma [con una casa]*  
 ‘#Laura lives in Rome [with a house]’
- f. *#Lara usava [una casa] per abitare a Roma*  
 ‘#Lara was using [a house] to live in Rome.’
- g. *#Lara pensava a Luca [con una canzone]*  
 ‘#Lara was thinking of Luca [with a song].’
- h. *#Lara usava [una canzone] per pensare a Luca*  
 ‘#Lara was using [a song] to think of Luca’
- i. *#Lara cadde [con una spinta]*  
 ‘#Lara fell [with a push]’
- j. *#Lara usò [una spinta] per cadere*  
 ‘#Lara used [a push] to fall’

The second restriction on the distribution of Instruments states that they require the presence of an Agent (Gruber 1965), whether explicit (10a) or implicit/unspecified (10b). Otherwise, the sentence is ungrammatical (10c).

- (10) a. *Alberto ha rotto la finestra [con un sasso]*  
 ‘Alberto broke the window [with a rock]’
- b. *La finestra è stata rotta [con un sasso]*  
 ‘The window was broken [with a rock]’
- c. *\*La finestra si è rotta [con un sasso]*  
 ‘\*The window broke [with a rock]’

Finally, since the Agent is always required and [+ animate] (Lakoff, 1968), the subject DP associated with the Agent must always be [+ animate]. Thus, Instruments can only occur with a [+ animate] subject DP, as the contrast between (11a, 11b) and (11c, 11d) shows:

- (11) a. *Gianni ha ucciso Lucio [con la dinamite]*  
 ‘Gianni killed Lucio [with the dynamite]’
- b. *Gianni ha usato [la dinamite] per uccidere Lucio*  
 ‘Gianni used the dynamite to kill Lucio’
- c. *\*L’esplosione ha ucciso Lucio [con la dinamite]*  
 ‘\*The explosion killed Lucio [with the dynamite]’
- d. *\*L’esplosione ha usato [la dinamite] per uccidere Lucio*  
 ‘\*The explosion used [the dynamite] to kill Lucio’

The three main restrictions seen so far show that Instruments – unlike pure adjuncts, e.g. temporals or locatives – cannot freely modify every verb. On the contrary, they can only occur with [+ dynamic] [+ causative] verbs, within events that imply an Agent, the latter being associated with a [+ animate] subject DP.

Crucially, such constraints apply to both Instruments of [+I] and [±I] verbs: the examples in (9a, 9b), (10a, 10b), and (11a, 11b) involve Instruments of [±I] verbs which, in spite of being semantically optional, behave like semantic arguments as far as the Semantic Specificity Criterion is involved. The same is true for Instruments of [+I] verbs, which in addition are semantically obligatory.

### III. Dependence on the head for the interpretation

The interpretation of arguments depends on the verbs that select them, while the interpretation of adjuncts is independent (Grimshaw 1990; Marantz 1984; Pollard & Sag 1987), see (12):

- (12) *Maria ha {regalato un libro/dato ascolto} a Marco ieri*  
 ‘Maria {gave a book / listened} to Marco yesterday’

While the interpretation of *ieri* ‘yesterday’ does not change depending on the predicate (‘to give’/ ‘to listen’), the interpretation of *a Marco* ‘to Marco’ does, depending on the verb by which it is selected. Thus, the PP is an argument, while the AdvP is an adjunct.

Different interpretations of Instruments have been identified as well, which depend on the predicate by which they are entailed (Marantz 1984; Dowty 1994; Schütze 1995).

- (13) a. *Avvistare la cometa con un telescopio*  
 ‘Sight the comet with a telescope’  
 b. *Spazzare il pavimento con la scopa*  
 ‘Sweep the floor with the broom’  
 c. *Attaccare un quadro con i chiodi*  
 ‘Fasten a picture with the nails’  
 d. *Coprire il tavolo con la tovaglia*  
 ‘Cover the table with the tablecloth’

In (13a), no change in the position of the Instrument is entailed; in (13b), the Instrument moves but not to a specific position; in (13c), the

Instrument ends up to a specific position; in (13d), the Instrument shows a higher degree of affectedness, comparable to that of direct objects. The four interpretations depend on the verbs and not on the Instruments themselves: *tovaglia* ‘tablecloth’ in (13d) could be replaced by *telescopio* ‘telescope’ and then the telescope would show a higher degree of affectedness than it does in (13a) (Dowty 1994).

Marantz (1984) and Schütze (1995) identified a difference in the semantic interpretation of Instruments that perhaps makes an even stronger case for the argumenthood of Instruments.

Instruments can be interpreted as playing either a causal role or a merely helping one (Marantz 1984). This distinction was confirmed by Ono (1992), who distinguished between intermediary and facilitating Instruments;<sup>169</sup> this was successively expanded on (Schlesinger 1995; Koenig *et al.* 2003; 2007), and three different types of Instruments were identified: intermediary (14a), enabling (14b) and ancillary (14c).

- (14) a. *Alberto ha tritato i funghi [con il coltello]*  
 ‘Alberto chopped the mushrooms [with the knife]’  
 b. *Lucia è entrata nella stanza [con la chiave]*  
 ‘Lucia entered the room [with the key]’  
 c. *Marco mangia la zuppa [con il cucchiaino]*  
 ‘Marco eats the soup [with the spoon]’

Only in (14a) the Instrument plays an actual causal role in that the act of the Agent (i.e. picking and moving the knife) directly causes the contact between the Instrument and the Patient (i.e. the knife and the mushrooms) which in turn directly causes a change of state or location in the Patient (here it causes the mushrooms to become chopped); in (14b), the Instrument (i.e. the key) enables Lucia to enter the room, but it surely does not cause her

<sup>169</sup> Ono (1992) hypothesized that intermediary Instruments are always semantically required by and included in the semantic representation of the verb, while facilitating (i.e. ancillary) are not. Nonetheless, the terms ‘semantically obligatory’ and ‘semantically optional’ cannot be considered as equivalent of ‘intermediary’ and ‘facilitating’, given that some causal instruments are semantically optional, e.g. the verb *rompere* ‘to break’ is considered to semantically allow but not require an Instrument; yet, when the Instrument is present, it always plays a causal role (but the opposite is not true, i.e. facilitating instruments are never obligatory). Therefore, it is better to maintain the two distinctions separate.

to enter it. In this case, the Instrument is a necessary precondition for the event to happen. In (14c), the Instrument (i.e. the spoon) just helps Marco to eat the soup, where the concept of “helping” is defined as follows (Koenig *et al.* 2007: 214):

«An eventuality  $e_1$  helps the occurrence of token  $e_2$  of the event category C [if and only if] (i) there is an ordering of tokens of C along a pragmatically defined scale (ease of performance, how good the resulting state is, fewer unwelcome ‘side effects’); (ii)  $e_1$  caused the token  $e_2$  of C to be higher on that ordering than it would otherwise have been.»

In (14c) the spoon is the eventuality that causes the event of eating (C) to be higher on some pragmatically defined scales, e.g. ease of performance (it is easier to eat a soup with a spoon than it would be with bare hands), fewer side effects (if the soup is hot, the eater could get burnt without a spoon), etc.

The interpretation of Instruments as intermediary, enabling or ancillary depends on the event by which the Instrument is implied and, thus, on the verb that selects it. Consequently, according to this criterion, both [I+] and [I±] Instruments behave like semantic arguments.

Recalling what was said in section 2.1, it is now clearer why a definition of Instruments in purely causal terms does not cover all the possible interpretations of Instruments.

In this section, we inspected the main semantic criteria for semantic argumenthood; according to the first criterion, two classes of Instruments were identified, one of semantically obligatory and one of semantically optional Instruments. Both classes are semantic arguments according to the second and third criteria as well.

Instruments of [+I] verbs are semantic arguments, Instruments [±I] verbs are secondary participants – because of their optionality – yet selected for and entailed by the verb.

### 3.2.2. Syntactic diagnostics

Following the assumption that the argument/adjunct distinction made at the level of argument structure reflects onto syntax (Levin & Rappaport 1988; Grimshaw 1990; Jackendoff 1990; Speas 1990), several syntactic diagnostics were proposed in order to establish the status of instruments (for a comprehensive review, see Schütze 1995).

### I. Iterativity test

This test is derived from the Theta-criterion (Chomsky 1981) and its subsequent formulations (Bresnan 1982; Pollard & Sag 1987). The iterativity test distinguishes arguments from adjuncts in that phrases expressing arguments cannot freely iterate (15a), whereas phrases expressing adjuncts can, as long as they differ in their level of granularity (Goldberg 2002), as (15b) shows (Rome – city > hotel – building > hall – room). If two phrases expressing adjuncts do not differ in their level of granularity, the iteration is disallowed, e.g. \**They met at the hotel at the restaurant* (hotel – building = restaurant – building).

- (15) a. \**Hanno prestato il libro a Maria a Laura*  
 ‘\*They lent the book to Maria to Laura’  
 b. *Li ha incontrati a Roma nell'albergo nella hall*  
 ‘She met with them in Rome in a hotel in the hall’

With respect to the iterativity test, instrumental phrases (inst-PP) behave like arguments: they cannot be iterated whether they are semantic arguments of [+I] verbs (16a) or secondary participants of [±I] verbs (16b).

- (16) a. \**Ha incollato i fogli [con la colla con lo scotch]*  
 ‘\*He glued the paper with the glue with the tape’  
 b. \**Ha colpito l'auto [con un bastone con un sasso]*  
 ‘\*She hit the car with a stick with a rock’

According to the iterativity test, inst-PPs expressing Instruments of either [+I] or [±I] verbs behave like arguments.

### II. (Revised) Ordering test

This test relies on the assumption that phrases expressing arguments typically precede phrases expressing adjuncts. Here, the revised version of this test will be applied, which states that a phrase expressing an adjunct cannot precede a phrase expressing an argument, even when the former is focused (Schütze 1995), as (17b) shows (the examples are translated from Schütze 1995):

- (17) a. *Chris ha messo un libro sul tavolo dopo pranzo*  
 ‘Chris put a book on the table after lunch’  
 b. \**Chris ha messo un libro dopo PRANZO sul tavolo (non dopo cena)*  
 ‘\*Chris put a book after LUNCH on the table (not after dinner)’

Both Instruments of [+I] and [±I] verbs can occur before phrases expressing arguments, as in (18b) and (18d), respectively:

- (18) a. *Sara ha legato il carretto al palo [con una corda]*  
 ‘Sara tied the handcart to the stake [with a rope]’  
 b. *Sara ha legato con una CORDA il carretto al palo (non con lo scotch)*  
 ‘Sara tied with a ROPE the handcart to the stake (not with the tape)’  
 c. *Sara ha mangiato la zuppa con il cucchiaino*  
 ‘Sara ate the soup with a spoon’  
 d. *Sara ha mangiato con il CUCCHIAIO la zuppa (non con le bacchette)*  
 ‘Sara ate with the SPOON the soup (not with the chopsticks)’

According to this test, inst-PP expressing Instruments of either [+I] or [±I] verbs behave like arguments.

### III. Anti-reconstruction effect test

The C condition of the binding theory differently applies to *wh*-phrases containing referential expressions depending on whether they are arguments or adjuncts, as shown by the contrast in (19) (the examples are translated from Ono 1992: 199):

- (19) a. *\*Quali foto di John<sub>i</sub> pro<sub>i</sub> ha distrutto?*  
 ‘\*Which pictures of John<sub>i</sub> did he<sub>i</sub> destroy?’  
 b. *?Quali foto vicino a John<sub>i</sub> pro<sub>i</sub> ha distrutto?*  
 ‘Which pictures near John<sub>i</sub> did he<sub>i</sub> destroy?’

The contrast between (19a) and (19b) shows that coreference is possible only when the referential expression is an adjunct (Speas 1990). Speas (1990) additionally points out that not all adjuncts show anti-reconstruction effects; that is, adjuncts which are assigned a theta role by the verb (“theta-marked adjuncts”) do not show anti-reconstruction effect, e.g. the benefactive *for Mary’s brother* in *\*For Mary<sub>i</sub>’s brother, she<sub>i</sub> was given some old clothes* (Speas 1990: 52), while “pure” adjuncts, i.e. those which are not assigned a theta role by the verb, do, e.g. the rationale *for Mary’s valor* in *For Mary<sub>i</sub>’s valor, she<sub>i</sub> was awarded a purple heart* (Speas 1990: 52).

This asymmetry in anti-reconstruction effect can be used to determine whether Instruments are arguments or adjuncts, or, more precisely, the hypothesis that instruments are “pure” adjuncts can be ruled out if they do not show anti-reconstruction effects. Consider (20):

- (20) a. \**Con quale coltello di Marta, lei<sub>i</sub> ha tagliato le verdure?*  
 ‘\*With which of Marta<sub>i</sub>’s knives, did she<sub>i</sub> chop the vegetables?’  
 b. \**Con quale cucchiaino di Marta, lei<sub>i</sub> ha mangiato la zuppa?*  
 ‘\*With which of Marta<sub>i</sub>’s spoons did she<sub>i</sub> eat the soup?’

As shown in (20a), Instruments of [+I] verbs do not display anti-reconstruction effects, which is consistent with the hypothesis that they are arguments. (20b) shows that Instruments of [±I] verbs do not display anti-reconstruction effects either. Thus, neither of them is a “pure” adjunct according to this test (Ono 1992).

#### IV. Pro-form replacement

The pro-form replacement test has been used to prove that Instruments are adjuncts, since they cannot be replaced by the verbal pro-form *farlo* ‘to do so’ (Levin & Rappaport 1988; Randall 1992; Rissman 2013). Consider (21):

- (21) a. *Sara ha riempito il camion di/con il fieno [con un forcone]*  
 ‘Sara loaded the truck with hay with a fork’  
 b. \**Sara ha riempito il camion di/con il fieno e Marco l’ha fatto di/con il cartone*  
 ‘\*Sara loaded the truck with hay and Marco did so with cardboard’  
 c. *Sara ha riempito il camion di fieno [con un forcone] e Marco l’ha fatto [con una gru]*  
 ‘Sara loaded the truck with hay [with a fork] and Marco did so [with a crane]’  
 d. *Sara ha attaccato il poster [con la colla] e Marco l’ha fatto [con lo scotch]*  
 ‘Sara attached the poster with the glue and Marco did so with the tape’

While the *locatum* in (21a) *di/con il fieno* ‘with the hay’ cannot be left out from the verbal pro-form, hence the ungrammaticality of (21b), inst-PP expressing Instruments of [+I] and [±I] verbs are left out from the pro-form, as in (21c) and (21d).

However, the pro-form substitution only singles out elements that are placed under the V’ node (Ono 1992). That is, the test does not demonstrate that the inst-PP are adjoined to VP; rather, it only excludes that they are under the V’ node. Their syntactic structure is something like (22):

- (22) Sara ha [<sub>VP</sub> [<sub>V’</sub> riempito il camion di/con il fieno] con un forcone]  
 ‘Sara [<sub>VP</sub> [<sub>V’</sub> loaded the truck with hay] with a fork]’

The representation in (22) is consistent with their being “theta-marked adjuncts”, secondary complements or quasi arguments.

### V. Cliticization

The possibility of being cliticized is perhaps the strongest case for considering instruments secondary complements (Cardinaletti, 2015; 2016; 2019). In Italian and other Romance languages, virtually every primary – (23a), (23a) – and secondary – (23b), (23b) – complement of the verb can be cliticized while VP-external adjuncts cannot – (23c), (23c). The following examples are adapted from Cardinaletti (2016: 50).

- (23) a. *Sara mangia un panino*  
 ‘Sara eats a sandwich’  
 a’. *Sara lo=mangia*  
 Sara it=eats’  
 ‘Sara eats it’  
 b. *Esco sempre con Gianni*  
 ‘I always go out with Gianni’  
 b’. *Ci=esco sempre*  
 with.him=go\_out always  
 ‘I always go out with him’  
 c. *Gianni è felice a casa dei genitori*  
 ‘Gianni is happy at his parents’  
 c’. \**Gianni ci=è felice*  
 Gianni there=is happy’  
 ‘\*Gianni is happy there’

Inst-PPs can be cliticized, whether they express Instruments of [+I] or [±I] verbs, as in (24a, 24a) and (24b, 24b), respectively.

- (24) a. *Taglio il pane con il coltello*  
 ‘I cut the bread with the knife’  
 a’. *Ci=taglio il pane*  
 with.it=cut.PRS.1SG the bread’  
 ‘I cut the bread with it’  
 b. *Ho aperto la scatola con le forbici*  
 ‘I opened the carton box with the scissors’  
 b’. *Ci=ho aperto la scatola*  
 with.them=have.PRS.1SG opened the box’  
 ‘I opened the carton box with them’

The possibility of being cliticized supports the view of instruments as secondary complements.

### 3.2.3. Summary

In this section, we applied three semantic criteria and five syntactic diagnostics to Instruments in Italian.

Two types of Instruments are identified by the Semantic Obligatoriness Criterion, i.e. Instruments that are semantically required and Instruments which are semantically allowed but not required by the verb; according to the other semantic criteria, Instruments of [+I] verbs are arguments and Instruments of [±I] verbs secondary participants, more strongly entailed by the verb meaning than adjuncts and assigned the theta role by it. Iterativity test – the syntactic diagnostics more intertwined with semantics – leads us to the same conclusion.

As for the other syntactic diagnostics, the pro-form replacement test only excludes that Instruments of [+I] and [±I] verbs are syntactic arguments, i.e. that they are under the V' node. However, according to the anti-reconstruction effect test, both of them behave like arguments, and the possibility of being cliticized excludes that they are adjuncts.

The tests applied so far converge in excluding that Instruments of [+I] and [±I] verbs are pure adjuncts and in demonstrating that they are secondary complements or quasi-arguments occupying an intermediate position between the V' and the VP node, and thus sharing more properties with arguments than with adjuncts.

## 4. The omission of Instruments: between semantics and syntax

The optionality of Instruments, i.e. the possibility of being omitted without the sentence becoming ungrammatical, has been considered the strongest argument in favour of their being adjuncts.

Nonetheless, after having demonstrated that Instruments of neither [+I] nor [±I] verbs are adjuncts, we claim that syntactic optionality of Instruments is not an argument in favour of their being adjuncts; instead, their syntactic omission should be considered as an instance of null arguments, something analogous to the null object or object drop, as will be discussed in detail in section 4.1.

## 4.1. Shadow, Default and Open Instruments

Schütze (1995: 101), paraphrasing Jackendoff (1977), observes that «arguments to a particular lexical head can be obligatory, whereas modifiers are (almost) always optional, but the converse is not true: there are **optional arguments**.» (our emphasis)

The most studied optional argument is the direct object (cf. Levin 1989; Resnik 1993; Cappelli & Lenci 2020; Cappelli 2022; among others). Its omission is called *object deletion*, *intransitivization*, and *object-drop*; the omitted objects are referred to as *null complements*, *implicit objects*, and *dropped objects*.

Verb meaning plays a major role in determining argument omission (Levin 1993). In particular, a “key determinant” (Cappelli & Lenci 2020: 131) factor in argument omission is their semantic recoverability (Hopper & Thompson 1980; Levin 1993; Resnik 1993, 1996; Conklin *et al.* 2004; Medina 2007; Glass 2020), determined by the verb meaning.

The notion of semantic recoverability, i.e. the possibility for the omitted argument to be interpreted by the listener through the semantics of the verb (and of the other arguments), is strictly related to the amount of information arguments and verbs provide about each other (Pustejovsky 1995). Pustejovsky (1991, 1995; for Italian, Ježek 2017, *this volume*) distinguishes among three types of arguments based on their informativity with respect to the verb and on their recoverability from the verb meaning alone.

We applied such a distinction to Instruments in Italian, identifying Shadow, Default and Open Instruments.

### 4.1.1. Shadow Instruments

In Italian, Shadow Instruments (SI) – as shadow arguments – are «semantically incorporated in the meaning of the verb» (Ježek 2017: 15)<sup>170</sup> and, thus, completely recoverable from the verb meaning alone and minimally informative with respect to the verb.

---

<sup>170</sup> We will only refer to semantic incorporation, in order to avoid the discussion on whether these verbs are examples of incorporation à la Baker (1988) or are derived from different syntactic/morphological/lexical processes, such as conversion or zero-suffixation (Hale & Keyser 1993; Fabrizio 2013). Albeit being extremely interesting, this discussion is beyond the scope of this work.

An example of a verb entailing a SI is *spazzolare* ‘to brush’, which semantically incorporates the Instrument *spazzola* ‘brush’ (25). Therefore, SI are always interpreted as a single instrumental lexical item, e.g. for the sentence *Maria martella il muro* ‘Maria hammers the wall’, the only possible interpretation is that she is doing it with a hammer (hence, the SI is interpreted as a single lexical item).

- (25) *Spazzolare* ‘to brush’  
 ⇒ INST: {*spazzola* ‘brush’}

Other examples of verbs entailing SI (henceforth, shadow-verbs) in Italian are *telefonare* ‘to phone’ (INST: {*telefono* ‘phone’}), *pettinare* ‘to comb’ (INST: {*pettine* ‘comb’}), *pugnalarlo* ‘to stab’ (INST: {*pugnale* ‘dagger’}), etc.

Since SI are uninformative with respect to the verb, they must be omitted in the surface structure, as shadow arguments must be (Pustejovsky 1995; Ježek 2017), in order to avoid redundancy: this is the Non-Redundancy Constraint (Fabrizio 2013).

- (26) a. *Sara spazzola suo fratello*  
 ‘Sara brushes her brother’s hair’  
 b. ??*Sara spazzola suo fratello con la spazzola*  
 ‘??Sara brushes her brother’s hair with the brush’  
 c. *Sara spazzola suo fratello con la spazzola che ha comprato ieri*  
 ‘Sara brushes her brother’s hair with the brush that she bought yesterday’

The comparison between (26a) and (26b) clearly shows that the acceptability of the latter is degraded. However, (26c) is acceptable: Shadow Instruments can be syntactically realized when they are modified, since only through modification they become informative with respect to the verb again.

#### 4.1.2. Default Instruments

Default Instruments (DI) – as default arguments – are selected by the verb and consist of a restricted and semantically coherent class of entities, as in (27):

- (27) *Tagliare* ‘to cut’  
 ⇒ INST: {*cutting\_objects*}

Other examples of verbs entailing DI (henceforth, default-verbs) are *sparare* ‘to shoot’ (INST: {firearms}), *bucare* ‘to pierce’ (INST: {piercing\_ objects}), *vestire* ‘to dress’ (INST: {clothes}), etc.

When referring to the entire class of entities, DI are uninformative with respect to the verb and maximally recoverable from its meaning alone; therefore, they are usually omitted. Default arguments are syntactically realized only when the speaker wants to refer to a single instance of the class, since only in that case they are informative with respect to the verb. Consider the examples in (28), for the verb *mangiare* ‘to eat’, whose default object is the class {food}. (28a) is perfectly fine: the default object is omitted, and interpreted as the whole class of entities, while (28b) is weird, because the syntactically realized object is redundant. (28c) is acceptable, since the speaker refers to a single instance of the class, making the default object informative:

- (28) a. *Ieri ho mangiato \_*  
 ‘Yesterday I ate \_’  
 ⇒ OBJ: {food}
- b. *??Ieri ho mangiato cibo*  
 ‘Yesterday I ate food’
- c. *Ieri ho mangiato la pizza*  
 ‘Yesterday I ate pizza’

On the contrary, DI are often shadowed (i.e. interpreted as a single instrumental lexical item, like SI) by the co-combination of the verb and its internal complement, as in (29). As a consequence, DI are omitted more often than, for instance, default objects.

- (29) a. *Tagliare le verdure*  
 ‘To chop the vegetables’  
 ⇒ INST: {coltello ‘knife’}
- b. *Tagliare il prato*  
 ‘To cut the lawn’  
 ⇒ INST: {tagliaerba ‘lawn mower’}
- c. *Tagliare i capelli*  
 ‘To cut someone’s hair’  
 ⇒ INST: {forbici ‘scissors’}

DI are slightly more informative than SI with respect to the verb meaning; however, they are still highly retrievable (i) from the verb alone,

which usually selects for a semantically coherent narrow class of entities, and (ii) from the direct object combined with the verb.

#### 4.1.3. Open Instruments

In Pustejovsky's words (1995), true arguments encode grammatically relevant participants, but not inherent properties of the verb; they must be syntactically realized in order for the verb meaning to be complete and for the verb to be interpreted.

As for Instruments, the third class is that of Open Instruments (OI) – instead of “True” – because OI, unlike true arguments, are still syntactically optional;<sup>171</sup> this label highlights the fact that OI are unrecoverable from the verb meaning alone – which usually selects for a broad (and not always semantically coherent) class of entities (30) – without entailing their syntactic obligatoriness.

- (30) *Rompere* ‘to break’  
 ⇒ INST: {*martello* ‘hammer’, *palla* ‘ball’, *mazza* ‘bat’, etc.}

Other examples of verbs that entail OI (henceforth, open-verbs) are *andare* ‘to go’, *distruggere* ‘to destroy’, *sporcare* ‘to dirty’, etc.

In some cases, OI can be recovered via pragmatic factors, such as the speakers' background knowledge. Compare (31a) with (31b):

- (31) a. *Marta ha distrutto la parete*  
 ‘Marta destroyed the wall’  
 ⇒ INST: {*martello* ‘hammer’, *dinamite* ‘dynamite’, *fuoco* ‘fire’, etc.}  
 b. *Marta è andata a Parigi*  
 ‘Marta went to Paris’  
 ⇒ INST: {*aereo* ‘airplane’, *treno* ‘treno’, *auto* ‘car’, etc.}

In (31a), the OI is unrecoverable when syntactically omitted. In (31b), if the listener knows the starting point of Marta's trip, the range of possible Instruments narrows down: for instance, if Marta went to Paris from Dourdon, then she is likely to have used the train or a car. If she went to Paris from Cape Town, then airplane is the likeliest Instrument.

<sup>171</sup> Such property is consistent with their syntactically being secondary complements.

OI are more informative with respect to the verb – and less recoverable from its meaning alone – than SI and DI, and therefore more likely to be syntactically realized.

#### 4.2. Relating semantics and syntax: the pattern of omission of Instruments

In the previous section, three classes of Instruments were identified:

- (i) SI are uninformative with respect to the verb and maximally recoverable from its meaning; they must be omitted unless they are further modified.
- (ii) DI are recoverable from the verb meaning – and uninformative – when the speakers refer to the entire class of entities, and thus omitted; unlike default arguments, they are also omitted when shadowed by the co-combination of the verb and the direct object.
- (iii) OI are maximally informative with respect to the verb and minimally recoverable from its meaning. They tend to be more often syntactically realized.

Hence, we have:

<b>Semantic recoverability:</b>	Shadow	>	Default	>	Open
	↓				
<b>Syntactic omission:</b>	Shadow	>	Default	>	Open

Omission of Instruments is principled; namely, it is ruled by semantic recoverability, as it happens for object drop; the more an Instrument is recoverable from the verb meaning, the more it is syntactically omitted. In addition, Instruments tend to be more frequently recoverable from the verb meaning than, e.g. direct objects, in that, for instance, DI can be shadowed by the co-composition between the verb and its internal argument, as in (29), and pragmatic factors can contribute to ease the recoverability of OI, as in (31b).

If Instruments are only produced when they are not recoverable from the verb meaning, we expect them to be rarely syntactically realized and to be produced more often in co-occurrence with open-verbs than with shadow- and default-verbs. Finally, we expect them to co-occur with shadow- and default-verbs only under specific circumstances, i.e. when they are not recoverable. For instance, we expect them to co-occur with shadow-verbs when they are

modified and to co-occur with default-verbs when the speakers want to refer to a single Instrument of the selected class, as long as it is not shadowed by the co-composition of the verb and its internal argument.

## 5. Corpus analysis

A comprehensive corpus analysis of Italian spontaneous speech was performed in order to investigate the pattern of production/omission of Instruments as well as to investigate the validity of our proposal; therefore, we focused on (i) the relative frequencies of occurrence of the inst-cl and the inst-PP among adult speakers; (ii) which verbs they co-occur with.

### 5.1. Selection of corpora and methodology

Three spoken corpora of Italian were analysed:

- **Lessico di frequenza dell'italiano parlato (LIP, DeMauro *et al.* 1993, VOLIP Alfano *et al.* 2014)**: only the sub-corpora containing spontaneous speech (conversations, phone conversations, debates, etc.) were considered. For the clitic pronoun *ci*, 1.683 utterances were analysed, 2.378 occurrences were manually classified depending on their function; for the *con*-PP, 932 utterances were analysed, 1.182 occurrences were manually classified depending on their function.
- **KIP (Mauri *et al.* 2019)**: only three types of texts were considered (i.e. conversations, semi-structured interviews and meetings between students and professors). For the clitic pronoun *ci*, 87 texts were analysed, 3.838 occurrences were manually classified depending on their function; for the *con*-PP, 88 texts were analysed, and 1.688 occurrences were manually classified depending on their function.
- **ParlaTO (Cerruti & Ballarè 2021)**: all the semi-structured interviews were considered. For the clitic pronoun *ci*, 49 texts were analysed, 5.265 occurrences were manually classified depending on their function; for the *con*-PP, 50 texts were analysed, 1.454 occurrences were manually classified depending on their function.

A total of 11.481 occurrences of the clitic *ci* and 4322 of the *con*-PP were found in the corpora and considered for the present analysis. Ten functions of the clitic pronoun were considered: (i) accusative (first person plural);

(ii) dative (first person plural); (iii) inflected forms of the verb *esserci* ‘to be there’; (iv) locative; (v) instrumental; (vi) comitative; (vii) inflected forms of the verb *averci* ‘to have’; (viii) complex verbs (where the clitic only has an intensifying function, but no meaning); (ix) reflexive/reciprocal; (x) clitic cluster *ci si*.

For the *con*-PP, on the contrary, the range of functions is narrower. *Con*-PPs cover less functions than the clitic *ci* in Italian. The main functions of *con*-PPs, in Italian, are comitative, manner and causal. We did not consider the causal function of *con*-PPs as a separate function in the classification, as it was outside of the scope of our work. Although a more fine-grained classification of *con*-PPs would be possible, the presence of idiosyncrasies in their use in Italian would have made it extremely difficult to elaborate such a classification. We therefore decided to maintain the categorization simple in order to focus on the *inst*-PP and avoid as many uncertainties as possible. Three functions are considered: (i) instrumental; (ii) comitative; (iii) modal; the occurrences that do not fall into either of these classes are coded as (iv) others.

## 5.2. Results

### 5.2.1. Frequencies of occurrence

In the LIP corpus, the *inst-cl* only represents the 0.38% of the total; in the KIP the 0.52% and in the ParlaTO the 0.53%; considering all the corpora, the *inst-cl* counts 57 occurrences out of 11.481 (Table 1). Concerning the *con*-PP, the frequencies of occurrence of its instrumental function are higher: the *inst-PP* represents the 7.58% of the total in the LIP corpus, the 5.63% in the KIP and the 6.19% in the ParlaTO, with 275 occurrences out of 4.322 considering all the corpora (Table 1).

Table 1. Absolute and relative (%) frequencies of the *inst-cl* and the *inst-PP*, in LIP, KIP, ParlaTO.

Corpora	<i>inst-cl</i> (absolute)	<i>inst-cl</i> (%)	<i>inst-PP</i> (absolute)	<i>inst-PP</i> (%)
<i>LIP</i>	9/2378	0.38	90/1182	7.58
<i>KIP</i>	20/3838	0.52	95/1686	5.63
<i>ParlaTO</i>	28/5265	0.53	90/1454	6.19
<i>All corpora</i>	57/11841	0.50	275/4322	6.36

Figure 1 shows that the instrumental function of the clitic pronoun *ci* is the least produced of all the functions considered. Figure 2 shows that the instrumental function of the *con*-PP – despite counting more occurrences than the *inst-cl* (s. Table 1) – is far from being frequently produced, when compared with the main other functions of the *con*-PP.

Figure 1. Frequency (%) of occurrence of the clitic pronoun *ci*, functions (i-x), all corpora.

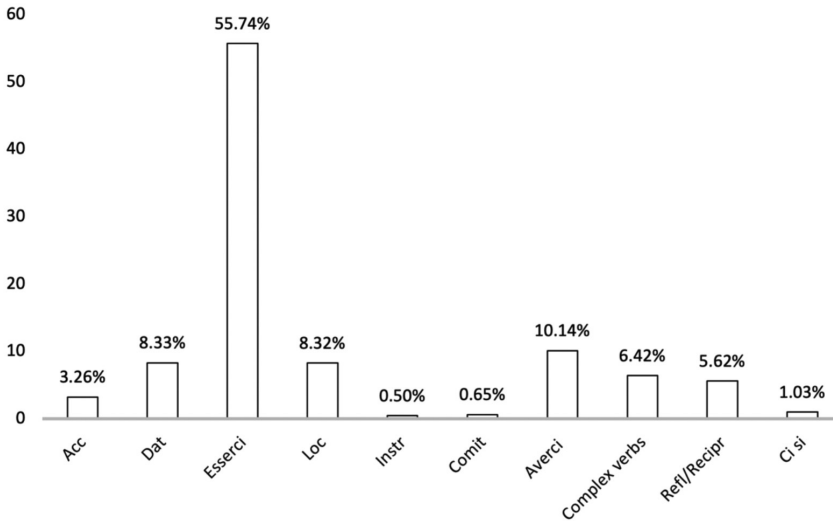
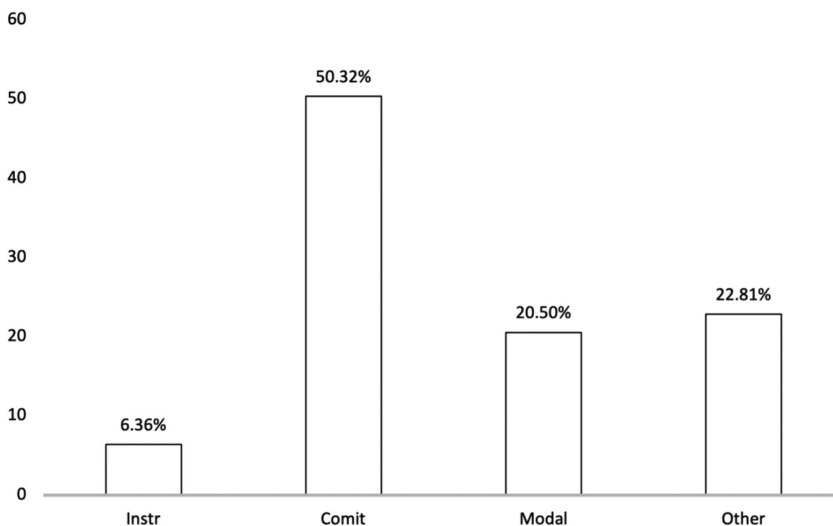


Figure 2. Frequency (%) of occurrence of the *con*-PP, functions (i-iv), all corpora.



The corpus analysis revealed that in spontaneous speech, both the inst-cl and the inst-PP are scarcely produced, with the inst-PP being more produced than the clitic; such data are consistent with our first prediction, i.e. Instruments are often recoverable from the verb alone and therefore omitted. A complete list of all the verbs that co-occur with inst-PPs and inst-cl is found in the Appendix with the total occurrences of each verb in the corpora.<sup>172</sup>

### 5.2.2. Types of verbs that co-occur with Instruments

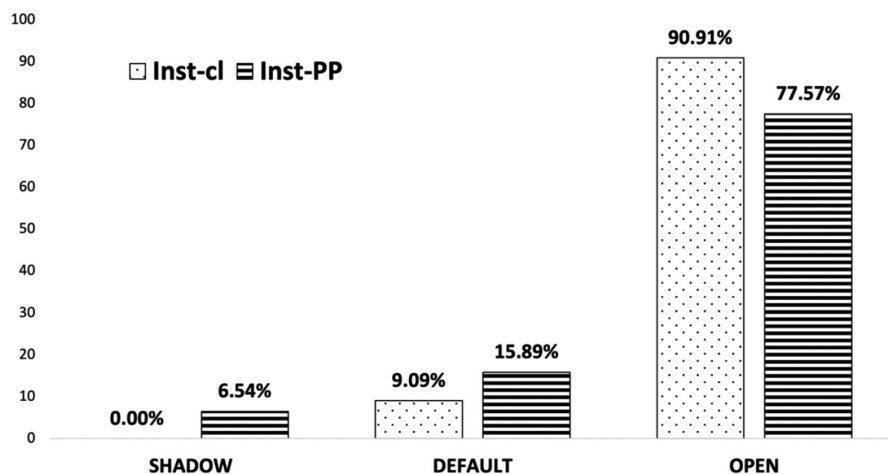
Information about which types of verbs (shadow-, default- or open-) Instruments tend to co-occur with is useful to have a better insight on their pattern of production/omission.

The inst-cl – which counts 57/11.481 occurrences – co-occurs with 10 different verbs; one out of 10 is a default-verb (*scrivere* ‘to write’), nine are open-verbs (e.g. *creare* ‘to create’, *giocare* ‘to play’, *costruire* ‘to build’, etc.), and zero co-occurrences with shadow-verbs are found.

The inst-PP – which counts 275/4.322 occurrences – co-occurs with a wider range of verbs, i.e. 107 different verbs: of these, 7 are shadow-verbs (e.g. *stampare* ‘to print’, *avvelenare* ‘to poison’, *telefonare* ‘to phone’), 17 are default-verbs (e.g. *scrivere* ‘to write’, *mangiare* ‘to eat’, *pescare* ‘to fish’, etc.) and 83 are open-verbs (e.g. *uccidere* ‘to kill’, *distruggere* ‘to destroy’, *aprire* ‘to open’, etc.). Figure 3 shows the percentages of co-occurrence with such verbs:

<sup>172</sup> If it is conceivable to hypothesize that inst-PPs/inst-cl are rarely produced because the verbs that entail them are less frequent than the verbs that entail, for instance, a locative or comitative argument/adjunct, the list in the Appendix shows that this is not the case. For instance, among the verbs that most frequently co-occur with an Instrument, *fare* ‘to do’ (13613 occurrences), *andare* ‘to go’ (7385 occurrences), and *venire* ‘to come’ (7385 occurrences) are found. Crucially, *andare* and *venire* require a locative argument.

Figure 3. Frequency (%) of co-occurrence of shadow-, default-, open-verbs with inst-cl and inst-PP, all corpora.



As Figure 3 shows, both the inst-cl and the inst-PP tend to co-occur more frequently with open-verbs than with shadow- and default-verbs: indeed, consistently with our proposal, they are more informative (and not redundant) with respect to open-verbs than they are with respect to shadow- and default-verbs.

### 5.2.3. Instruments co-occurring with shadow- and default-verbs: special circumstances

In order to provide a better insight of the production/omission pattern of Instruments in Italian, the cases where inst-PP and inst-cl co-occur with shadow-, default-, and open-verbs must be investigated in detail. This analysis is necessary for two reasons: on the one hand, not taking into consideration the frequency of occurrence of shadow-, default- and open-verbs (as in Figure 3) within the three corpora does not provide an exhaustive picture of the instrument distribution; on the other hand, looking at the raw proportion of co-occurrence of an inst-PP/cl with a verb relative to the total occurrences of that verb may be misleading. See Table 2 and Table 3, where these proportions are shown for the inst-PP and the inst-cl, respectively.<sup>173</sup>

<sup>173</sup> As stated in section 2.2, 'inst-PP' and 'inst-cl' refer to *con*-PPs and the clitic pronoun *ci* when they are used in their instrumental function.

In Tables 2 and 3, the column “Total occurrences of the verbs” refers to the total occurrences of the verbs with which inst-PPs and inst-cl co-occur, not to all verbs in the corpora.

Table 2. Proportion co-occurrence with an inst-PP/total occurrences, for shadow-, default-, open-verbs, all corpora.

Verb type	Number of verbs (lexemes)	Co-occurrence with an inst-PP (occurrences)	Total occurrences of the verbs	Proportion of co-occurrence/total occurrences
<i>Shadow-verbs</i>	7	11	412	0.027
<i>Default-verbs</i>	17	35	4736	0.007
<i>Open-verbs</i>	83	229	48138	0.005
<i>Total</i>	107	275	53283	0.005

Table 3. Proportion co-occurrence with an inst-cl/total occurrences for shadow-, default-, open-verbs, all corpora.

Verb type	Number of verbs (lexemes)	Co-occurrence with an inst-cl (occurrences)	Total occurrences of the verbs	Proportion of co-occurrence/total occurrences
<i>Shadow-verbs</i>	0	---	---	---
<i>Default-verbs</i>	1	1	897	0.001
<i>Open-verbs</i>	10	56	16770	0.003
<i>Total</i>	11	57	17667	0.003

As suggested above, Table 2 and Table 3 provide relevant information about the production/omission of Instruments: they show that Instruments (both inst-PPs and inst-cl) rarely co-occur with open-verbs, too. Indeed, the co-occurrence with Instruments represents a small fraction of the total occurrences of such verbs (0.005 for inst-PPs, 0.003 for the clitic). This is consistent with our proposal, since – as shown in (31b) – OI can be made recoverable by pragmatic factors, i.e. by both the linguistic and the extralinguistic context. Looking at the proportion of co-occurrence with the inst-cl relative to the total occurrences of the verbs shown in Table 3, the distribution of the inst-cl is also consistent with our proposal, i.e. open-verbs occur with the inst-clitic more frequently (0.003) than shadow- (0 occurrences) and default-verbs (0.001) do.

However, looking at the same proportion in Table 2, one could conclude that actually shadow-verbs tend to co-occur more frequently with inst-PP (0.027) than default- (0.007), and default-, in turn, more often than open-verbs (0.005). This is why the raw proportion may be misleading, if not complemented by additional data on the circumstances where the inst-PP co-occur with the three types of verbs.

The core of our hypothesis is that Instruments are not syntactically realized unless they are informative with respect to the verb, i.e. when they are not recoverable from the verb alone. Indeed, a generalization can be made about the conditions that allow the production of inst-PP with shadow- and default-verbs, which do not exclude each other. Inst-PP can be syntactically realized:

- i. with shadow-verbs if they are modified;
- ii. with shadow- and default-verbs when they are atypical, i.e. they do not realize the (class of) Instrument(s) selected by the verb;
- iii. with shadow- and default-verbs when the context is pragmatically marked, e.g. with irony, contrastive focus, etc.
- iv. with default-verbs when one wants to mention a single element of the entire class selected by the verb (with the Instrument not being shadowed by the verb+internal complement).

The circumstances in (i-iv) decrease the recoverability of Instruments making their syntactic realization necessary. Let us now see some examples taken from the corpora (32):

- (32) a. *Sai la sera che mi ero **avvelenata con la pizza marcita*** (KIP, convers. id: BOA3004)  
 ‘You know, the evening when I poisoned myself with the rotten pizza’  
*Avvelenare* ‘to poison’ – Shadow (INST: {*veleno* ‘poison’})
- b. *Ha **murato la porta con assi e chiodi*** (KIP, convers. id: TOD2002)  
 ‘He bricked up the door with boards and nails’  
*Murare* ‘to wall/brick up’ – Shadow (INST: {*muro* ‘wall’})
- c. *Si che puoi **scrivere con la mano*** (KIP, convers. id: TOA1005)  
 ‘Yes, of course you can write with your hand’  
*Scrivere* ‘to write’ – Default
- d. *Insomma **giocavamo non so con mh ...***  
*Con i pattini con la bicicletta* (KIP, convers. id: TOD2012)  
 ‘Well, [we] played with I don’t know mh ... With our skates with our bicycle’  
*Giocare* ‘to play’ – Open

e. A: *Durante il cammino sentivam **delle frasi** che ci colpivano*

(KIP, convers. id: BOD2010)

‘During the walk we heard some sentences that impressed us’

B: ***ci=creavamo** *delle canzoni**

*with\_them=create.IMPF.1PL some songs’*

‘We used to create some songs with them’

*Creare* ‘to create’ – Open

In (32a), the inst-PP *con la pizza* ‘with the pizza’ co-occurs with a shadow-verb: it is further modified by the adjective *marcia* ‘rotten’; it is also atypical, i.e. it does not realize the Shadow Instrument selected by the verb, i.e. *veleno* ‘poison’. Moreover, the context is pragmatically marked with irony. In (32b), the inst-PP *con assi e chiodi* ‘with boards and nails’ co-occurs with a shadow-verb (*murare* ‘to brick up’) and, again, it is atypical, i.e. it is not *con il muro* ‘with the wall’, the Instrument selected by the verb. In (32c), the inst-PP *con la mano* ‘with your hand’ co-occurs with a default-verb: it realizes an atypical Instrument because one cannot usually write with their hand, and the context is pragmatically marked: the speaker is correcting a wrong presupposition of the interlocutor. In (32d) and (32e), the inst-PP and the inst-cl, respectively, co-occur with open-verbs and do not need to be modified, be atypical or occur in a pragmatically marked context, since they are unretrievable from the verb meaning alone.

The examples in (32a), (32b) and (32c) are representative of many instances of co-occurrences of inst-PP and shadow- or default-verbs found in the corpora. In Table 4, the number of modified (i), atypical (ii) or single elements of the class (iv) inst-PPs in co-occurrence with shadow- and default-verbs are reported, in order to provide a more fine-grained picture of the distribution of inst-PPs in co-occurrence with such verbs. Open-verbs are left out since no instances of circumstances (i-iv) are found for them.

Table 4. Number of atypical, modified, “single element of the class” inst-PPs for shadow-, default-verbs, all corpora.

Verb type	Modified INST	Atypical INST	Single element of the selected class (only for Default)	Proper Shadow/ Default INST	Total number of inst-PPs
<i>Shadow-verbs</i> (7)	2 (+ 1 out of 8 atypical INST are additionally modified)	8	---	1	11
<i>Default-verbs</i> (16)	(2 out of 19 single elements are additionally modified)	14	19	2	35

As can be seen in Table 4, the only proper SI which is pronounced is *lingua* ‘tongue’ for the verb *leccare* ‘to lick’: the sentence is about a cat which licks her kittens to clean them (as opposed to the necessity for the owner of the cat to clean them). Therefore, the context is slightly pragmatically marked since the speaker is correcting a wrong presupposition of their interlocutor. Consistently with our hypothesis, other inst-PPs are either atypical (8/11), e.g. for verbs like *stampare* ‘to print’, *tappare* ‘to cork’, *avvelenare* ‘to poison’, or modified (2/11), e.g. for *registrare* ‘to record’, *telefonare* ‘to phone’ (or both, e.g. 1/11 for *avvelenare* ‘to poison’).

Regarding DI, consistently with our proposal, 19/35 inst-PPs realize a single element of the class selected by the verb (and the DI is not shadowed in context), e.g. for verbs like *comprare* ‘to buy’, *pagare* ‘to pay’, *decapitare* ‘behead’; 14/35 inst-PPs are atypical, e.g. for verbs like *disegnare* ‘to draw’, *pesicare* ‘to fish’, *scrivere* ‘to write’. Only 2/35 are proper DI, in that the inst-PP expresses the class of entities selected by the verb. In one case, this happens in the sentence *Lavarla con quali prodotti?* ‘Which products can I wash it with?’: the context is pragmatically marked in that it is a question. In the other case, the sentence is *Allora aveva comprato le magliette con i soldi invece di prendere i soldi ha preso le maglie* ‘So he bought the t-shirts with the money, instead of taking the money he took the t-shirts’: here, the subject of the sentence was expected to keep the money instead of buying some t-shirts with them. Thus, the amount of money is specific and was mentioned before; moreover, the speaker is highlighting that buying the t-shirts instead of keeping the money was an unexpected and surprising action.

It is now possible to present a revised version of Table 2, i.e. a more precise mean proportion of co-occurrence with inst-PP relative to the total occurrences in the corpora for the three types of verbs, excluding atypical SI and DI, as in Table 5:

Table 5. Proportion co-occurrence with an inst-PP/total occurrences for shadow-, default-, open-verbs, all corpora, revised version.

Verb type	Number of verbs (absolute)	Co-occurrence with a proper SI, DI and OI (absolute)	Total occurrences of the verbs (absolute)	Proportion of co-occurrence/total occurrences
<i>Shadow-verbs</i>	7	1	412	0.002
<i>Default-verbs</i>	17	21	4736	0.004
<i>Open-verbs</i>	83	229	48138	0.005
<i>Total</i>	107	251	53283	0.01

The predictions that stem from our proposal are confirmed by the data collected through the corpus analysis: Instruments are rarely produced because they are easily recovered from the verb meaning ( $\pm$  internal argument;  $\pm$  pragmatic factors); for the same reason, they tend to co-occur more frequently with open-verbs than they do with shadow- and default-verbs; finally, when they co-occur with shadow- and default-verbs, they do so under specific circumstances (i-iv), all of which decrease their recoverability, making their syntactic realization more likely.

Furthermore, the contexts outlined in (i-iv) explain why the inst-cl is less produced than the inst-PP: inst-cl cannot occur in such contexts. Indeed, clitics cannot be modified (the context in (i) is thus ruled out); if the Instrument is atypical or the speaker wants to mention a single element of the class selected by the verb, the full PP needs to be spelled out (ruling out the contexts in (ii) and (iv)); clitics cannot occur in pragmatically marked contexts, e.g. they cannot be focused, etc. – the context in (iii) is ruled out as well.

## 6. Conclusions

A long debate has surrounded the argument status of Instruments over the years. Nonetheless, no straightforward conclusions have been reached. The main issue, when considering Instruments is the mismatch between semantics and syntax: indeed, Instruments that are semantically required by the verb remain syntactically optional; in turn, their optionality has been used as a proof of their being adjuncts. Our study addressed both issues, which are interrelated.

First, we applied the main semantic and syntactic diagnostics for argumenthood to Instruments in Italian, in order to determine whether they are arguments, secondary complements or adjuncts; we concluded that semantically obligatory Instruments are arguments, while semantically optional Instruments are secondary participants – entailed and selected for by the verb, albeit more weakly. Syntactically, both types of Instruments are secondary participants/quasi-arguments, selected and assigned their theta-role by the verb.

We then turned to the issue of their optionality: since Instruments are not adjuncts, their omission is not random, but principled, and it is to be treated as an instance of argument omission. We claimed that it is ruled by semantic recoverability, which depends on the amount of information a verb and its arguments provide about one another. A slightly revisited version of Pustejovsky's (1995) classification of arguments was then applied to Instruments in Italian: Shadow, Default and Open Instruments were identified. The three classes differ in the amount of information they add to the amount already provided by the verb and in their recoverability from the verb meaning, i.e. Shadow Instruments are uninformative and maximally retrievable; Open Instruments are maximally informative and minimally recoverable, while Default Instruments (in isolation) occupy an intermediate position.

Afterwards, we linked semantic recoverability of these classes with the omission of Instruments, claiming that the more recoverable an Instrument is, the more it is syntactically dropped: Shadow Instruments are maximally omitted, Open Instruments are maximally likely to be produced, while Default Instruments behave more like Shadow ones in context – in that they tend to be shadowed by the co-composition of the verb and its internal argument.

Finally, we analysed three corpora of spoken Italian, in order to test our predictions about the pattern of production/omission of Instruments; our data showed that Instruments are rarely produced as both inst-PP and inst-cl (with the latter being produced less frequently than the former); that they tend to co-occur with verbs that entail Open Instruments and that they co-occur with verbs that entail Shadow and Default Instruments only under very specific circumstances, i.e. when they are not recoverable from the verb meaning. The results of the corpus analysis are consistent with our predictions.

To sum up, our proposal organically accounts for both the debated argument status of Instruments and the mismatch between semantics and syntax, providing a solid explanation for its syntactic omission which is consistent with the real linguistic datum.

## Appendix

### Acknowledgments

We would like to thank two reviewers for their comments on an earlier version of the paper. We would also like to thank the audience of the conference “Zadar Linguistic Forum 2022 - Predicate: from polysemy to arguments” for the insightful comments they provided to our talk, on which this paper is based.

### Abbreviations

#	semantically anomalous sentence
*	completely ungrammatical sentence
?/??	sentences whose acceptability/grammaticality is degraded, at different degrees
DI	default instrument
INST	instrument
inst-cl	instrumental clitic
inst-PP	instrumental PP
OI	open instrument
SI	shadow instrument

### Complete list of shadow-, default- and open-verbs found in KIP, ParlaTO and LIP

If nothing is added, the verb only co-occurs with an inst-PP; otherwise, it is specified if it co-occurs /with both inst-PPs and inst-cl or with the inst-cl only.

Verb	Type	Co-occurrence with an inst-PP/inst-cl	Total occurrence
Avvelenare 'to poison'	Shadow	2	2
Leccare 'to lick'	Shadow	1	8
Murare 'to brick/wall up'	Shadow	1	4
Registrare 'to record'	Shadow	1	119
Stampare 'to print'	Shadow	3	91
Telefonare 'to phone'	Shadow	1	181
Tappare 'to cork'	Shadow	2	7
Asciugare 'to dry/wipe'	Default	1	7
Cacciare 'to hunt'	Default	1	9
Comprare 'to buy'	Default	6	304
Coprire 'to cover'	Default	3	35
Dipingere 'to paint'	Default	1	7
Disegnare 'to draw'	Default	2	51
Lavare 'to wash'	Default	3	149
Mangiare 'to eat'	Default	3	506
Osservare 'to watch/inspect'	Default	1	24
Pagare 'to pay'	Default	4	460
Parlare 'to speak/communicate'	Default	1	2034
Pescare 'to fish'	Default	2	32
Ragionare 'to think/discuss'	Default	1	26
Richiamare 'to call again/to attract'	Default	1	57
Scrivere (both inst-PP and inst-cl) 'to write'	Default	4 (3 inst-PP + 1 inst-cl)	847
Suonare 'to play an instrument, music, ...'	Default	1	187
Decapitare 'to behead'	Default	1	1
Accompagnare 'to escort'	Open	1	52
Accontentare 'to please'	Open	1	1

Affrontare 'to face'	Open	1	51
Aggiungere 'to add'	Open	2	65
Ammazzare 'to kill violently'	Open	4	60
Andare 'to go'	Open	23	7385
Angosciarsi 'to agonize'	Open	1	3
Annunciare 'to announce'	Open	1	8
Aprire 'to open'	Open	3	239
Arricchirsi 'to get rich'	Open	4	13
Arrivare 'to arrive'	Open	10	1253
Articolare 'to enunciate'	Open	1	8
Ascoltare 'to listen'	Open	2	234
Attaccare 'to attack'	Open	1	58
Attirare 'to attract/lure'	Open	1	11
Attrezzare 'to equip'	Open	1	9
Aumentare 'to increase'	Open	1	45
Battere 'to beat up'	Open	1	52
Campare (both inst-PP and inst-cl) 'to live + to economically sustain oneself'	Open	4 (1 inst-PP + 3 inst-cl)	11
Chiudere 'to close'	Open	2	231
Cogliere 'to hit'	Open	1	10
Colpire 'to hit'	Open	3	48
Combattere 'to fight'	Open	1	17
Comunicare 'to communicate'	Open	1	37
Consolare 'to cheer up'	Open	1	2
Contribuire 'to contribute'	Open	1	14
Correggere 'to correct/adjust'	Open	1	25
Costruire (both inst-PP and inst-cl) 'to build'	Open	2 (1 inst-PP + 1 inst-cl)	66
Creare (inst-cl only) 'to create'	Open	2	158
Dimostrare 'to demonstrate'	Open	2	49
Dire 'to say/to communicate'	Open	2	6325
Distuggere 'to destroy'	Open	1	24
Divertirsi 'to enjoy oneself'	Open	1	55
Entrare 'to enter'	Open	4	469

Fare (both inst-PP and inst-cl) 'to do' (this verb has a very general meaning, which is further specified by its internal complement)	Open+	60 (44 inst-PP + 16 inst-cl)	13613
Fissare 'to secure/attach/fasten'	Open	1	28
Formarsi 'to form'	Open	1	22
Giocare (both inst-PP and inst-cl) 'to play'	Open	26 (5 inst-PP + 21 inst-cl)	332
Girare 'to wander'	Open	4	312
Guadagnare (both inst-PP and inst-cl)	Open	17 (1 inst-PP + 16 inst-cl)	75
Guardare 'to look at/to watch'	Open	4	860
Imparare 'to learn'	Open	1	232
Inviare 'to send'	Open	1	34
Isolarsi 'to isolate oneself'	Open	1	20
Lavorare 'to work/to do a job'	Open	1	1283
Macchiare 'to stain'	Open	1	4
Mandare in paranoia 'to fester into paranoia'	Open	1	1
Martirizzare 'to martyr'	Open	1	1
Mettere 'to put'/metterci 'to take time'	Open	7	1725
Minacciare 'to threaten'	Open	1	3
Misurare 'to measure'	Open	1	10
Modificare 'to change/modify'	Open	1	32
Muoversi 'to move/to go somewhere'	Open	9	86
Partire 'to leave'	Open	3	554
Passare 'to go through some place'	Open	3	816
Pigliare (both inst-PP and inst-cl) 'to take' (informal)	Open	2 (1 inst-PP + 1 inst-cl)	106
Portare 'to bring/to carry'	Open	2	1027
Prendere 'to take'	Open	8	1074
Provare 'to try'	Open	1	435
Rafforzare 'to strengthen'	Open	1	4
Rapire 'to kidnap/abduct'	Open	1	6
Reggere 'to bear/carry/handle'	Open	2	4
Rendere 'to change a thing'	Open	1	94
Riempire 'to fill'	Open	1	39

Rientrare 'to enter again'	Open	1	62
Sbloccare 'to unlock'	Open	1	6
Scendere 'to go down'	Open	2	134
Seguire 'to follow'	Open	1	115
Sistemare 'to fix'	Open	1	37
Spalmare 'to spread'	Open	1	2
Spostarsi 'to move'	Open	3	281
Svuotare 'to empty'	Open	1	9
Toccare 'to touch'	Open	2	69
Tormentare 'to torment'	Open	1	1
Tornare 'to go back'	Open	1	585
Trastullarsi 'to enjoy oneself'	Open	1	1
Trattenere 'to hold/restrain'	Open	1	12
Trovare (both inst-PP and inst-cl) 'to find'	Open	3 (2 inst-PP + 1 inst-cl)	981
Uccidere 'to kill'	Open	2	41
Urtare 'to unwillingly touch/to irritate'	Open	1	3
Uscire 'to get out'	Open	2	592
Vedere 'to see/to watch'	Open	2	2238
Venire 'to come'	Open	4	1890
Vincere (inst-cl only) 'to win'	Open	3	1371
Vivere (both inst-PP and inst-cl) 'to live/here: to make a living'	Open	6	1317

## References

- Alfano, I., Cutugno, F., De Rosa, A., Iacobini, C., Savy, R. & Voghera, M. (2014). VOLIP: a corpus of spoken Italian and a virtuous example of reuse of linguistic resources. In: Calzolari, N., Choukri, K., Declerck, T., Loftsson, H., Maegaard, B., Mariani, J., Moreno, A., Odijk, J. & Piperidis, S. (eds) (2014) *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, 3897–3901. Reykjavik: European Language Resources Association (ELRA).
- Abney, S. P. (1987). *The English noun phrase in its sentential aspect*. PhD thesis, Massachusetts Institute of Technology.
- Baker, M. C. (1988). *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.
- Bresnan, J. (1982). Control and complementation. In: Bresnan, J. (ed) (1982). *The mental representation of grammatical relations*, 292–390. Cambridge (MA): MIT Press
- Brunson, B. A. (1992). *Thematic discontinuity*. PhD thesis, University of Toronto.
- Brunson, B. A. (1993). The Instrumental Role: Argument or Adjunct? *Toronto Working Papers in Linguistics*, 12(1), 13–25.
- Cappelli, G. (2022). *Implicit indefinite objects at the syntax-semantics-pragmatics interface: a probabilistic model of acceptability judgments*. “Tesi di perfezionamento”, Pisa: Scuola Normale Superiore.
- Cappelli, G. & Lenci, A. (2020). PISA: A measure of Preference In Selection of Arguments to model verb argument recoverability. In: Gurevych, I., Apidianaki, M. & Faruqui, M. (eds) *Proceedings of the Ninth Joint Conference on Lexical and Computational Semantics*, 131–136. Association for Computational Linguistics.
- Cardinaletti, A. (2015). Syntactic Effects of Cliticization. In: Kiss, T. & Alexiadou, A. (eds) *Syntax – Theory and Analysis. Volume 1*, 595–653. Berlin: De Gruyter Mouton.
- Cardinaletti, A. (2016). Sui limiti dei pronomi clitici: inventario ed estrazione. In: Ledgeway, A., Cennamo, M. & Mensching, G. (eds) *Actes du XXVIIe Congrès international de linguistique et de philologie romanes (Nancy, 15–20 juillet 2013). Section 4: Syntaxe*, 49–62. Nancy: ATILF/SLR
- Cardinaletti, A. (2019). Cliticization as Extraction: The Big DP Hypothesis Revisited, *Revista da Associação Portuguesa de Linguística*, 5, 1–16.
- Cerruti, M. & Ballarè, S. (2021). ParlaTO: corpus del parlato di Torino. *Bollettino dell’Atlante Linguistico Italiano (BALI)*, 44 (2020), 171–196.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Conklin, K., Koenig, J-P. & Mauener, G. (2004). The role of specificity in the lexical encoding of participants. *Brain and Language*, 90(1/3), 221–230.
- Croft, W. (1991). *Syntactic Categories and Grammatical Relations*. Chicago: Chicago University Press.
- Croft, W. (2001). *Radical construction grammar: syntactic theory in typological perspective*. Oxford/New York: Oxford University Press.
- DeArmond, R. & Hedberg, N. (1998). On Complements and Adjuncts. In: *Proceedings of the Canadian Linguistics Association*, 1–18.
- De Mauro, T., Mancini, F., Vedovelli, M. & Voghera, M. (1993). *Lessico di frequenza dell’italiano parlato*, Milano: Etaslibri.

- Dowty, D. (1982). Grammatical relations and Montague grammar. In: Jacobson, P. & Pullum, G. (eds) (1982). *The nature of syntactic representations*, 79–130. Dordrecht: Reidel.
- Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 67, 547–619.
- Dowty, D. (1994). Adjunct-to-argument reanalysis in a model of grammar growth. Paper presented at ESCOL '94, University of South Carolina.
- Dowty, D. (2003). The dual analysis of adjuncts/complements in Categorical Grammar. In: Lang, E., Maienborn, C. & Fabricius-Hansen, C. (eds) *Modifying adjuncts*, 33–66. Berlin: Mouton de Gruyter.
- Ernst, T. (2001). *The syntax of adjuncts*. Cambridge: Cambridge University Press.
- Fabrizio, C. (2013). The meaning of a noun converted into a verb. A semantic exploration on Italian. *Rivista di Linguistica*, 25(2), 174–219.
- Fillmore, C. (1968). The case for case. In Bach, E. & Harms, R. (1968), *Universals in Linguistic Theory*, 1–87. New York: Holt, Rinehart, and Winston.
- Foley, W. & Van Valin, R. (1984). *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Glass, L. (2020). Verbs describing routines facilitate object omission in English. In: Farrell, P. (ed) *Proceedings of the Linguistic Society of America*, 5(1), 45–58.
- Goldberg, A. E. (2002). Surface generalizations: an alternative to alternations. *Cognitive Linguistics*, 13, 327–56.
- Grimshaw, J. (1990). *Argument structure*. Cambridge (MA): MIT Press.
- Gruber, J. S. (1965). *Studies in lexical relations*. PhD thesis, Massachusetts Institute of Technology.
- Hale, K. & Keyser, S. J. (1993). On Argument Structure and the Lexical Expression of Syntactic Relations. In Hale, K. & Keyser, S. J. (eds) *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 53–110. Cambridge (MA): MIT Press.
- Haspelmath, M. (2014). Arguments and Adjuncts as Language-Particular Syntactic Categories and as Comparative Concepts, *Linguistic Discovery*, 12(2), 3–11.
- Hopper, P. J. & Thompson, S. A. (1980). Transitivity in Grammar and Discourse. *Language*, 56(2), 251–299.
- Jackendoff, R. (1977). *X-bar syntax: A study of phrase structure*. Cambridge (MA): MIT Press.
- Jackendoff, R. (1990). *Semantic structures*. Cambridge (MA): MIT Press.
- Ježek, E. (2017). Generative Lexicon Theory and Lexicography. In: Hanks, P. & De Schryver, G. M. (eds) *International Handbook of Modern Lexis and Lexicography*, 1–21. Berlin: Springer.
- Johns, A. (2017). Noun incorporation. In: Everaert, M. & van Riemsdijk, H. (eds) *The Wiley Blackwell Companion to Syntax*, 2<sup>nd</sup> edition, 1–27. Hoboken: Wiley.
- Koenig, J.-P. & Davis, A. R. (2006). The KEY to lexical semantic representation. *Journal of Linguistics*, 42, 71–108.
- Koenig, J.-P., Mauner, G. & Bienvenue, B. (2003). Arguments for adjuncts. *Cognition*, 89, 67–103.
- Koenig, J.-P., Mauner, G., Bienvenue, B. & Conklin, K. (2007). What with? The anatomy of a (Proto)-role. *Journal of Semantics*, 25, 175–220.
- Lakoff, G. (1968). Instrumental adverbs and the concept of deep structure. *Foundations of Language*, 4, 4–29.

- Langacker, R. W. (1987). *Foundations of cognitive grammar*. Stanford: Stanford University Press.
- Levin, B. (1989). *Towards a lexical organization of English verbs*. Evanston: Northwestern University.
- Levin, B. (1993). *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.
- Marantz, A. (1984). *On the nature of grammatical relations*. Cambridge (MA): MIT Press.
- Mauri, C., Ballarè, S., Gorla, E., Cerruti, M. & Suriano, F. (2019). KIParla corpus: a new resource for spoken Italian. In: Bernardi, R., Navigli, R. & Semeraro, G. (eds) *Proceedings of the 6<sup>th</sup> Italian Conference on Computational Linguistics CLiC-it*, paper 45. Associazione Italiana di Linguistica Computazionale.
- Medina, T. N. (2007). *Learning which verbs allow object omission: Verb semantic selectivity and the implicit object construction*. PhD thesis, The Johns Hopkins University.
- Miller, P. (1997). Compléments et circonstants: une distinction syntaxique ou sémantique? In: Souesme, J.-C. (ed) *Actes du 37<sup>ème</sup> Congrès de la SAES*, 91–103. Nice: Presses Universitaires de Nice.
- Ono, N. (1992). Instruments: a case study of the interface between syntax and lexical semantics. *English Linguistics/Journal of the English Linguistic Society of Japan*, 9, 196–222.
- Pollard, C. & Sag, I. (1987). *Information-based syntax and semantics*. Stanford: CSLI.
- Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge (MA): MIT Press.
- Randall, J. H. (1992). The catapult hypothesis: An approach to unlearning. In: Weissenborn, J. Goodluck, H. & Roeper, T. (eds) *Theoretical issues in language acquisition: Continuity and change in development*, 93–138. Hillsdale: Lawrence Erlbaum.
- Rappaport-Hovav, M. & Levin, B. (1988). What to do with  $\theta$ -roles. In: Wilkins, W. (ed) *Syntax and semantics. Volume 21*, 7–36. San Diego: Academic Press.
- Resnik, P. (1993). *Selection and Information: A Class-Based Approach to Lexical Relationships*. PhD thesis, University of Pennsylvania.
- Resnik, P. (1996). Selectional constraints: An information-theoretic model and its computational realization. *Cognition*, 61(1-2), 127–159.
- Rissman, L. (2013). *Event participants representations and the instrumental role: a cross-linguistic study*. PhD thesis, John Hopkins University.
- Rissman, L. & Rawlins, K. (2017). Ingredients of Instrumental Meaning. *Journal of Semantics*, 34, 507–537.
- Schlesinger, M. I. (1995). *Studies in English language. Cognitive space and linguistic case: Semantic and syntactic categories in English*. Cambridge: Cambridge University Press.
- Schütze, C. (1995). PP attachment and argumenthood. *Papers on language processing and acquisition*, 26, 95–151. Cambridge (MA): MIT Press.
- Speas, M. (1990). *Phrase Structure in Natural Language*. Dordrecht: Kluwer.
- Talmy, L. (1976). Semantic Causative Types. In: Shibatani, M. (ed) *The Grammar of Causative Constructions. Volume 6*, 43–116. New York: Academic Press.
- Talmy, L. (2000). *Toward a cognitive semantics*. Cambridge (MA): MIT Press.
- Van Hooste, K. (2018). *Instruments and Related Concepts at the Syntax-Semantics Interface*. Düsseldorf: Düsseldorf University Press.

# 11. For a cognitive continuum underlying the distinction between arguments and adjuncts

ALESSANDRO BUTTINELLI

Keywords: *argument; adjunct; subjectivity; objectivity; cognitive linguistics; linguistic continua.*

The main aim of this paper is to resystematize the traditional discrete distinction between arguments and adjuncts within the cognitive continuum of subjectivity (Langacker 1990, 2007). From this perspective, the traditional discrete distinction between argument and adjuncts can be reinterpreted in a continuistic way (Mereu 2020). To achieve this, we will qualitatively analyse the behaviour of an Italian adverbial preposition of limitation, *riguardo (a)* NP ‘regarding NP’, showing how the different roles it assumes with respect to the argument structure can be traced back to a different subjective or objective construction of the relation it establishes.

*Original scientific paper*

## 1. Introduction

The dichotomy between arguments and adjuncts is the natural outcome of the structuralist tradition. More recently, the post-structuralist literature, characterized by an emergentist and usage-based approach to grammar (Hopper 1987; Bybee 2010, among others), has attempted to overcome the classical approach in favour of a perspective that does not see the different features of language as the result of binary oppositions, but rather as focused areas within broader grammatical continua. The linguistic structures that form the speaker’s grammar would emerge as the outcome of functionally motivated but blurred distinctions.

We want to reconsider the traditional dichotomy between arguments and adjuncts, adopting an approach that takes these two categories as extreme

poles of a continuum (Mereu 2020).<sup>174</sup> More specifically, the distinction can be grasped in terms of different construals of the conceptual representation of the sentence. By construal we mean «our multifaceted capacity to conceive and portray the same situation in alternate ways» (Langacker 2007: 435).

The continuum that we believe best characterizes the distinction between arguments and adjuncts is that of subjectivity-objectivity. Central to the construal of a proposition is indeed «the extent to which a particular entity is subjectively or objectively construed» (Langacker 2007: 435). We can describe the two categories in these terms: «an entity is said to be objectively construed to the extent that it goes ‘onstage’ as an explicit, focused object of conception. An entity is subjectively construed to the extent that it remains ‘offstage’ as an implicit, unselfconscious subject of conception» (Langacker 2006: 18). We will thus use the subjectivity-objectivity continuum to re-propose an internal distinction at the adjunct pole. This distinction is roughly equivalent to the traditional difference between sentence modifiers and predicate modifiers. In the first case (sentence modifier) it is the entire event/state of affairs presented in the sentence that is modified. See for example: *Theoretically, man can arrive on Mars*. Thus, the fact that man can arrive on Mars is only presented theoretically. In the second case it is only the predicate that is modified: *Luke eats fast*. Here it is only the action of eating performed by Luke that is modified, specifically the fact that he does it quickly. Placed on a cognitive level, the difference can be explained on a different construal of the proposition.

In order to show how the continuum of subjectivity-objectivity responds to different construals of the proposition, we analyse the behaviour of a rather marginal complement in traditional grammatical literature: the complement of limitation in Italian *riguardo (a) NP* ‘regarding NP’. The class of complements of limitation is very broad. A non-exhaustive list<sup>175</sup> of its instances is given in Table 1.

---

<sup>174</sup> See Birtić & Brač (*this volume*, Chapter 9) for a survey of tests for distinguishing between arguments and adjuncts.

<sup>175</sup> Providing a complete list of the possible realizations of the complement of limitation is beyond the scope of this paper, although it represents an interesting object of analysis that deserves a dedicated work. Already a non-exhaustive list nevertheless provides a first picture of the variety and complexity of complement’s combinations.

Table 1. A non-exhaustive list of complements of limitation in Italian

Combination <sup>176</sup>	Instance
P	<i>per; in; su; circa; secondo</i>
V	<i>considerando</i>
N + P	<i>riguardo a; rispetto a</i>
V + P	<i>parlando di</i>
Adv + P	<i>relativamente a; limitatamente a</i>
P + N + P	<i>a proposito di; in relazione a; in riferimento a; in quanto a; in merito a; in fatto di; a livello di; in tema di; in rapporto a; in connessione con; con attinenza a</i>
P + Pro + Rel + V	<i>per quanto riguarda; per quello che concerne; per ciò che si riferisce</i>

In Table 2 we also provide a non-exhaustive list of possible realisations of the complement of limitation in English. In the examples we will always propose a translation of *riguardo (a)* as ‘regarding’ thus favouring the criterion of structural similarity between the two, but the other forms are also possible translations of the limiting meaning of the complement.

Table 2. A non-exhaustive list of complements of limitation in English

Combination	Instance
P	<i>on; for; of; about</i>
V	<i>regarding; concerning; respecting; touching</i>
P + P	<i>as for; as to</i>
V + P	<i>relating to; dealing with; referring to; speaking of; according to; concerned with</i>
Adj + P	<i>relative to; germane to; relevant to</i>
P + N + P	<i>with reference to; with respect to; in reference to; in relation to; in respect of</i>

All realizations of this class of complements share the same limiting function, hence the possibility of treating them as belonging to a single, functionally homogeneous class. In this sense, *riguardo (a)* NP has not been

<sup>176</sup>The different combinations are represented here by word classes and not as syntagmatic representations. We therefore intentionally use N to denote a noun and not a noun phrase.

chosen because it is endowed with special characteristics with respect to the other elements of the class, but precisely because it is its member. Moreover, the choice of studying the behaviour of complements of limitation with respect to the distinction between arguments and adjuncts is based on the fact that this type of complement never has the status of argument, but shows a variety of realisations as different types of adjuncts, thus allowing us to complexify a category that seems by nature discrete. This does not distinguish it from other types of complements that are semantically different but with similar syntactic status. The results we will arrive at would thus theoretically be extendable to other classes of complements, but a separate work will be necessary to corroborate this hypothesis.

In general, this complement limits the propositional content to a domain it introduces:

- (1) *La Adam Air è ora al centro di polemiche riguardo allo stato di manutenzione della sua flotta.*  
‘Adam Air is now at the centre of controversy regarding the state of maintenance of its fleet.’
- (2) *Riguardo la dissipazione, la CPU sopportava temperature fino a 100 °C e non consumava oltre i 31 Watts.*<sup>177</sup>  
‘Regarding dissipation, the CPU could withstand temperatures of up to 100 °C and consumed no more than 31 Watts.’

Starting from the qualitative discussion of real examples, we will show how such a complement can place itself between and outside the margins of the sentence. We will discuss how this ambiguity is the product of interacting factors: semantics, sentence syntax and pragmatic roles (i.e. topic vs. comment). The distinction between arguments and adjuncts and between different adjuncts will thus be determined by the interface between different levels of linguistic analysis, motivated at a cognitive and functional level (Mereu 2020).

All the examples used in the analysis are collected from the PAISÀ (*Piattaforma per l’Apprendimento dell’Italiano Su corpora Annotati*) corpus of written Italian of the web (Lyding *et al.* 2014). The corpus contains a

---

<sup>177</sup> In this example it can be seen that the simple preposition *a* is not obligatory. In a process of simplification we have forms of the same complement represented only by *riguardo* NP, which are considered equally grammatical as their counterpart *riguardo a* NP. For this reason, we have represented the element *a* as an optional element in round brackets.

total of about 380,000 documents from about 1,000 separate sites with a total of about 250 million words, collected between September and October 2010. About 260,000 documents come from the Wikipedia, about 5,600 from other Wikimedia Foundation projects. About 9,300 documents come from Indymedia, and an estimated 65,000 documents come from blogs. We have chosen this corpus over other larger corpora exclusively for a practical reason: PAISÀ is freely available online at <https://www.corpusitaliano.it/> without any kind of subscription. Moreover, its size still guarantees an adequate representation of the written Italian of the web.

## 2. Subjective and objective construal

Within a cognitivist view of grammar, the meaning of a sentence is given by the way we construct the conceptual representation of the entities, states, events, and processes involved. In this sense, the meaning conveyed by linguistic expressions reflects how we see them and how we want the addressee to see them. Meaning is therefore not exclusively given by the correspondence between a linguistic and a real dimension, but rather actively constructed by speakers during the interaction. To represent this view of things, Langacker (1990) introduces the term “construal”:

«A speaker who accurately observes the spatial distribution of certain stars can describe them in many distinct fashions: as a constellation, as a cluster of stars, as specks of light in the sky, etc. Such expressions are semantically distinct; they reflect the speaker’s alternate construals of the scene, each compatible with its objectively given properties.»  
(Langacker 1990: 61)

The choices that the speaker makes when selecting a linguistic expression are reflected at all levels of the language: from inflectional morphemes (temporal and aspectual relations with respect to the communicative situation) and derivational (consider the effect elicited by diminutives with respect to their normal forms), to the choice of word class (see, for example, nominalizations of verbs), up to broader syntactic structures (e.g. cleft or passive structures or word order). The use of language, as an activity, is in fact always set in a precise spatio-temporal context and situated at the convergence between at least two individuals, each endowed with respective

knowledge and, in addition, in a continuous operation of calibration of meaning according to shared knowledge.

One of the operations of construal is that of perspective: the speaker builds the conceptual representation of the message from a certain point of view, which may be more or less subjective or objective. In the meaning proper to cognitive linguistics, these two terms must be understood as two extreme poles of the same continuum reflecting the way in which a particular entity can enter the process/event described by the sentence. A subjective entity or relation between entities will therefore be «a locus of consciousness that is not itself conceived» (Langacker 2007: 436) as part of the described action. Rather, it will be external to it, as a point of view that is on the side of the conceptualizer – which can be understood here as the speaker – of the described event, reflecting the perspective it takes on the event and on its linguistic representation. In contrast, an objective entity or relation between entities will be a part of the described event, an entity or relation involved in the scene represented by the meaning of the sentence. We then have a whole series of realizations characterized by intermediate degrees of subjectivity separating the two extreme poles of the continuum.

For the purposes of our analysis, this continuum can help to enlighten the traditional distinction between arguments and adjuncts, and within these, between different types of adjuncts, from a more cognitivist perspective. An argument, due to its distributional and structural properties, should be seen as a participant in the event presented with the highest degree of objectivity, as a necessary component of the conceptual representation. In contrast, an adjunct is usually constructed with a maximum degree of subjectivity, as an optional component that contributes to the representation by presenting unnecessary participants or relations that reflect a subjective point of view of the speaker on the event. As they are not part of the objective representation, they can be omitted.

We must not forget how subjectivity and objectivity are gradual properties, which gives us the opportunity to consider the whole series of realizations that show behaviour somewhere between argument and adjunct, such as adverbial modifiers. In order to better understand this, however, we need to introduce the different linguistic metafunctions; indeed, depending on the level of metafunction to which a linguistic element can be traced, it will show a greater or lesser degree of subjectivity.

### 3. Language metafunctions

In the perspective presented by Halliday (1970, 1994), and continued in works such as Dik *et al.* (1990) and Hengeveld (1989), the sentence is the result of three different structures each belonging to three different linguistic metafunctions. Each component then fulfils different roles in the economy of the message and expresses a different meaning with which it contributes to it. Thus, we have: an ideational level, within which the meaning of the sentence is represented as a representation of the event described, with its participants and the different circumstances that contribute to defining the event. It is thus the level in which participants, events and relations are objectively constructed. Then, we have an interpersonal level, which manages how the speaker stands in relation to the representation of events (think of marker of evidentiality or epistemicity) and the discursive dynamic between speaker and listener. This level is construed on a higher degree of subjectivity and a lesser degree of objectivity, in fact it has less to do with the representation of the message than with the perspective the speaker takes on it. Finally, there is a textual level, which handles thematic progression, how the sentence binds to the preceding and following context in the broader textual and discursive organisation. This level possesses the highest degree of subjectivity, as it exclusively reflects the speaker's point of view for his communicative purposes, it is thus entirely determined by the perspective taken by the speaker.

The distinction between arguments and adjuncts can then be appropriately interpreted if we take this perspective: depending on the different level, we have arguments or adjuncts. Arguments are in fact part of the ideational level, they are necessary participants in the represented event. Some adjuncts can also be part of this level, namely predicate and predication modifiers (Dik *et al.* 1990), whose role for the former is to add participants or relations that characterize and enrich the conceptual representation, while for the latter it is to specify the spatial, temporal, or conceptual coordinates within which the event takes place. At a higher level of subjectivity and meta-functionality, on the other hand, we find additions that do not directly participate in the event but indicate an attitude or state of the speaker towards what is being represented, or that frame the event in an interpretative frame of reference. Finally, at the textual level we usually find discursive signals such as connectives or topicalization markers. Different syntactic realizations thus pertain to different levels.

For the purposes of our analysis, we will therefore resort to a particular category, that of adverbial modifier, which better than others shows different behaviours depending on the level of metafunctions it modifies. For this reason, it constitutes an interesting observation point from which we can test our interpretation in terms of the subjectivity and objectivity of the construal. We shall do this by observing the behaviour of a *riguardo (a)* NP as an adverb.

#### 4. Adverbial modifiers

It is necessary now to better understand to which syntactic category *riguardo (a)* NP belongs to. Once this has been clarified, it will be easier to explain its behaviour. The definition of *riguardo (a)* NP in terms of discrete syntactic categories is indeed problematic. In fact, it is clearly a prepositional phrase, characterized, as usual, by a high cohesion between the prepositional element and the nominal one, such as not to allow the insertion of other linguistic material between the two. In terms of a preposition, it will be more appropriate to describe *riguardo (a)* as a complex preposition composed by a nominal element (*riguardo* ‘regarding’) and a simple preposition (*a*). It should be noted here that the simple preposition *a* may not appear in some cases and the preposition is limited to its nominal element.

However, once we move on to consider the function of this complex preposition in the economy of the sentence, we can realize how its function is that of an adverbial modifier. Indeed, it modifies the propositional content and truth value of the sentence, limiting them in relation to the dimension represented by the NP introduced by the preposition. An adequate paraphrase of *riguardo (a)* NP may in fact be *limitatamente a* NP ‘limitedly to NP’. Adverbs, indeed, constitute a structurally heterogeneous class, characterized by a typically prototypical configuration (Croft 1998), endowed with its own centre, represented by the most frequent and representative cases of the class (such as adverbs in *-mente* for Italian) and a whole constellation of more marginal cases, both in frequency and representativeness, also consisting of other word classes that may have modifier functions.

In this sense, *riguardo (a)* NP as an adverbial modifier is certainly a non-prototypical case of an adverb. In contrast to prototypical adverbs, the complex preposition cannot appear on its own (*\*riguardo a*), the presence of the NP is necessary. In this sense, it can be seen as a case of argument-

taking adverb (Ramat & Ricca 1998: 188). *Riguardo (a)* works as the head of the prepositional phrase, we can then define it as an adverbial preposition (Ramat & Ricca 1998: 188). As is typical of all adverbs, however, it can modify a noun, an adjective or more complex syntactic structures such as predication or the entire process (Ramat & Ricca 1998: 187):

- (3) *Il tutto è successo per colpa di una lamentela da parte di Giancarlo verso Davide riguardo ad un presunto danneggiamento della sua vettura e un divieto di posteggio nel cortile del condominio.*  
‘The whole thing happened because of a complaint by Giancarlo to Davide regarding alleged damage to his car and a parking ban in the condominium’s courtyard.’
- (4) *Le applicazioni più elastiche riguardo alla perdita dei dati e dipendenti dal tempo si affidano invece a UDP.*  
‘Applications that are more elastic regarding data loss and time-dependent rely instead on UDP.’
- (5) *Non sappiamo nulla riguardo l’aspetto di questa prima chiesa né del primo nucleo del convento.*  
‘We know nothing about the appearance of this first church or the first nucleus of the monastery.’
- (6) *Riguardo a tale opera è sorta una disputa.*  
‘Regarding this work a dispute has arisen.’

Having clarified the category to which *riguardo (a)* NP pertains, we must now take a closer look at how this adverbial preposition behaves with respect to the distinction between arguments and adjuncts of the argumentative structure. In fact, in general, adverbs are considered as adjuncts, by virtue of the freedom of movement and on the consideration that they constitute an optional part of the sentence, which is not required to have a grammatical sentence, but merely enriches its meaning, modifying it. They can therefore be omitted without affecting the grammaticality of the resulting sentence. This is also true with *riguardo (a)* NP, which thus behaves like most adverbs in this respect:

- (7) *Riguardo all’aspetto lirico, il Midwest rap è dominata dall’arte di concetto.*  
‘Regarding the lyrical aspect, Midwest rap is dominated by concept art.’
- (8) *Hynek era scettico riguardo molte testimonianze.*  
‘Hynek was sceptical regarding many testimonies.’

In (7) and (8), the sentences resulting from the omission of the adverbial preposition are syntactically correct, while being semantically more generic: the propositional content and truth value of the sentence is no longer interpreted limited to one dimension but applies more generally.

Although not a prototypical case of an adverb, we can then observe how the adverbial preposition under investigation has the typical functions of an adverb, as well as its behaviour as an adjunct. This can also be observed from another property typically assigned to adverbials and adjuncts: the possibility of occupying different positions within the syntagmatic continuity of the sentence:

- (9) *Riguardo il titolo dell'album ci fu una polemica.*  
 'Regarding the title of the album there was a controversy.'
- (10) *Un genitore indignato riguardo al cartone ha scritto una lettera al Washington Post.*  
 'A parent outraged about the cartoon wrote a letter to the Washington Post.'
- (11) *In molte transazioni, una parte ha maggiori informazioni riguardo al prodotto, rispetto all' altra parte.*  
 'In many transactions, one party has more information about the product than the other party.'
- (12) *Va però segnalato che diverse volte in passato si sono diffuse notizie infondate riguardo la sua morte.*  
 'It should be noted, however, that several times in the past there were unfounded reports about his death.'

However, the apparent freedom of movement with which the adverbial preposition is endowed is not the result of a free choice of the speaker but depends on several factors. These, as we will see in section 5, include the level at which the adverbial is modified.

Before proceeding, it is first necessary to provide a terminological clarification of the use of the term adjunct and the term modifier. In this paper the former is used in its traditional sense, as summarized by Lyons (1968: 334): «adjuncts (of place, time, manner, reason, etc.) are optional, or structurally dispensable, constituents of the sentence: they may be removed without affecting the remainder of the sentence». For the syntactic tradition, an adverbial modifier is an adjunct. For the cognitivist tradition, the term adjunct is reserved for those syntactic structures that do not further elaborate

the sentence structure by completing or modifying it (Langacker 1991: 430). This entails a potential ambiguity with regard to our use of the term adjunct and the term modifier. Since the aim of this paper is primarily to describe the traditional discrete distinction between argument and adjunct as a continuum, we are interested here in maintaining the structuralist meaning of the term adjunct. In contrast to the cognitivist position, adverbial modifiers here are considered adjuncts, although recognising the inadequacy of the term to describe a wide range of phenomena united only by their optionality.

## 5. Tests for subjectivity and objectivity

In this part of the paper, we will resort to a series of tests to show how the adverbial preposition *riguardo (a)* can be characterized according to different degrees of subjectivity or objectivity, depending on the level of linguistic metafunctions it contributes to modify. We will resort to a series of tests of a different nature, both more strictly syntactic and also semantic and pragmatic, as we believe that the behaviour of a linguistic expression can only be explained starting from a holistic approach, which sees the different levels of linguistic analysis not in isolation but in their natural interaction, without thereby running the risk of confusing the specificities proper to each level.

In proceeding, we shall follow an increase in the subjectivity with which the limiting relation established between *riguardo (a)* NP and the modified element is constructed. At first, we shall in fact show how *riguardo (a)* NP is placed at an ideational, hence objective, level of the sentence. In such cases it can modify a noun, an adjective or the predication. It acts as a modifier of predication when it modifies the state of affairs represented by the predication (consisting of predicate followed by its possible arguments) with respect to a spatial, temporal or conceptual dimension (Dik *et al.* 1990: 176). In general, these types of modifiers «specify the setting within which an SoA occurs» (Dik *et al.* 1990: 176), but they can be characterized according to the type of setting they impose. Among the most frequent and widespread are certainly temporal and spatial settings, while *riguardo (a)* NP imposes a setting that we can define as conceptual, presenting the dimension (represented by the NP) limited to which the predication is to be interpreted as relating.

Starting from a series of examples collected from the Paisà corpus, we will show how these uses must be kept distinct from more subjective constructions of the relation established by *riguardo (a)* between the introduced NP and the broader process depicted by the sentence. In these cases, *riguardo (a)* NP must be seen as an adverbial modifier of the entire process, establishing a domain within which it is to be interpreted. Therefore, we call this type of adverbs domain adverbs (Bellert 1977; Bertuccelli Papi 1992; Ramat & Ricca 1998, among others). Such adverbs restrict the domain within which the sentence must be interpreted in order to be considered true. In this sense, this type of adverbs does not express a manner in which the process takes place, «but denotes an extra-linguistic domain, for which the utterance is presented as being objectively valid by the speaker» (Grübl 2018: 66). Compared to the cases of noun, adjective, or predication modifier, we find ourselves here observing relations endowed with a higher degree of subjectivity. The level of linguistic metafunctions they modify is no longer the ideational one, but the interpersonal one. They can be paraphrased by expressions such as *parlando* ‘speaking’ plus an adverb, in the case considered by *parlando limitatamente a N* ‘speaking limitedly to N’, in this sense making explicit their interpersonal metalinguistic function: *I clarify to my interlocutor the conceptual domain within which I am going to place what I am going to say*. This metalinguistic function is a subjective construal operation that clearly has nothing to do with the representation of the event itself. As we shall see, domain adverbs constitute an interesting category that displays typical behaviours of modifiers of the conceptual level, such as modifiers of predication, such as the possibility of being posited in focal structures, while not sharing others, such as the possibility of being questioned or denied (Ramat & Ricca 1998: 191). In this sense, it emerges «the difficulty of drawing sharp distinctions between different layers or, better, at the fact that the same adverb may perform different functions» (Ramat & Ricca 1998: 193).

The last class of adverbial modifiers that *riguardo (a)* NP can be a part of is that of connectives (Berretta 1984; Schiffrin 1987; Bazzanella 2006; Fischer 2006; among others), endowed with a maximum degree of subjectivity, as elements proper to the textual metafunction. Contrary to the cases seen above, they do not participate in any way in the truth-value of the sentence, but merely play a cohesive role within the text and conversation, referring back to the preceding context and giving coherence to the following one, which is segmented into a thematically coherent unit, i.e. one that speaks of

the same thing. In this sense, *riguardo (a)* NP as a connective serializes the text into subunits. With this function, adverbial modifiers are constructed with the highest degree of subjectivity, as they encode a relation that can only be on the side of the conceptualizer: it is in fact him the one who decides on the organisation of the text, reflecting his communicative goals. These elements thus make explicit precisely this operation of constructing the textual texture, without in any way modifying what is said. Given these properties, we expect that from the proposed tests *riguardo (a)* NP as a connective will show a different behaviour from that observed for its uses as an objective modifier of nouns, adjectives, and predications and as a subjective domain modifier. In particular, it will be impossible for it to be focalized, as a properly thematic element, and its omission will entail no loss for the propositional content and truth value of the sentence.

### 5.1. *Riguardo (a)* NP as an adverbial of the predication and as a domain adverb

Let us begin by distinguishing between the role of *riguardo (a)* NP as modifier of noun, adjective and predication and as modifier of domain. We mentioned earlier how in the former case, the limiting relation established by the adverbial preposition is constructed objectively, whereas in the latter case subjectively. This distinction in the perspective with which the relation is constructed leads to different results when placed under negation. In fact, only what is part of the objective representation of the sentence can be negated, whereas what is outside this cannot be. In this regard, see the following examples:

(13) *Non si sa nulla di Lord Millenial riguardo alla sua vita prima di diventare apprendista.*

‘Nothing is known about Lord Millenial regarding his life before he became an apprentice.’

(14) *Riguardo l’anno esatto della fondazione non c’è certezza, ma tradizionalmente la si fa risalire al 1229.*

‘Regarding the exact year of the founding, there is no certainty, but traditionally it is dated to 1229.’

In (13), the modification of predication falls under negation, the possibility of limiting the positive content of the sentence (knowing something) in relation to the dimension introduced by the argument taking

adverbial is in fact represented as negative. On the contrary, in (14), the conceptual domain represented by *riguardo (a)* NP is outside the scope of negation, which merely negates the predicate of the sentence. It is only this that is negated, and not the relation established by the adverb, is shown by the fact that the following context, itself framed by a domain adverb as *tradizionalmente* ‘traditionally’, provides us with new positive information about the domain presented by *riguardo (a)*.

The same behaviour is shown by the two different modifiers with respect to the possibility of being questioned. In this sense, *riguardo (a)* NP, when objectively construed, can be made the object of questioning, on the contrary when it is subjectively construed, it necessarily remains outside it. Consider the following examples:

(15) *Riguardo l'Italia, il partito della Margherita aveva aderito ed attualmente è confluito nel PD?*

‘Regarding Italy, did the Margherita party join and is it now part of the PD?’

(16) *Stein iniziò a investigare riguardo alla struttura delle proteine?*

‘Did Stein begin to investigate about the structure of proteins?’

In (15), *riguardo (a)* NP used as domain adverbial is not affected by the question, it is not the domain relation that is questioned. One could therefore not answer this question by stating: *no, not regarding N*. On the contrary, when *riguardo (a)* NP is objectively constructed as in (16), it falls under the question. An answer, this time appropriate, to the question could then be: *no, not regarding the protein structure*.

The observed difference for questions and negation is due once again to the difference observed above between subjective and objective construction of an entity or relation. When we have a subjective construction, it will necessarily also be presupposed, as it does not reflect a state of affairs, but an attitude that the speaker takes with respect to what is being said. Conversely, when it is objectively constructed, it constitutes asserted information, which is part of the information that is not shared between speakers and can therefore be denied or tested through questions. Thus, while the assertion can be denied or inquired into, the presupposition cannot be; not being part of the objective representation of what is said, it necessarily remains outside of denial or question.

This difference between presupposed and asserted information is reflected in the assignment of pragmatic roles that can be assigned to the NP

introduced by the adverbial preposition. As a domain adverb and in initial position, *riguardo (a)* NP marks the NP as topic (i.e. a point of reference for the message, Langacker 1993: 25). The topic is an extrasentential category and is part of the subjective construction of the message. In final position and as a modifier of predication, *riguardo (a)* NP is part of the focus (i.e. unshared information conveyed by the message, Lambrecht 1994: 206). Focus is a sentential category and thus part of the objective construct of the message:

- (17) *Riguardo a questa Cappella, dalle scarse notizie che rimangono possiamo dir che ha avuto un'esistenza travagliata. Nel 1797 e nel 1799 è stata chiusa al culto ed è servita per le adunanze consiliari: è stata la prima sede del municipio. In seguito si era cercato di adibirla a uso di forno, finché è stata riconsacrata e benedetta il 23 aprile 1840. Infine dopo tante vicende è stata demolita nel 1916 e il quadro del Santo Protettore è stato destinato all' Oratorio.*

'Regarding this Chapel, from what little is known we can say that [it] had a troubled existence. In 1797 and 1799 [it] was closed to worship and used for council meetings: [it] was the first seat of the municipality. Later, attempts were made to use it as a bakery, until [it] was re-consecrated and blessed on 23 April 1840. Finally, after many vicissitudes, [it] was demolished in 1916 and the painting of the patron saint was used as an oratory.'

- (18) *I Padri conciliari poi ordinarono una citazione perentoria contro i francesi riguardo la Sanzione Pragmatica. Quest'ultima fu solennemente revocata e condannata e venne approvato il concordato con Francesco I di Francia.*

'The Council Fathers then ordered a peremptory summons against the French regarding the Pragmatic Sanction. This last one was solemnly revoked and condemned and the concordat with Francis I of France.'

In (17), the NP *questa Cappella* 'this Chapel' is taken up in subsequent sentences by a series of high continuity coreferential pronouns: zero anaphora pronouns. In contrast, in (18), the name *la Sanzione Pragmatica* 'the Pragmatic Sanction' is not retrieved by a coreferential anaphora in the sentence in which it appears but can be retrieved by a full lexical element (*this last one*) in the following context.

Another aspect reflecting the difference in the subjective and objective construction of *riguardo (a)* NP, is the noticeable difference in the extent to

which the adverbial preposition modifies the domain in one case and a noun, adjective or predication in the other. In the former, since it is subjectively construed, the scope of the modification can be extensive, stretching over several sentences placed in sequence:

- (19) *Riguardo a Fisker, gran macchina la Karma e spero che il nuovo progetto sia così entusiasmante... chissà che non sia la vera rivelazione e magari quando il petrolio finisce, questi fanno le scarpe a tutti i marchi più conosciuti.*

‘Regarding Fisker, the Karma is a great car and I hope the new project is as exciting as it is... who knows, maybe it will be the real revelation and maybe when the oil runs out, they will outdo all the better-known brands.’

If, on the other hand, we consider cases in which *riguardo(a)* NP modifies the ideational level of the language and is thus objectively constructed, the scope of the modification cannot extend beyond the limited boundaries of the element being modified.

- (20) [*Gli zaibatsu*] *Erano largamente considerati benefici e l’opinione pubblica giapponese oscillava tra il disinteresse e la disapprovazione riguardo a un’eliminazione degli zaibatsu. Attualmente gli antichi sistemi di controllo amministrativo e finanziario sono stati distrutti.*

‘The zaibatsu were widely regarded as beneficial and Japanese public opinion swung between disinterest and disapproval regarding an elimination of the zaibatsu. At present the ancient systems of administrative and financial control have been destroyed.’

As can be seen, in (20) *riguardo(a)* NP modifies only the predication within which it is found, as it is part of it. The subsequent context is in no way modified by the adverbial preposition. In this sense, as an objective part of the representation, it cannot be extended outside it, while when subjectively construed, it can do it.

A final useful test for the distinction between predication modifiers, which pertain to the ideational level, and domain modifiers, which pertain instead to the interpersonal level, sees the possibility for the former to be moved into an independent sentence as modifiers of the predicate *farlo* ‘to do it’ (Prandi & De Santis 2006; Dik *et al.* 1990; Hengeveld 1989). This can in fact be used as a verbal proforma, to indicate an action in its entirety that

can still be modified internally by adjuncts. In contrast to this, a verb such as *accadere* ‘to happen’ denotes a closed process, completed within it, which can only be modified by elements external to it, which therefore do not participate in its conceptual representation and can therefore be described as subjective. Thus, *riguardo (a) NP* as a domain modifier can be moved into an independent sentence with *accadere*, whereas it cannot be so with a predicate such as *farlo*. Conversely, *riguardo (a) NP* as a predicate modifier can be moved into an independent sentence with *farlo*, but not with *accadere*, which describes a closed process that cannot be modified from within. As expected, the adverbial preposition behaves differently when modifying predication from within and when modifying the domain from outside. In (21a) and (22a) we provide the examples in their original form from the corpus, while in (21b) and (22b) we give the modified version with *farlo*, and finally in (21c) and (22c) the modified version with *accadere*:

- (21) a. [Adolphe Quételet] *Ha influito sul sistema di Émile Durkheim riguardo l’affermazione che i fenomeni sociali hanno una natura non riconducibile ai singoli elementi di cui sono composti.*  
 ‘Adolphe Quételet influenced Émile Durkheim’s system regarding the assertion that social phenomena have a nature that cannot be traced back to the individual elements of which they are composed.’
- b. [Adolphe Quételet] *Ha influito sul sistema di Émile Durkheim. E lo ha fatto riguardo l’affermazione che i fenomeni sociali hanno una natura non riconducibile ai singoli elementi di cui sono composti.*  
 ‘Adolphe Quételet influenced Émile Durkheim’s system. And he did regarding the assertion that social phenomena have a nature that cannot be traced back to the individual elements of which they are composed.’
- c. [Adolphe Quételet] *Ha influito sul sistema di Émile Durkheim.? E ciò è accaduto riguardo l’affermazione che i fenomeni sociali hanno una natura non riconducibile ai singoli elementi di cui sono composti.*  
 ‘Adolphe Quételet influenced Émile Durkheim’s system. ?And this happened regarding the assertion that social phenomena have a nature that cannot be traced back to the individual elements of which they are composed.’

- (22) a. *Riguardo l'attività cinematografica, [Emma Danieli] è ricordata per la sua interpretazione nel film di fantascienza-horror "L'ultimo uomo della Terra".*  
 'Regarding film activity, [Emma Danieli] is remembered for her performance in the science fiction-horror film "The Last Man on Earth".
- b. *[Emma Danieli] è ricordata per la sua interpretazione nel film di fantascienza-horror "L'ultimo uomo della Terra". E ciò accade riguardo l'attività cinematografica.*  
 '[Emma Danieli] is remembered for her performance in the science fiction-horror film "The Last Man on Earth". And this happens regarding film activity.
- c. *[Emma Danieli] è ricordata per la sua interpretazione nel film di fantascienza-horror "L'ultimo uomo della Terra". ?E lo fa riguardo l'attività cinematografica.*  
 '[Emma Danieli] is remembered for her performance in the science fiction-horror film "The Last Man on Earth". ?And it does regarding film activity.'

In this section we have brought a good deal of evidence in favour of the different functions the adverbial preposition can fulfil according to the different degree of subjectivity with which the speaker constructs it. We can now move on to discuss the differences between these realizations and those with the highest degree of subjectivity as connectives.

## 5.2. *Riguardo (a)* NP as a connective

In section 3, we argued how the textual metafunction is the one endowed with the highest degree of subjectivity, in that it reflects the speaker's intentions and purposes independently of the specific content of what the speaker wants to say. In this sense, it does not participate in any way in the truth-value and propositional content of the sentence, in contrast to what *riguardo (a)* NP does, albeit with differences, at the conceptual and interpersonal levels.

In order to demonstrate this difference, we can first present the results of a test for the omission of the adverbial preposition, observing whether there are any differences in meaning and truth value with respect to the original sentence:

- (23) *Riguardo allo scaricare, io sono un vero purista per quanto riguarda i film che voglio proprio vedere.*  
 ‘Regarding downloading, I am a real purist when it comes to films I really want to see.’
- (24) *Riguardo le poste italiane, non sono affatto d’accordo.*  
 ‘Regarding the Italian postal service, I do not agree at all.’
- (25) *Sono state istituite numerose cattedre riguardo questa materia.*  
 ‘Numerous chairs have been established regarding this subject.’

Earlier, we were able to observe how omission without affecting the grammaticality of the sentence is a typical property of adverbs as adjuncts. This is confirmed again by this set of examples. What results, however, when the omission test is applied, is a difference in meaning with respect to the original sentence. Indeed, note the outcome to which the omission of *riguardo (a)* NP gives rise in examples (24) and (25). In these cases, compared to the original sentence, the sentence can no longer be interpreted as referring to the dimension introduced by the preposition, but must be read in general. The semantic value of limitation inherent in this expression, which modifies the truth value by limiting it to the presented dimension, is thus lost. In this sense, both as an objective modifier of the predication and as a more subjective modifier of the domain, *riguardo (a)* NP still partakes of the propositional content. Compare, however, this behaviour with the one shown in (23), where *riguardo (a)* NP is a connective. Its omission does not entail any loss at the level of propositional content, nor does it change its truth values. The relation it establishes is thus not that of semantic limitation to a dimension represented by the NP, but must be read as a purely cohesive relation between the different parts of the text. It therefore does not participate in the semantics of the sentence, but in the pragmatics and organization of the text. In this sense, it is constructed with a maximum degree of subjectivity, higher even than that with which it is constructed as a domain adverb, as demonstrated by the different outcome of the omission in the cases considered.

If the dimension of relevance proper to these realizations of *riguardo (a)* NP is pragmatic and not semantic or syntactic, we can see how its behaviour differs from its other uses by using precisely pragmatic tests. The first of these has to do with the shifting of *riguardo (a)* NP from the initial position, typically associated with a thematic element, to the final position, associated

with the focal element, and vice versa. With respect to the same examples proposed above, this is the outcome of the shift test:

- (26) *\*Io sono un vero purista per quanto riguarda i film che voglio proprio vedere riguardo allo scaricare.*  
 ‘I am a real purist when it comes to the films I really want to see regarding downloading.’
- (27) *Non sono affatto d’accordo riguardo le poste italiane.*  
 ‘I do not agree at all regarding the Italian postal service.’
- (28) *Riguardo questa materia, sono state istituite numerose cattedre.*  
 ‘Regarding this subject, numerous chairs have been set up.’

In (27) it can be observed how the shifting of the domain adverb, as seen above associated with presupposed and thematic information, to a final position does not give rise to any problems. It can thus be focused and made part of the representation of the event. In fact, as we mentioned earlier, domain adverbs can be focalized. The same applies to the predication modifier in (28), which can be thematized from an initial focus position. This relative freedom with which with *riguardo (a)* NP can move from topic to focus is evidence that it, even as a domain adverb, still has a certain degree of objectivity sufficient to be made part of the event. In contrast, *riguardo (a)* NP as a connective in (26) cannot be shifted without resulting in an inappropriate sentence. This is because when it is used as a connective, it is constructed with the highest degree of subjectivity, such that its inclusion within the objective representation of the sentence is impossible. It must therefore be understood as an exclusively pragmatic and thematic element.

In this regard, a final test of the pragmatic nature of the different uses of *riguardo (a)* NP has to do with the possibility of its appearance within a cleft structure. Such types of structures are typically seen as focusing strategies and in Italian have the schematic structure *è X che Y* ‘is X that Y’, in which element X is the focus, while Y represents presupposed information that is given about X. Assuming what we have said and seen above, we expect that *riguardo (a)* NP as a modifier of predication and as a domain modifier can be focused, whereas this would be impossible for its uses as a connective:

- (29) *\*È riguardo allo scaricare che io sono un vero purista per quanto riguarda i film che voglio proprio vedere.*  
 ‘It is regarding downloading that I am a real purist when it comes to films that I really want to see.’

- (30) *È riguardo le poste italiane che non sono affatto d'accordo.*  
 'It is regarding the Italian postal service that I do not agree at all.'
- (31) *È riguardo questa materia che sono state istituite numerose cattedre.*  
 'It is regarding this subject that numerous chairs have been established.'

The hypothesis seems to be confirmed once again by the impossibility of *riguardo (a)* NP as a connective to appear in focal position in (29), contrary to what can be observed for example in (30) and (31), in which the adverbial preposition can be focalized without giving rise to an inappropriate sentence. The reason for this divergence in usage is once again to be ascribed to the different degree of subjectivity with which the relation is construed. By placing itself at the highest level of subjectivity, *riguardo (a)* NP cannot enter an objective relation with the representation of the event, thus giving negative results to all tests that increase the objectivity of the relation, such as those of focalization.

A final test<sup>178</sup> to show the different degree of subjectivity with which *riguardo (a)* NP is constructed as a connective with respect to the other cases has to do with the possibility for the NP to be retrieved by an anaphoric pronoun when it represents an argument of the predicate:

- (32) *Riguardo le ipotesi sui metodi di esecuzione, esse sarebbero con un colpo di mazza alla nuca o per fucilazione.*  
 'Regarding the hypotheses on the methods of execution, these would be a blow with a club to the back of the head or shooting.'

In (32) *le ipotesi sui metodi di esecuzione* is retrieved by the personal pronoun *esse* which must necessarily be interpreted as coreferent of the NP introduced by *riguardo (a)*. This is only possible for the connective, whereas it does not occur in the case where it is a modifier of predication, thus constructed with a higher degree of objectivity:

- (33) *?Storici autorevoli come Thomas Hodgkin e John Bagnell Bury hanno espresso chiaramente il loro scetticismo su essa riguardo a questa identificazione.*  
 'Influential historians such as Thomas Hodgkin and John Bagnell Bury have clearly expressed their scepticism about it regarding this identification.'

---

<sup>178</sup>We thank the anonymous reviewer who suggested this additional way to observe the different behaviour of *riguardo (a)* NP depending on the degree of subjectivity with which it is constructed in the conceptual representation of the event.

It is interesting to note that in (33) the sentence is agrammatical only if we decide to consider *essa* coreferential with *questa identificazione*. In order to give rise to an appropriate reading, the pronoun will necessarily have to be interpreted as coreferential with another NP. This is because *riguardo (a)* NP it is already part of the objective representation of the sentence and therefore cannot be retrieved by means of a cataphoric pronoun. On the contrary, in cases where *riguardo (a)* is connective, anaphoric retrieval is possible precisely because there is no conflict between the anaphora and the NP introduced by *riguardo (a)*, since the latter is absent from the conceptual representation. The anaphoric pronoun serves precisely as a way of mediating the presence of the referent in the representation of the event without directly involving the connective.

The presence of the anaphoric coreferential pronoun with the NP of *riguardo (a)* constructed with a maximum degree of subjectivity is therefore only in apparent contradiction with what we have said about the status of maximum subjectivity possessed by the connective. Indeed, for the referent evoked by the NP to appear in the objective representation of the event as an argument required by the predicate, it must necessarily be duplicated. However subtle, the contradiction is thus only apparent.

## 6. Conclusions

We can now draw some initial conclusions from this attempt to redefine the discrete difference between arguments and adjuncts within a cognitive background. We began by presenting in more detail the cognitive continuum of subjectivity as an aspect of construal that can explain the observable differences in the behaviour of arguments and adjuncts and of different types of adjuncts. To better understand the modifier role that *riguardo (a)* NP plays, we introduced a distinction of the different linguistic metafunctions. This was necessary to better understand which of these dimensions our expression modifies. Drawing on the tradition inaugurated by Halliday (1970), we distinguished between an ideational, an interpersonal, and a textual metafunction, each characterized by an increasing degree of subjectivity.

We then described in more detail the expression object of our analysis (*riguardo (a)* NP). On the one hand, it can be seen as a prepositional phrase. From a functional point of view, however, it must be described as an adverb,

since its role is to modify a noun, an adjective, or more complex syntactic structures. For this reason, we have adopted the term prepositional adverb or argument-taking adverb. These background considerations of the relationship between the role of adverbial modifiers and degree of subjectivity allowed us to present a qualitative analysis of a few examples of *riguardo* (a) NP taken from the PAISÀ corpus (Lyding *et al.* 2014). Starting from a series of tests of a syntactic, semantic, and pragmatic nature, we were able to describe some instances of *riguardo* (a) NP as modifiers of predication, thus pertaining to the level of message representation. Pertaining instead to the interpersonal level are those uses of *riguardo* (a) NP as domain adverb. Finally, it was possible to observe how *riguardo* (a) NP in initial position does not even pertain to the interpersonal level, but rather to the textual level, fulfilling the function of a textual connective. The gradualness of the subjectivity continuum thus allowed us to interpret the different modifier roles played by the adverbial preposition as different kind of adjuncts characterized by different degrees of subjectivity.

## Appendix

### Acknowledgments

We would like to thank the anonymous reviewers for their comments and suggestions, which helped us to improve the previous version of the article.

### Abbreviations

Adj	adjective
Adv	adverb
NP	noun phrase
P	preposition
Pro	pronoun
Rel	relative pronoun
V	verb

## References

- Bazzanella, C. (2006). Discourse markers in Italian: towards a 'compositional' meaning. In: Fischer, K. (ed) *Approaches to discourse particles*, 449–464. Amsterdam: Elsevier.
- Bellert, I. (1977). On semantic and distributional properties of sentential adverbs. *Linguistic Inquiry*, 8, 337–351.
- Berretta, M. (1984). Connettivi testuali in italiano e pianificazione del discorso. In: Coveri, L. (ed) *Linguistica testuale. Atti del XV congresso internazionale della Società di Linguistica Italiana* (Genova–Santa Margherita Ligure, 8–10 maggio 1981), 237–254. Roma: Bulzoni.
- Bertucelli Papi, M. (1992). Determining the proposition expressed by an utterance: the role of 'domain adverbs'. *Textus*, 5, 123–140.
- Bybee, J. L. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press
- Croft, W. (1990). *Typology and universals*. Cambridge: Cambridge University Press.
- Dik, S. C., Hengeveld, K., Vester, E. & Vet, C. (1990). The hierarchical structure of the clause and the typology of adverbial satellites. In: Nuyts, J., Bolkestein, M. & Vet, C. (eds) *Layers and levels of representation in language theory*, 25–70. Amsterdam/Philadelphia: John Benjamins.
- Fischer, K. (ed) (2006). *Approaches to discourse particles*. Amsterdam: Elsevier.
- Grübl, K. (2018). On the rise of domain adverbials in Italian: the history of the -mente parlando construction. *Linguistik online*, 5, 65–86.
- Halliday, M. A. K. (1970). Functional diversity in language as seen from a consideration of modality and mood in English. *Foundations of Language*, 6, 322–365.
- Halliday, M. A. K. (1994). *An Introduction to Functional Grammar*. London: Edward Arnold.
- Hengeveld, K. (1989). Layers and operators in Functional Grammar. *Journal of Linguistics*, 25, 127–157.
- Hopper, P. J. (1987). Emergent grammar. In: Aske, J., Beery, N., Michaelis, L. & Filip, H. (eds) *Annual Meeting of the Berkeley Linguistics Society. Volume 13*, 139–157. Berkeley: Berkeley Linguistics Society.
- Lambrech, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
- Langacker, R. W. (1990). Subjectification. *Cognitive Linguistics*, 1(1), 5–38.
- Langacker, R. W. (1991). *Foundations of cognitive grammar: Descriptive application. Volume II*. Stanford: Stanford University Press.
- Langacker, R. W. (1993). Reference-point constructions. *Cognitive Linguistics*, 4(1), 1–38.
- Langacker, R. W. (2006). Subjectification, grammaticization, and conceptual archetypes. In: Athanasiadou, A., Canakis, C. & Cornillie, B. (eds). *Subjectification: Various paths to subjectivity. Volume 31*, 17–40. Berlin/New York: Mouton de Gruyter.
- Langacker, R. W. (2007). Cognitive Grammar. In: Geraaerts, D. & Cuyckens, H. (eds) *The Oxford handbook of cognitive linguistics*, 421–462. Oxford: Oxford University Press.
- Lyding, V., Stemle, E., Borghetti, C., Brunello, M., Castagnoli, S., Dell'Orletta, F. & Pirrelli, V. (2014). The PAISÀ Corpus of Italian Web Texts. In: Bildhauer, F. & Schäfer, R. (eds) *9th Web as Corpus Workshop (WaC-9)@ EACL 2014*, 36–43. Association for Computational Linguistics.

- Lyons, J. (1968). *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
- Mereu, L. (2020). *Semantica della frase*. Roma: Carocci.
- Prandi, M. & De Santis, C. (2006). *Le regole e le scelte. Introduzione alla grammatica italiana*. Torino: UTET.
- Ramat, P. & Ricca, D. (1998). Sentence adverbs in the languages of Europe. In: Van der Auwera, J. (ed) *Adverbial Constructions in the Languages of Europe*, 187–275. Berlin: Mouton de Gruyter.
- Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.