UNIVERSITÀ
DEGLI STUDI
FIRENZE

PhD in

Philology, Italian Literature, Linguistics

Curriculum: Digital Humanities

CYCLE XXXVI

COORDINATOR Prof. Francesco Bausi

From Events to Language: Categorisation and Construals of Breaking Events

---

Academic Discipline (SSD) L-LIN/01

Doctoral Candidate

Cacioli Caterina

(signature)

Supervisor

Prof. Panunzi Alessandro

(signature)

Coordinator

Prof. Bausi Francesco

(signature)

2020-2024

Table of Contents

Abstract vi
Riassunto viii
Introduction 1
Chapter 1. Theoretical background 7
1. Introduction 7
2. Categorisation 9
2.1. Classical categorisation 9
2.2 The cognitive approach 12
2.3 Basic-level categories 13
2.4 Ad hoc categories and lexicalisation 15
2.5 Towards event categories 16
3. Construal 17
3.1 Dimensions of construal 18
3.2 Construals cross-linguistically 20
4. Events 21
4.1 Event categorisation 23
4.2 Event Cognition 29
5 Conclusion 38
Chapter 2. Events and the Breaking Domain 41
1. Introduction 41
2. Events, actions, and action verbs 42
2.1 Actions and conceptualisations 43
2.2 Towards linguistic action encoding 45
2.3 Action verbs 49
3. The breaking domain 52
3.1 Break verbs in language acquisition 54
4. Linguistic studies on cutting and breaking 57
4.1 Fillmore (1970) 58
4.2 Levin (1993) and following work 60
4.3 Event structure 67
4.4 Constructional approaches 69
4.5 FrameNet 73
5. Semantic categorisation of cutting and breaking 77
5.2 Majid et al. (2007) 78

5.3 Follow-up studies 83
5.4 Categorisation of closely related languages and codability 85
5.4 Stathi (2023) 86
6. Selecting the domain for investigation 87
6.1 Operational definition of breaking events 87
6.2 Languages under study 88
6.3 Cross-linguistics considerations 91
7. Conclusions and next steps 92
Chapter 3. Cross-linguistic Equivalences and Semantic Boundaries 95
1. Introduction 95
2. Operationalize the study of semantic boundaries 96
3. Imagact 98
3.1 Local equivalence 99
4 Breaking events in Imagact 102
4.1 Locally equivalent verbs 102
5 Corpus study 111
5.1 Verb selection 112
5.2 Method and procedure 112
5.3 Annotation schema and parameters 114
5.4 Results 117
5.5 Discussion 125
6 Conclusion 127
Chapter 4. Event Categorisation and Construals 129
1. Introduction 129
2. Procedure and Data 130
2.1 Step 1: Visual inspection 130
2.2 Step 2: Linguistic inspection and selection of data 131
2.3 Step 3: Elicited descriptions 132
3. Results 135
3.1 Lexical diversity 135
3.2 Codability 138
3.3 Categorisation through semantic maps 140
4. Discussion 144
5. Conclusion 146
Chapter 5. Language Planning and Production 148
1. Introduction 148

2. Finding the action encoding window 149
3. Research questions, design, and hypothesis 150
3.1 Selection of target events 151
3.2 Experiment Design 152
3.3 Hypotheses 152
3.3.1 Hypothesis on Hand condition 153
3.3.2 Hypothesis on Tool condition 153
3.3.3 Speech planning and fixations 153
4. Building the experiment 155
4.1 Stimuli 155
4.2 Data collection 159
4.3 Protocol 160
4.4 Measures and windows of analysis 161
5. Results 163
5.1 Data coding 163
5.2 Data cleaning 163
5.3 Linguistic production and verb onsets 164
5.4 Fixations and Multilevel Mixed Models (MLM) Analysis 171
5.4.1 Early time window 172
5.4.2 Late time window 174
5.4.3 Middle time window 176
5.5 Timecourse visualisation and Divergence analysis 178
6. Discussion 185
6.1 Limitations of the study and future directions 188
7. Conclusion 188
General Conclusions 190
Appendix A 195
Appendix B 196
Appendix C 197
Appendix D 198
Appendix E 200
Appendix F 201
References 205

Abstract

Events segment the world into units of experience (Yates, Sherman & Yousif, 2023). Our ability to categorise the world into events allows us to abstract away from details and structure the information that we receive from the world into relevant and meaningful units. The ability to categorise events plays a pivotal role in our cognitive processes and how we make meaning: events, as shared experiences, form the basis of our understanding of the world. Categorisation is relevant at the linguistic level as well. Assuming the theoretical framework of Cognitive Linguistics, categorisation underlies processes of conceptualisations and is reflected in form-meaning mappings. Through construals, we express our conceptualisations.

Languages differ significantly in their lexical and syntactic material, categorising the world differently and showing construal preferences (Talmy, 2000). Thus, investigating construals is a way of tapping into our linguistic categorisation of events. Furthermore, experimental methods, such as eye tracking, can lead to an understanding of our processing of events and language at the moment they occur by showing the aspects of events that are fixated during spoken production.

This work investigates categorisation through linguistic construals of events in three languages (English, Italian, and Swedish). Specifically, we focus on causative breaking events, a category comprising events of change of state in which an object undergoes a loss of integrity and loss of functionality. This category brings together individual events that are different in their action schemas, force dynamics, spatial and temporal characteristics. These events, which have been part of our collective history for millennia and are crucial in early experiences of the world's dynamics as young children (Bowerman, 2005), show a great range of cross-linguistic diversity in their lexical realisations (Majid et al., 2007).

In three studies, we explore this topic using corpora and elicited event descriptions (written and spoken) as linguistic data, analysed from both qualitative and quantitative points of view. In the first study, we investigate category boundaries, examining fuzziness and typicality through corpora and annotating a set of parameters on occurrences of verbs that describe breaking events. In the second study, we focus on elicited written descriptions of a sample of breaking events shown in videos and analyse the construals' diversity using codability scores and semantic maps. In the third study, we test the differences in how Italian and Swedish speakers visually attend events to shed light on the areas speakers inspect before committing to a particular construal. The results of the studies show the typicality and degree of fuzziness in verb categories in English and Italian while also discussing the degree of equivalence of Italian and English pairs of verbs. Additionally, it is shown how the linguistic categorisation of breaking events changes

across the three languages under investigation through the construals used to describe them. Lastly, we found that the linguistic differences between Swedish and Italian are reflected in different fixations to event components in language planning during a production task.

Riassunto

Le nostre esperienze nel mondo, che si svolgono in maniera continuativa, possono essere segmentate in *unità esperienziali* (Yates, Sherman & Yousif, 2023), dette *eventi*, tramite processi di categorizzazione. La nostra abilità innata nel categorizzare gli eventi ci permette di organizzare le informazioni che riceviamo dal mondo in modo da costituire unità semantiche. Infatti, l'abilità nel categorizzare è fondamentale per i processi cognitivi e per l'attribuzione di significati: gli eventi, in quanto unità semantiche esperienziali, sono alla base delle nostre conoscenze. I processi di categorizzazione sono cruciali anche a livello linguistico, per il modo in cui associamo forme linguistiche a significati e a eventi nel mondo. In Linguistica Cognitiva, la categorizzazione sottende i processi di concettualizzazione e si riflette nel mapping tra forma e significato, il quale è espresso attraverso i *construal* linguistici.

Le lingue differiscono in modo significativo nella varietà lessicale e nelle costruzioni a disposizione per creare categorizzazioni di eventi. Studi linguistici e cognitive hanno dimostrato come lingue tipologicamente diverse usano diversi *construal* per codificare le informazioni, mostrando alternanze e preferenze (Talmy, 2000).

In questo lavoro si utilizza l'analisi di *construal* di eventi per esplorare la categorizzazione di eventi in ottica contrastiva. Nello specifico, si analizza la categorizzazione di eventi causativi rappresentanti una modifica materiale e funzionale dell'oggetto (*breaking events*) in inglese, italiano e svedese. Questa categoria raccoglie eventi caratterizzati da diversi schemi di azione, forze, elementi spaziali e temporali, che sono esperiti dall'uomo da millenni (Bowerman 2005) e che mostrano una grande variazione nel modo in cui vengono descritti in lingue diverse (Majid et al., 2007). La ricerca si articola in tre studi, che utilizzano dati da corpora e da elicitazioni di descrizioni di eventi (scritte e orali), analizzate con metodi statistici e combinate a metodi sperimentali cognitivi (tracciamento oculare). Il primo studio affronta il tema della categorizzazione in relazione ai confini semantici delle categorie, tra verbi di una lingua e di lingue diverse (inglese e italiano) attraverso dati da corpora. Nel secondo studio ci si concentra sulle elicitazioni di descrizioni scritte di un campione di video che mostra eventi di manipolazione di oggetti per analizzare la differenza di varità lessicale, codabilità delle azioni e risultante categorizzazione degli eventi tramite mappe semantiche confrontando i *construal* prodotti da parlanti di inglese, italiano e svedese. Nel terzo studio, tenendo in considerazione le differenze tra lingue e tra categorie emerse negli studi precedenti, si testa sperimentalmente le differenze di pianificazione di produzione linguistica tra parlanti di italiano e svedese al momento della verbalizzazione di descrizioni durante l'ispezione di eventi. I *construal* sono analizzati insieme ai movimenti oculari dei partecipanti durante l'ispezione di video rappresentanti gli eventi di interesse.

Introduction

The way we interpret the world and manage to communicate about what we see and do has sparked researchers' curiosity for centuries. As what we experience in the world is complex and rich with information, we manage to capture aspects of it that are relevant for communicating and carrying out activities. The primary mechanism that grants us this capacity is our categorisation ability: we abstract away from details and operate categorisations – at the conceptual and linguistic level – to structure the continuous flow of information from the world. A primary categorisation that we operate to confer structure to the world is that of segmenting and creating *events* – that is, units of experience (Yates, Sherman & Yousif, 2023).

This notion has become a core research subject in cognitive sciences, psychology, and linguistics. It is explored as a way of tapping into our mental organisation of the world and studying how we navigate the world by categorising experiences. What defines events and how they emerge is currently an object of research, but one of the more relevant aspects is that events are shared experiences in the world and in communities. Events of eating are common experiences for all living entities, while events of *flamenco* dancing are a category relevant for some communities only. It is likely that in the stream of events in the world, many people categorise some actions of eating as a *unit of experience*, while a less consistent part of humanity categorises dancing-type movements as flamenco. Events have a shared, intersubjective dimension that extends between universality and specificity, with an additional subjective component: individual minds make up for a degree of subjectivity in structuring reality. Language is an explicit means through which we render our knowledge about events accessible and the means we use to communicate with each other. Moreover, since events are related to the world, they have a primary role in the study of vision and how we pay attention to what we see, as much as the gaze strategies we adopt to interpret and give meaning to events.

Categorisation is a crucial cognitive process for our daily life – an ability we use to organise the world efficiently and make sense of the reality around us. It encompasses all areas of our knowledge, from cognition to language: we instinctively categorise objects, events, and experiences into mental groups based on similarities between them. With our categorisation skills, we learn new things: the first time we see an unknown object, we try to figure it out by comparing it to objects we know. If we ask others for information about the new object, we receive information about what it is similar to, in an attempt to provide an explanation.

Languages differ greatly in their syntactical, morphological, and lexical characteristics, packing meanings in a variety of ways. These differences are historically stratified and conventionalised, offering speakers pre-packed structures to use and share in an inter-subjective

system, as well as enough productive material to express a degree of subjectivity. *Construals* are cognitive operations of perspectivisation, change of focus, or selection that are reflected in language in the way we choose to express ourselves. The term *construal* refers to the way language users interpret and conceptualise information. In Cognitive Linguistics, the term emphasises the subjective nature of cognition, suggesting that there is a degree of subjectivity to the way we create meaning.

Events and their linguistic descriptions have undoubtedly attracted significant attention over the years and provided evidence of cross-linguistic shared structures, typological differences, and lexicalisation patterns (among others, Talmy 2000; Malt et al., 2008, 2014; Slobin, Ibarretxe-Antuñano, Kopecka, & Majid, 2014). In the world, events are composed of a stream of segments that together generate a recognisable situation: when we verbalise what we see, we provide a construal of our experience. Looking at this phenomenon from a linguistic perspective, rather than referent events in the world, verbs can generate multiple expectations regarding an action event (Hung, Edmonds & Reilly 2016) through *construals* since they can be associated with different actors, agents, patients, manner, results, type of themes, instruments, or locations.
Categorisation and construals are integral components of event cognition. Event cognition refers to the mental processes involved in perceiving, interpreting, categorising, and remembering events, which are studied in experimental tasks. The cognitive planning that leads to a construal of an event can be investigated to retrieve information on categorisation processes and the cognitive mediation between what we see and how we talk about it.

In this work, we aim to investigate event categorisation through linguistic construals. We adopt a multidisciplinary approach and operate through the investigation of verbs' semantics. We focus on the boundaries between verbs that describe a group of events, the levels of abstractions that emerge from lexical choices, and the planning of a construal for linguistic production. We examine these topics in three studies, focusing on a set of verbs in three languages (English, Italian, Swedish) describing *breaking events*. These events have been experienced by humans for millennia (in contrast, for example, to writing activities). They are a part of our life experiences that children learn early (Bowerman, 2005) and can be thought of as events that are shared by communities worldwide. However, unlike events of food consumption events (mentioned above), which tend to be visually similar to each other, breaking events encompass a great range of action schemas, forces applied, results and a great range of cross-linguistic diversity (Majid et al., 2007).

Research questions

The research questions addressed in this thesis are long-standing issues in linguistics and psycholinguistic research concerning the mapping between events and language, as mediated by cognition thorough conceptualisation. Cross-linguistics is the perspective often taken in tackling these issues, as the differences in languages enhance the possibilities for the study of meaning-making. In this regard, our research questions can be spelt out as follows:

1. What is the linguistic categorisation of breaking events in different languages?
2. How are breaking events described by verbs?
3. What are the boundaries between *break* verbs within a language?
4. How do we plan and utter the description of a breaking event?

On a broader perspective, we ask what the relation between categories is in the way it emerges linguistically (how actions are grouped together by the same linguistic elements) through construals of events and to what extent these categories fit into cognitive theories of categorisation – such as whether there are prototype effects or basic level categories. Finally, we ask whether cross-linguistic differences are related to the way visually attend to events: which areas speakers inspect before committing to a particular construal?

A focus of this research lies also in the methodological research questions. As a matter of fact, meaning eludes quantifications and poses problems for objective analysis. Research in semantics needs operationalizations that allow to conduct work that goes beyond introspection and competence. Therefore, one research question asked here is which are the most efficient methods for semantic investigation. To answer this question, we explore a combination of different data types, data collections and analyses.

Structure of the thesis

The thesis is articulated in 5 chapters. Chapter 1 presents the theorical framework under which the research is undertaken, and the three aspects related to cognition that the thesis focuses on: categorisation, construals and events. We introduce categorisation and construals under the framework of Cognitive Linguistics, discussing how categorisation is relevant for studies in linguistics and the link between conceptualisation of events and construals. We then introduce events and focus on event categorisation and cognition. Event categorisation is discussed with a review of the literature and the individuation of gaps in the literature that will be addressed with empirical research. Event cognition is discussed taking into consideration the phases that lead

from event apprehension to speech production. The chapter concludes with an introduction to the processes of planning and spoken production of event descriptions, focusing on how these processes can be investigated with the use of eye-trackers. Chapter 2 zooms into breaking events and discuss the aspects connected to it from a linguistic perspective. The point of view taken is that of linguistic research in a broad range of theoretical perspectives which have dealt with this category – from formal approaches to cognitive approaches. We discuss the grammar of the verb break, verb classes, event structure, FrameNet, semantic categorisation and experimental approaches. We then describe and present the languages under investigation in this work: Italian, English and Swedish. We proceed by proposing an operational definition of breaking events, to be used as a way of delimiting the area of research and provide comparable data across the languages. Chapter 3, 4, and 5 describe the empirical and experimental studies conducted, presenting their aims, method, materials, and results.

Empirical and experimental studies

This thesis proposes three studies, which are linked to each other and one the continuation of the previous.

In Study 1 (chapter 3) we investigate the semantic boundaries between a sample of semantically related verbs in English and Italian to explore the ways they partition the world and describe events. To do this, we first provide an overview of the verbs describing a set of videos depicting breaking events and extract parameters that define those events. We then use corpora occurrences and manually annotate them with the parameters of the breaking events to gather a quantitative perspective on the distribution of those parameters in the group of verbs. In this way, we set the ground to compare the two languages and asses the degree of correspondence between pairs of English and Italian verbs, as well as verbs referring to the same events within each language. This method helps bringing out within-language and across-language differences in the investigation of semantic boundaries and explore categorisation through a semantic mapping, but it requires a careful and time-consuming manual labour.

In Study 2 (chapter 4) we use the stimuli from the previous study for a written event description task. Given the type of task, it was possible to carry out this study with three languages – we collect data from English, Italian, and Swedish participants to focus on their construals of the events from a qualitative and quantitative point of view. We explore lexical diversity, codability of events and semantic maps to find insights into levels of generality and specificity in these language's lexicalizations and how they relate to their event categoriesation.

In Study 3 (chapter 5) we turn our attention to event cognition with an online experimental method to investigate the planning and production phases of event descriptions

motivated by the language differences discussed in the previous studies. We test Swedish and Italian participants in an eye-tracking experiment, thus excluding English, as these two languages are the most different from each other for their lexical variety and construal structures for which it might be reasonable to find planning and production differences. The experiment aims at exploring the patterns of inspection of events – how speakers focus to event participants and find the relational information to encode in the verb.

Methods and analyses

Throughout the studies we use different types of data and employ different methods of data collection and analysis. First, we use a video database annotated with linguistic descriptions (Imagact). Then, we use corpora to provide a usage-based account of verbs' semantics and access a larger sample of natural-occurring data. Moving forward from corpora and zooming in in our focus of study we collect written descriptions of events. As a last step, we collect spoken descriptions of a further delimited sample of events. Lastly, we use an online experimental method, eye-tracking, to explore research questions on cognition. This is viewed as a necessary step in order to put *more cognitive* into Cognitive Linguistics, as Dabrowska (2016) and Divjak, Levshina & Klavan (2016) encouraged to do.

A crucial element of this work is the used of videos depicting events. Using videos, we represent the meaning by representing its referents but taking a distance from a referential approach to meaning.

Concerning data analysis, we provide a combination of qualitative and quantitative data. Corpora occurrences and linguistic construals (both spoken and written) are analysed qualitative, providing a sketch of the verbs semantics and construal typology. We rely on semantic maps as a way of visualizing meaning and we perform statistical testing on the effects of language and construal types, speech onsets and eye movements.

Scope and Theoretical impact

Categorisation, construal, and event cognition are interconnected cognitive processes that shape how we interpret the world. Studying the interconnection of these processes enhances our understanding of cognition. With regards to the area of research of this work and on the linguistic side, we hope to provide a comprehensive overview on the semantics of the verbs used for the description of breaking events in English, Italian and Swedish, with their differences in categorisation and lexical variety. Theoretically, we want to provide a contribute to the studies in categorisation by focusing on event categories and levels of abstraction, drawing the attention to

what appears to be a, still, understudied area. Lastly, we aim at contributing to research in speech planning and models of spoken production, highlighting the role of video-based elicitation.

Chapter 1. Theoretical background

1. Introduction

This chapter presents the topics of this thesis – categorisation, construals and events – to nestle this work in the framework of Cognitive Linguistics and introduce studies in Event Cognition.

The study of categorisation in Cognitive Linguistics goes beyond the classification of language systems (such as word classes), extending to the mental concepts and our ontological structures (Paradis, 2005). Construals represent the flexible way in which language users interpret the world – using categories and *categorising*. They emphasise the dynamicity of cognitive processes in linguistic expression and account for a degree of subjectivity in the interpretation of the world. Within this research, the focus on Event Cognition plays a pivotal role in understanding the cognitive processes that lead from visual perception to the interpretation of the world to linguistic expression. Framing our discussion in Cognitive Linguistics, we assume the basic tenets of the discipline, concisely summarised in three points – *embodiment*, *conceptualisation* and *meaning* – relating them to the aspects that concern the semantic investigation of event descriptions.

Cognitive linguistics put forward the idea that our understanding of the world is mediated by our bodies (Johnson, 1987; Lakoff, 1987; Lakoff & Johnson, 1980); thus, our basic mental concepts emerge based on our physical experiences and what our sensory-motor system perceives. The categories under which we operate daily, as well as the meanings we create, are grounded and embodied. In cognitive linguistics, it has often been contended that meaning *is* conceptualisation (Langacker, 1987; Croft & Cruse, 2004) – a view that moves away from classical (philosophical and linguistic) perspectives which equate meaning with reference or meaning with truth conditions. Concepts are instead seen as mental entities that *pack* our knowledge through a process that Langacker (1987) calls *conceptualization.* Lakoff (1987: 280–281) argued that conceptualisation is the ability to form symbolic structures, metaphorical projections, complex concepts, and general categories. Evans' (2019) most recent view on conceptualisation is "the ways in which we construe or 'see' the range of sensations, experiences, reflections and so on, that make up our mental life" (2019, p. 7), mirroring Croft and Cruse (2004) view of construal as conceptualisation. The equation of *meaning* with *conceptualisation* is, however, not taken as the equation of linguistic meaning with concepts, as this would have strong deterministic consequences. A one-to-one mapping would imply that people form conceptual structures based

on their native language(s), and people speaking different languages have different conceptual structures – which is a problematic claim (Murphy, 2002). Rather, meanings are mapped to certain forms, and these are part of the conceptual structure, forming concepts, but do not represent the full set of our concepts (Langacker, 1987; Evans & Green, 2007). The idea in Cognitive Linguistics (and specifically in Cognitive Semantics) of this partial overlap and the relation between lexical items and concepts has gained territory in experimental research to investigate the influence of lexical concepts on conceptualisation – as in the hypothesis of *thinking for speaking* (Slobin, 1996) for which linguistic categories can affect conceptualisation under certain tasks. Under this view, lexical items are conceptual categories.

Paradis (2015) claims that words *evoke* conceptual structures, and word forms are cues to structures in a conceptual space. The use potential of a word is not static, as it necessarily is when captured in lexicographic approaches. In principle, it is constantly developing as an effect of the influence of new experiences from language use. And as we experience new things in the world and use potentials of a word change, so do concepts (Scott, 2011). As concepts are closely connected to categories, to explain *how words mean*, Paradis (2005, 2015) posits the interplay of conceptual structures (ontologies) and construal operations. Construal operations are ways of structuring domains and conceptual information. They reflect cognitive abilities – in Langacker (1999): specificity, background, perspective, scope, and prominence. Along with this, construals *tap into* categorisation processes*:* a construal must select a level of generality or specificity, a part-whole relation or a gestalt (for example, *poodle*, *dog*, or *animal or 'put a hat on the rack' vs 'hang the hat on the rack')*.

Conceptual structures related to events in the world are evoked by action verbs, and construal operations determine which aspect of the events is described. Categorisation, as we see it from a linguistic surface, belongs to the operating system of construals rather than conceptual structures (Paradis, 2005).

Having briefly laid out some theoretical premises, we describe categorisation, construals and events more in detail in the following sections. In Section 2, after framing categorisation as a cognitive process, we describe the classical categorisation theory and the more recent works that contributed to an update of the theory. In Section 3, we describe construals, presenting definitions and possible classifications of their dimensions (Langacker, 1987; Talmy, 2000; Paradis, 2005). In Section 4, we provide a definition of events and discuss event categorisation. Despite the interest in the field of categorisation and events from a linguistic perspective, event categories are still underrepresented in the research. Next, we provide a brief review of studies in event cognition to introduce linguistic planning and production in the study of event descriptions.

2. Categorisation

This section reviews studies that investigate how linguistic structures are acquired and organised cognitively, and how concepts of the world are formed and interact with form-meaning mappings. This question has interested linguists, philosophers of language, psychologists, and cognitive scientists. As Harnad puts it in the title of a work, to *cognize is to categorize* (Harnad, 2017).

Categories contribute to giving meaning to events and establishing socially acceptable ways of acting and interacting. Categorisation is a basic cognitive ability highly exploited in scientific knowledge, such as in plant taxonomies, as well as in daily tasks – we distinguish fruits, vegetables, flowers, berries, etc. This organisation means we can judge whether something belongs to a category or not based on similarity (Radden & Dirven, 2007). The ability to form categories (with the process of *categorisation*) is one of humans' most basic cognitive skills. As Lakoff (1987: 5-6) writes, when we speak and produce an utterance, we use dozens of categories that space from sounds to words to constructions up to conceptual categories. Organising the world into categories has obvious advantages: it makes the world more predictable and understandable. More than this, it allows us to treat objects new to us in an efficient way.

2.1. Classical categorisation

The interest in the relationship between entities and labels to refer to actions goes back to Aristotle, where entities were seen in terms of properties and relations between them. This view is known as *classical categorisation*, also referred to by Lakoff (1987) as the *objectivist theory.* The assumptions of the classical model rest on logic, and classical semantics claims that the logical structure of the world is reflected in language, which acquires meaning by fitting the logical structures. In the classical view of categorisation is rule-based: the world's entities fit into categories, whose members have set and defined identifiable features, which make them predictable and stable. When classical categorisation came to be applied to linguistics, it had success in phonology (Taylor, 1989) but encountered two main problems with semantics: it was found that category members do not always share necessary and sufficient conditions and that categories are inherently vague, with fuzzy boundaries. A consequence is that category members do not have the same status in a category. If there are strict conditions to fulfil to be a member of a category but some members do not share all those conditions, they are still regarded as members, but with less typicality. Moreover, if boundaries are not clear-cut, it means some members are at the periphery of a category.

Concerning *category membership*, a renowned example is that of the word *bachelor* in English. If we look for necessary and sufficient conditions to describe the concept of bachelor we get to the following traits (Katz & Postal, 1964: 13): [+human], [+male], [+adult] and [+never married]. According to these conditions, we can call a Catholic priest a *bachelor*. In the practice of our daily life, we do not do so despite meeting the criteria listed above. These necessary and sufficient conditions work from a logical point of view: Catholic priests are human adult males who have never been married, and these conditions can, in principle, form a category. But we need to take into consideration the value of these conditions in our culture: as marriage is such an important social construction, it has been relevant to categorise people as married or not married, but there are categories of people for which marriage is not defining (such catholic priests). Thus, we know through culture that we do not use the word *bachelor* for a Catholic priest. This example shows that necessary and sufficient conditions do not always work – and that listing necessary and sufficient conditions is an impossible task to carry out.

We can exemplify *typicality* in category membership with the bird category: we call birds animals with wings and feathers, that lay eggs, such as a robin or a blue jay. When we are asked for an example of a bird, we might, in fact, reply with 'robin'. In Western culture, it is typical to see these birds and learn BIRD (as a concept) according to their characteristics. We might be more hesitant to call a *bird* a penguin because some of a penguin's perceptual characteristics deviate from our experience of the typical bird.

About *category boundaries*, Wittgenstein example of the word *game* is one of the early pieces of evidence of the problems with classical categorisation and fuzziness of category boundaries. Wittgenstein (1953) notably put forward a discussion on the word game and the impossibility of pinning down the necessary and sufficient conditions for a game to be called such. There seems to be no definite set of properties that define the word *game*, which is shared by all the things we refer to as games. The linguistic category *game* applies to diverse entities that constitute a net of overlapping similarities – like that of a family resemblance (Wittgenstein, 1953). Most importantly, it appears that we know what a game is without knowing an explicit definition of it. Linguistically and conceptually, such categorisation of the word *game* is learnt through exemplars we encounter since we start interacting in the world. Labov (1973) provides the first piece of experimental evidence in the fuzziness of categories: asking people to name a set of drinking vessels, he found that the categorisation operated by participants was, in fact, fuzzy. The boundaries between types of different drinking vessels (e.g., naming them glasses, cups, bowls, or vases) were not the same for all participants: in the continuity of materials for shape and dimensions, there is not a specific point where a glass stop being a glass and becomes a vase (Figure 1). What is more to mention is that the categorisation changed not only according to the visual properties of the objects but also according to the context. In the case of Labov's experiment,

the function of the object in context. Labov asked participants to name drinking vessels giving different instructions, priming participants with a clue on the use of the objects (e.g., 'name these objects used for food' or 'name these objects used for drinking').

Figure 1. Drinking vessels in Labov (1973)

He found that the switch point of category from cup to bowl changed its boundaries depending on whether the viewer thought the object was for drinking or eating. Since category membership and fuzziness were acknowledged, the classical categorisation was called into question, and two similarity-based categorisation theories emerged: the prototype theory and the exemplar theory. These theories take into consideration the role of our body and our minds in meaning-making processes and attempt to explain the way world entities are conceptualized and mapped into linguistic entities.

2.2 The cognitive approach

Cognitive linguistics follows this change of perspective and challenges the notion of static and rigid categories of the classical categorisation to propose theories in which categories are flexible and dynamic, to account for fuzziness and typicality effects.

The awakening is led by a range of disciplines working in cognition (such as cognitive psychology or the neurosciences). Linguistics was especially influenced by the ground-breaking works of Rosch (1978), Barsalou (1983) and Lakoff (1987). Cognitive Linguistics research in categorisation has looked at semantic diversity across languages, particularly for discrete entities and in domains such as artefact names (Malt et al., 1999), spatial relations (Bohnemeyer, Donelson, Tucker, & Benedicto, 2014; Levinson, 2003; Levinson, Kita, Haun, & Rasch, 2002; Pederson et al., 1998), numbers (Frank et al., 2008; Gordon, 2004), colour terms (Berlin & Kay, 1969). Two basic principles for the formation of categories are discussed by Rosch (1978):

1. Cognitive economy: the function of categorisation is to provide for the maximum information with the least effort. Categorisation helps organise the information from the world individuating differences between stimuli and recognizing them as similar. To be efficient, categories cannot provide differentiations that are not relevant to our purposes.
2. Perceived world structure: the world's knowledge is perceived as correlated structures of information. The example provided by Rosch is that for a person in this world, it is an empirical fact that wings cooccur with feathers rather than, for example, fur. As cognizers, we are able to structure the world according to what we know may or may not occur to different degrees (e.g., what is highly correlated, what is a low probability and what is impossible).

These two principles operate together: when information is perceived as structured, the first principle of cognitive economy is realised, and vice versa. These two principles allow for the efficient organisation of knowledge, even in terms of linguistic encoding. According to Croft and Cruse (2004: 74), categories serve four essential functions in language: learning, planning, communication, and economy.

That items belong to a category according to different degrees is explained by the two different models mentioned above: the *prototype model* and the *exemplar-based model*. Tere are not necessary and sufficient conditions, rather a net of properties that may or may not apply to a label, bringing into the theory the idea of *degree* of membership to a category. The prototype model is based on family resemblances and the idea that categories are formed abstracting similarities from individuals encountered in the world, structured according to a prototype of the

most recurrent features, to which each new individual is compared. In the exemplar structure (Hintzman, 1986; Nosofsky, 1986) a category's mental representation encodes each exemplar that constitute a category. Exemplars are built gradually, keeping traces of all the instances of an item encountered (e.g., all chairs a person has seen, experienced or even heard about). To decide whether a new entity is a member of a category, one compares it to its stored exemplars rather than to the prototype of the encountered individuals.

2.3 Basic-level categories

Berlin and Kay (1969) presented seminal work on colour terms, demonstrating that various languages have distinct boundaries within the colour spectrum but that a higher-order categorisation of "basic colour terms" emerges, representing the most prototypical examples for different colour terms. The prototype of a category is a representation of the most common features associated with a category (Rosch & Mervis, 1975). It follows that in learning about the world, categories are formed with experience and based on encounters with the individual instantiations of a category (such as in the BIRD example). Once prototypes have emerged, categorisations can be predicted on a similarity-judgement basis, comparing new individuals to a prototype: the likelihood of placing an object into a category increase as it becomes more similar to the prototype (Mervis & Rosch, 1981). The notion of a prototype as the foundation of category formation also led to observe, in a series of experiments (Rosch et al., 1976; Mervis & Rosch, 1981), that, as much as in colour terms, there is a perceptual level in the world's discontinuities that is most relevant for cognitive and linguistic purposes. The idea of a basic level of categorisation rests on the relevance of perceptual characteristics on that category members on cognitive operations (Rosch & Lloyd, 1978; Murphy & Smith, 1982, Murphy, 2002). If cats and dogs are basic level, the differences between cats and dogs are more relevant than the differences between members of each of these categories (e.g., for cats: Siamese and ragdoll; for dogs: Pitbull and golden retriever).

From here, categories have been discussed with what is below (that is, more specific) the basic level (e.g., a ragdoll cat) and what above, that is, more general (e.g., a mammal) in a three-way hierarchical organisation: a superordinate, a basic and a subordinate. This hierarchy represents levels of abstraction and schematicity in the conceptualisation of categories. Using a different example, 'chair' is a basic level category (a generalization of specific types of chairs), 'kitchen chair' is a subordinate, and 'furniture' is the superordinate (generalizing over lower-level categories).

In Rosch and colleague's words, basic level categories are "information-rich bundles of attributes that form natural discontinuities" (Rosch & Mervis, 1975: 586). The basic level is the level at which "categories maximize within-category similarity relative to between-category similarity" (Rosch et al., 1981: 92). As for their characteristics, basic level categories share motor actions for interacting with category members, members have similar overall shape, and a mental image can reflect the entire category (Rosch et al., 1981). The basic level enables us to identify and recognize a category as quickly as possible: recognizing and categorizing 'chair' is quicker than 'furniture' and 'kitchen chair'. The basic level is defined as such because of its relevance in cognition as the most efficient way for us to interact with the environment. For this reason, it is argued that it is learned early by children, before the superordinate and subordinate levels. Such as FURNITURE, ANIMALS or CLOTHING are higher-level categories: we do not have the same motor programs or knowledge to use with all pieces of furniture, animals, or pieces of clothing. Superordinate categories are too general to be efficient or cognitively relevant for our daily life purposes. Additionally, we cannot form a mental picture of a general piece of furniture, but we can picture a chair, a table, a desk, or a wardrobe. From a sensori-motor point of view, we have a conventional way to interact with chairs, but we do not have a generalized motor schema or knowledge to interact with all pieces of furniture in the same way. However, there is a common ground that brings together all pieces of furniture that determine a higher-level categorisation – the fact that they are found in homes or offices, they are functional for daily activities, etc.

Taking the perspective of language, categories are used to *communicate* about our daily activities and the basic level is assumed to be the most relevant for linguistic purposes (Rosch, 1978) and the most relevant part of our form-meaning mappings. This implies that basic level categories are coded in languages and, according to Rosch (1978: 35), emerged historically prior to the terms denoting superordinate or subordinate categories. The different levels of abstraction are reflective of our complex environment (Borghi, Parisi & di Ferdinando, 2005): with our language, we can establish higher and lower categories that go beyond the perceptual factor to incorporate cultural or functional aspects (Györi, 2013). At the superordinate level, perceptually different entities are brought together: it is the case of FURNITURE, which emerged in a cultural environment where different types of objects are used inside homes to provide for storing space and sitting areas. At the subordinate level, perceptually similar items with different functionality, tactile feel or goals can be differentiated (kitchen chairs and dining chairs may be visually similar but culturally different).

2.4 Ad hoc categories and lexicalisation

Experimental work carried out in the last decades enhanced the view of categorisation as dynamic and categories as inherently flexible (Barsalou, 1983; Smith & Samuelson, 1997; Whittlesea, 1997, Mauri & Sansò, 2018). In fact, categorisation as a process is constantly at play when we speak and interact with others (Edwards, 1991; Goria, 2020). A substantial contribution to this field was made by Barsalou (1983) with the introduction of the notion of *ad hoc* categories to capture the dynamicity and flexibility of categorisation. Ad hoc categories are context-dependent and serve us for short tasks (thus, without the need to store them in memory). Their members might not share visual properties or functional properties and typically do not have lexicalisations. Although this work does not examine ad hoc categories, their flexible nature and the stress on the lack of lexicalisations of ad hoc categories are useful for the purpose of discussing event categories.

Let's consider the ad hoc category *things I bring to the beach:* its members may include a towel, sunscreen, bathing suit, sunglasses, and folding deck chairs. If we imagine ourselves listing the members of this category, each of us might have different ideas of what to bring depending on one's needs and habits (someone might not need a deck chair!) but may also agree to others' ideas – find them plausible and motivated. This reflects the potentially subjective nature of ad hoc categories – and by extension of categorisation processes – but also of the inter-subjective aspect of it. Moreover, the members do not share any physical or sensory features and individually serve different functions (e.g., protect from the sun, provide a place to sit…) but are brought together by their higher-order function, that is, aiding us in our trip to the beach. The formation of the category is prompted by the activation of a frame, resorting to one's knowledge of the world and in reference to an activity relevant to the speaker (Goria, 2019). The idea of category members not having shared features is what characterize the superordinate level in the basic level theory. In both cases (ad hoc categories and superordinate level), the formation of the category is motivated by a higher-order functional aspect that captures abstract shared characteristics rather than the physical ones.

Lexicalisation is the other aspect particularly relevant in discussing *ad-hoc* categories, and relevant for event descriptions. It emerges from the studies on *ad-hoc* categories that their flexible and temporary nature is testified by the fact that they generally lack a lexicalisation, hinting at the idea that a category is less of a category if we do not have a label. The term *lexicalisation* is widely recognized and used in linguistics to refer to the process of integration of concept or meaning into a lexical item. Two main principles that guide lexicalisations are described in the literature (Stathi, 2023). The first, (*Relevant Lexicalization*) predicts that a language lexicalises those components that are relevant for a group (Gast et al., 2014). The second (*Intrinsic Lexicalization*) predicts that semantic components lexicalise when they have a mutual influence (Bybee, 1985), exemplified by

motion verbs where certain meaning components are integrated with the movement (such as manner of motion) but not others (English motion verbs do not integrate motion and weather conditions - Stathi 2023). Lexicalisations thus show levels of generality and specificity in encoding semantic components into lexical items, which can be interpreted in terms of levels of abstraction for categorisation purposes.

As far as the break category is concerned, it will be shown in Chapter 2 that the languages under investigation in this work differ in their lexicalisations and form-meaning mappings: Swedish does not lexicalise in a verb the break category in the same way as English or Italian do, even though the concept of object disintegration or malfunction is a relevant one for (at least) Western culture. This cross-linguistic difference and its possible implications for categorisation theory will be addressed in Chapter 4.

2.5 Towards event categories

Even though in Cognitive Linguistics the prototype theory of categorisation and basic level categories are popular and vastly researched, event categories are still an under-researched area. The work on prototype theory and basic level categories relies heavily on discrete entities (such as objects) and the primacy of visual perception in human's ability to categorise the world (such as with colour terms – Berlin & Kay 1969). Work in this field has suggested that there is no univocal hierarchical way of organizing other domains of experience, such as actions and events, that would reflect our cognitive organisation (Vigliocco et al., 2004). Ungerer and Schmidt (2006) note that the analysis of action categories in terms of a three-level hierarchy is applicable for some activities (e.g., with the verbs *eat* or *walk*): it is possible to distinguish most prototypical actions (e.g., walking down the street) from the less prototypical ones (e.g., walking backwards), and positing a hierarchy. For example, in the hierarchy presented in (1) *walk* is the basic level, and as such is supposed to be the highest level at which one can form a mental image of the category (Taylor, 2018: 48). Its superordinate is the verb *move,* and at the subordinate level, we can find specific ways of walking (*stroll, amber, saunter).*

(1)[move > walk > stroll, amber, saunter]

Event categories are a complex of actions and discrete entities, which consider events as wholes, made of participants (as objects or living entities), actions, spaces, and time. Taking an event category, we can turn to events of stealing (Lalljee, Lamb & Abelson, 1992). Events of stealing are characterised by the type of person who performs the act, what is typically taken, how it is done,

the goal of the action and its typical explanation. They can be described in terms of prototypes which form different *event prototypes*[1], consisting of different features and requiring specific knowledge of the world (a specific society where these features together are recognized as an act of theft). In the study, there were first identified different stealing events, and then participants were asked to describe the features of these events. It was found that participants systematically organised these events in terms of prototypes: the data was analysed with a correspondence analysis, which showed which events were closer together based on their features, which comprised objects, actions, and people. However, the literature on event categories is still lacking systematic work. This topic, central to this work, is discussed in Section 4.1 and is the focus of the study reported in Chapter 4.

3. Construal

Construal is a crucial term in Cognitive Linguistics to account for the cognitive basis of semantics. It captures the idea that the way in which we describe an event reflects our conceptualisation of it and accounts for a certain degree of subjectivity in linguistic expression and categorisation. Languages provide us with a variety of ways to express a single event – that is, to categorise events (Verhagen, 2010). A classic example is that of the passive-active alternation (2) for syntactic strategies that affect construals, but construal operation can also be based on semantic choices (3) or that of

(2)a. John ate the apple
b. The apple was eaten by John
(3)a. John wiped the table
b. John cleaned the table

In example (2), the syntax changes the focus. In example (3), the verb choice provides a different selection of event components: if someone was rubbing a dirty table with a cloth, I could either say that they wiped the table or that they cleaned the table—and I would be focussing on a different aspect of the event, the rubbing motion or the resulted clean state of the table.

The referent in the world of the examples (2-3) is one but the two linguistic possibilities for expressing one and the same situation are two ways of *construing* the situations that are compatible with the reality we see. Construal analysis has been vastly used - especially in

[1] Event prototypes are complex schemata, which are not distinguished from each other by single features, but form wholes of related features that people drawn upon to explain category membership of events.

discourse studies (Möttonen, 2016) to investigate the conceptualisation of life domains (e.g., in press reports of violence, Hart 2013) and applied to education context (e.g., construals of narrative texts in Creanga 2018).

With *construal* Langacker refers to "the mental ability of a speaker to construe a conceived situation in many alternate ways" (Langacker, 1987: 487), that is conceptualized as such and not imposed by the event itself. The use of a linguistic expression reflects a construal and the speaker's experience of the world. As such, differences in construal are differences in meaning. Construal operations are the source of all readings of a word's evoked meaning, conventional as well as idiosyncratic (Paradis, 2015). Underlying this concept is the idea, then, that true synonymy does not exist: each change in levels of specificity or constructions (e.g., such as alternations) is the carrier of a different conceptualisation. Under this view the fundamental role in meaning making is played by the language users themselves, who through construals (and categorisation) can negotiate their experiences in the world and their narration (Divjak, Milin & Medimorec, 2020). Thus, the focus in cognitive linguistics to study meaning is put on the subjects and the use, rather than the *objects* alone (Verhagen, 2007).

3.1 Dimensions of construal

Langacker (1991) writes that a chosen linguistic expression imposes an image on the content it evokes. The semantic value of linguistic expressions is to be found in dimensions of construal imposed on a situation or event, which are, as explained above, realisations of a speakers' conceptualisations.

The specific properties of construals, often referred to as *dimensions of construal* are most notably discussed by Langacker and Talmy, which posit a classification that is partially overlapping. The classification emerges from examples and observations of language, however not systematized. Their classifications have been criticized for appearing to be somewhat arbitrary (Croft & Cruse, 2004; Verhagen, 2010). However, being these classifications based on empirical data and the bast attempts at structuring and organizing the construal theory, they have high relevance in Cognitive Linguistics. Both Langacker and Talmy revised their classifications over the course of time (Langacker, 1987 then 2010; Talmy 1985, then 2000) and the large overlap of their classifications is taken as a consistency in the models and a confirmation of the relevance of the dimensions. We review the dimensions that capture the cognitive features of construals according to the latest work by Langacker (2010):

- *Perspective*. It is characterised by further distinctions, which are Figure/Ground alignment, Viewpoint, Deixis, and Subjectivity/Objectivity.
- *Selection*. It has levels based on granularity (specificity or schematicity). This dimension is related to the linguistic level: the fact that lexical categories can form taxonomic hierarchies which consist of levels of specificity, such as hyponymy-hypernymy relations (Golden retriever > dog > mammal > animal).
- *Prominence*. It includes profiling - foreground vs background of a figure. This is the concept of prominence of a figure, which receives attention if it is in the foreground. Generally, it is not predetermined by the event or situation but reflects the speaker's will to direct the attention to an entity.
- *Dynamicity*. About this, Langacker stresses that language and conception are about things that happen and evolve through time.
- *Specificity*. The level of detail included in a linguistic expression or a form-meaning mapping. This dimension needs to be understood as context-dependent – that is, in reference to another expression. Context-independent specificity is an ideal concept and may only be judged imprecisely (Möttonen, 2016).

For this work, specificity is a key feature of event descriptions. Different levels of specificity in construals describing the same event alter the schematicity at which the event is described. The same event can be described with the minimal syntactic components (SVO in case of causative events) or with a more thorough elaboration of the actions underlying the event (syntagmatic elaboration – Möttonen, 2016). For example, the same event can be described as in example (4).

(4) a. Mary prepares tea

b. Mary puts water in the electric kettle and turns it on, takes a mug from the shelf, puts a teabag in it and when the water boils, pours the water in the mug.

Preparing tea does not equate to the list of procedures in (4) (tea can also be prepared following other procedures) but includes them. The two construals give different pieces of information, one more schematic than the other.

Talmy's (2000) Dimensions, listed below capture different aspects of construal operations:

- Configurational Structure (comprising e.g., boundedness)
- Perspective
- Distribution of Attention

- Force Dynamics

Fundamental to this classification is the relevance given to Force Dynamics, thus capturing the physical dynamics of actions and events – in a view that can be said to be *embodied*. Force dynamics are hypothesized that force dynamics underly grammatical categories (e.g., in modals) as well as lexical categories, playing "a structuring role across a range of language levels" (Talmy, 2000: 410). An example proposed by Talmy (1988) is that of sentence (8): the difference between a. and b. is that while a. describes a state, b. communicates that John's willingness to go out of the house is blocked by some blocking force (Talmy, 1988: 52) – be it physical or metaphorical (example 5).

(5)a. John doesn't go out of the house
b. John can't go out of the house

Paradis (2005) proposes a model of lexical meaning based on ontology structures and construal. In this model, concepts form the ontological basis of lexical knowledge (which involves both encyclopaedic and linguistic knowledge). Construals are the operating system of the conceptual structure. They are not conceptual in nature but serve to structure conceptual domains: speakers use them to give a structure to their experiences according to cognitive strategies—focusing attention and choosing configuration.

Möttonen (2016) highlights the nature of construal as inevitable: every linguistic expression is necessarily *construed* by the cognizer, so dimensions of construals are to be seen as properties of language use. Dimensions of construals thus inform us that there is no objectivity in linguistic expressions and no single way to describe a situation, and the concept of construal itself serves more to the scope of investigating why specific types of construal structure semantic information in the way they do (Möttonen, 2016). This view stresses the subjectivity of linguistic production and situation framing, underlying that meaning-making is dependent on the cognizer.

3.2 Construals cross-linguistically

It has been discussed how construals differ within a language. Given the already discussed variety of linguistic means giving rise to form-meaning mappings, we consider now cross-linguistic diversity and the diversity of possible construals across languages.

Typological distinctions may correspond to different construals that form different patterns in different languages. These studies have flourished with Talmy's investigations on motion descriptions (Talmy, 2000), which first theorized the typological difference between satellite-framed and verb-framed languages for the encoding of motion events. Notably, Germanic languages belong to the first kind and Romance languages belong to the second. Languages, in fact, have different means to organise spatial relations but they also follow preferred patterns when more framing possibilities are available. Taking an example from motion events and Talmy's distinction, English features possible construals of both satellite-framed and verb-framed languages (see examples 6a-b). Nonetheless, according to experimental evidence, the preferred framing in language use is the one that conforms to the satellite-frame configuration (Slobin, 2017).

(6)a. Mary enters the room dancing
b. Mary dances into the room

What is the role of subjectivity in these cases? There are conventionalized patterns where the cognizer has little or no freedom of choice in picking a construal according to a dimension, as sometimes and for some domains, languages may lack some of the patterns (Verhagen, 2010). Typological research has also shown the influence of conventional patterns in a language on cognitive processing (Levinson, 2003). With respect to the satellite-framed and verb-framed language dichotomy, Slobin (1996) has developed the idea of 'thinking for speaking': patterns of one's native language force its speakers to habitually pay attention to certain aspects of events that are that are necessary for linguistic expression. In fact, it is relevant to add that cross-linguistic investigation on construals can reveal patterns of conventionalization in languages. Conventionalization in the framing of a situation or event comes at the expense of the subjectivity of construals and dynamicity of categorisation: high degrees of conventionalisation or entrenchment are a sort of cognitive routinisation which establishes a mapping that is stored and accessed as a whole (Langacker, 2008; Schulze, 2009).

4. Events

The discussion carried out in the previous sections highlighted the role of categorisation and construal in cognition, with a focus on linguistic expression. In this section, we move to discuss events and event cognition to set the grounds for the investigation of event descriptions.

Chapter 1. Theoretical Background

Events can be defined according to different perspectives. As pointed out in the introduction to this work, events are *units of experience* (Yates, Sherman & Yousif, 2020). This term captures that events are perceived as an independent frame of space and time that has meaning on its own. Other definitions capture other aspects, such as causality or the temporal dimension:

- "The set of participants, activities, and outcomes that are bound together by causal interrelatedness" (Elman, 2009: 572)
- "a segment of time at a given location that is conceived by an observer to have a beginning and an end" (Zacks and Tversky, 2001: 3).

The observer is a crucial part of the definition of events: an event is recognized as such on an intersubjective agreement on time-spatial boundaries (Davidson, 1969). Concerning categorisation and events, it is still to be thoroughly investigated what could be a possible relation between language and event categories: since events are perceived as units, we ask whether the way they are described, with higher or lower specificity, reflects the hierarchical organisation.

Cognitive science, psycholinguistic and neurolinguistic experiments carried out in event cognition look at how semantic meaning is processed in online tasks and in our brain, how it works together with other cognitive abilities and how it is linked to categorisation processes that occur in language and across different languages. These studies were aimed at understanding how languages segment reality and to what extent language interacts with (which) cognitive abilities. Describing events requires understanding many pieces of information and a conceptualisation of them: understanding and interpreting semantic roles, the relation between them, the pragmatic context, and providing a construal. How do we select and interpret these pieces of information and encode them in language? How do typologically different languages differ in patterns of event descriptions?

Among all experimental methodologies to investigate these questions, eye-tracking is an online research methodology widely applied: it taps into cognitive processing as it occurs. Research questions explored in this area look at how speakers convert event representations into utterances (that is, how we transition from event conceptualisation to sentence planning), how attention is allocated when speakers are planning to describe an action for verbalization purposes, or how we decide what information to include in an event description. A central interest is in the processes involved in the interplay between the conceptualisation of events in the world and how we describe them. The phase between information uptake and description is that of sentence planning: speakers extract meaning from what they are seeing and plan a message. This process is extremely fast: studies show that the gist of a scene is extracted after 300ms of exposure (Griffin

& Bock, 2000). The phase of the very initial information extraction is referred to as *event apprehension*. Eye-tracking is a convenient methodology to investigate where the speaker turns their attention in the first milliseconds of exposure (where is the first fixation allocated) and how language is linked to said location. The following sections are devoted to what concerns events categorisation and event cognition, discussing first event apprehension, and second speech planning.

4.1 Event categorisation

The categorisation of discrete entities is thus *grounded* in its relation between the concepts and the linguistic form, being driven by sensory information and according to the way we interact with them (Harnad, 1987; 1990). Categories are formed for relational concepts as well – such as the interaction between entities. Consider the event of a person breaking a vase, which can be expressed as 'the boy breaks the vase'. Construals of this event presuppose the knowledge of the concepts of BOY, BREAK and VASE That is, that the boy is an animate agent, capable of carrying out actions, that the vase is a brittle object, and that *break* is an action producing material destruction or non-functionality. It also presupposes the relation between these three elements within a specific context, which needs to be coherent (e.g., the vase breaks the boy is not coherent; experimental work in coherence and incoherence of events: Dobel et al., 2007). The three processes are all linked to each other in meaning-making, assuming that words evoke conceptual structures (Paradis, 2005) and that in virtue of such conceptual knowledge language users create categorisations and apply construals.

When thinking of the broader category of events the scenario comes to include not only perceptual information, but also motor schemas, force schemas and spatial schemas along with the temporal dimension. First, events are referred to with linguistic items that encompass a range of different features: as an example, the category "actions humans perform on objects to access a content", which is generally referred to in English with the verb *open*, includes events such as opening a window, a door, or a jar, which are all visually different from one another. However, there is not a complete match between this category and the English verb *open,* which can be used for events such as *opening an umbrella*, which involves the expansion of an object in a space. Second, an action unfolds in time and the unfolding is defined by all the entities involved: an agent or an actor, a patient, the background, the context, a manner of performance, an instrument. Contrasting the lexicalisations of different languages becomes key: different languages communicate about events using different lexicalisations within and across domains (Talmy, 2000; Stathi, 2023). Going back to the case of *open*, *opening an umbrella* in Swedish is not

described with the cognate English verb *open*, (Swedish *öppna*), rather with the verb *fälla upp* or *veckla ut* (lit. 'spread up', 'unfold').

Studies in categorisation moved to test events (Rifkin, 1985; Morris & Murphy, 1990) and more specifically the role of action (Barsalou, 2021; Anelli, Nicoletti & Borghi, 2010), function (Chaigneau & Barsalou, 2008) and goals (Gorgakopoulos & Härtl, 2020; Ratneshwar et al., 2001; Hamilton & Grafton, 2006) or motion (Slobin et al., 2014; Malt et al., 2014; Talmy, 2000) and also, more generally, verb categories (Fillmore 1978 with the semantics of the verb *climb*; Coleman and Kay 1981 with the verb *lie*). While some studies argue for event categorisation to operate with the same abstraction levels as object categories (Rifkin, 1985), others acknowledge that event categories are not organised in a hierarchical three-level model (Taylor, 2018; Györi, 2019). In linguistic studies event categories have been mainly investigated through verbs, but specifically focusing on *activities* (Vendler, 1969; Croft, 2012) such as *eating, sleeping* and *walking.* As an example, the category *eat* in English seems to conform to the hierarchical categorisation (Györi, 2019; Ungerer & Schimdt, 2006): the verb *eat* is a relevant verb to our daily life and our lexicon, is acquired early by children and is easily described by a set of perceptual attributes ('picking food up', 'opening one's mouth', 'putting the food into one's mouth', 'biting', 'chewing', 'swallowing'), which relate to ways of using objects to achieve a purpose (feed oneself) in precise motor schemas (lifting the object, moving our jaws) and/or spatial schemas (directionality of the movement towards the mouth). The activity of eating is composed of individual actions in a part-whole relation (*bite*, *chew* and *swallow*), and it is possible to link it to gestalt perception (there are prototypical instances, such as eating a sandwich and peripheral instances, such as feeding another person). Györi (2019) suggests a possible hierarchy where *ingest* is the superordinate term, *eat* and *drink* are basic level and among the subordinate of the basic level *eat* are *gobble*, *nibble*, *devour* (7).

(7)

INGEST

EAT - DRINK

Gobble, nibble, devour

In linguistic terminology, the superordinate and subordinate relations are termed hypernyms (*ingest*) and hyponyms (*gobble, nibble, devour*). More specifically, in the case of verbs, the subordinate terms show a relation of troponymy (Fellbaum, 1990), which indicates '*a manner of doing something*'. That event categories involve the same hierarchical organisation as object categories is also argued in Zhuang & Lingnau (2022), although taking a different perspective. As shown in Figure 2, their proposed subordinate level is not based on linguistic relations of

hyponymy or troponymy. It is individuated as single event instantiations which extend from a single verbal root (e.g., drink > drink water, drink beer, drink coffee) where no actual categorisation (=grouping based on similarity) is retraceable at the conceptual or lexical level. The asymmetry between the conceptual and the linguistic forms – in particular, in lexicalisations – can be investigated by taking linguistic diversity as a proxy in the operationalization of our research question.

Superordinate level	Basic level	Subordinate level
Locomotion (sich fortbewegen)	To go (gehen)	To walk (spazieren gehen)
		To hike (wandern)
		To walk a dog (Gassi gehen)
	To drive (fahren)	To drive a car (Auto fahren)
		To ride a bike (Fahrrad fahren)
		To take a bus (Bus fahren)
	To swim (schwimmen)	To swim front crawl (Kraulschwimn
		To swim breaststroke (Brustschwim
Ingestion (Nahrung aufnehmen)	To drink (trinken)	To drink water (Wasser trinken)
		To drink beer (Bier trinken)
		To drink coffee (Kaffee trinken)
	To eat (essen)	To eat an apple (einen Apfel essen)
		To eat cake (Kuchen essen)
	To cook (kochen)	To cook noodles (Nudeln kochen)
		To cook soup (Suppe kochen)

Figure 2. Zhuang & Lingnau's (2022) hierarchical model of event categories.

We can attempt to posit a hierarchical organisation that takes into consideration single instantiations of events (as in Zhuang & Lingnau, 2022), with a hierarchy built on lexicalisations at the basic level and constructs of verb and thematic roles at the subordinate level. In this case, we would have the situation in Table 1 for English and that of Table 2 for Swedish. The two languages display two different categorisations and basic levels.

Superordinate	[change of state]
Basic	Break
Subordinate	Break a vase break an arm break an egg

Table 1. Tentative categorisation of breaking events in English

Superordinate	[change of state]	[change of state]
Basic	Ha sönder (Eng. break)	Bryta (Eng. snap/break)
Subordinate	ha sönder burken (Eng. break a vase) ha sönder koppen (Eng. break a cup)	Bryta armen (Eng. break an arm) Bryta pennan (Eng. break a pen)

Table 2. Tentative categorisation of breaking events in Swedish

Although it seems reasonable to consider constructions of verbs and their themes rather than roots only (we do not expect people to verbalize 'John breaks' when witnessing a breaking event), the question can be asked as to how the subordinate level emerges in this model. The model is not based on similarity or language-specific verbal lexicalisations: drinking water and drinking beer are two different subordinates even though they share perceptual similarities and verbal lexicalisations (there is not a verb that lexicalises 'drinking water' and another that lexicalises 'drinking beer'), thus appearing arbitrary. From a contrastive perspective, it could capture linguistic differences in verb-themes relations: in the English and Swedish table, we see that the English 'break a vase' and 'break an arm' are included in Swedish as subordinate to two different basic levels. As these cases will be discussed in Study 2, we can add here that 'break the egg' in Swedish would also belong to a different basic level as this event is referred to with a more specific verb *knäcka äggen* ('break the egg').

Moneglia (1998; 1999) points attention to *general verbs* (referring to achievements and accomplishments, in Vendler's terms) rather than activity verbs. As discussed above, activity verbs have perceptually similar action schemas and, as such, extend to actions similar to each other. These verbs (e.g., *walk, eat*) have prototypical instances and peripheral ones, and it is possible to retrace a hierarchical structure. General verbs, instead, refer to a greater variety of possible events in the world. Lacking perceptual similarity, Moneglia (1998) argues, they cannot constitute *basic level categories*. However, *general verbs* such as *break* or *open* are learned early by children (e.g. before *unlock)* and are the most efficient ways to structure reality (as basic level categories). According to this, we could posit three different outcomes in event categorisation: i. general verbs reflect a basic level relevant in linguistic terms, but there is not a superordinate category in event categorisation; ii. general verbs are a superordinate level and in the case of event categories, children learn the superordinate level first; iii. general verbs are neither basic nor superordinate and event categories have a different organisation.

The principle of inclusiveness of superordinate, basic and subordinate levels is reflected in linguistics in hyponymy, hypernymy and troponymy (Fellbaum & Miller, 1990: 566 "the semantic relation denoting a way of doing something" that is described by a more general verb). György (2019) investigates inclusiveness of verbs in these lexical relations to find hyponymy/hypernymy relations of inclusiveness in the category of INGESTION, indicating a vertical

relation: if I eat something, I am ingesting something (see example (8): *devour* is inclusive of eating and, in turn, ingesting). At the same time, the troponymic relation is not based on inclusiveness (example 9) and it is not found in object categories[2] (example 10).

(8) I always devour my meals after the gym

(9) a. Did you devour the chips?

b. No I just ate them. (Györig 2019)

(10) a. Which dog did you get? Did you get a Bulldog?

b. ??No, I just got a dog (adapted from Györig 2019)

In contrast with example 10, example 9 shows that eating is a basic way of doing something that can stand on its own in a way that 'dog > bulldog' cannot.

General verbs such as *open* and *break* pose some challenges to categorisation studies for their status at the basic level and for which verbs could be their superordinates and subordinates. Starting from the superordinate level, we ask which is a more abstract level of the break category. At the conceptual level, we can zoom out to a more general category, including *break* and *open* in the *change of state* superordinate category. As far as lexicalisations are concerned the process is not straightforward: *separate* could be posited as a higher level term (the category *separation events* is used in Majid et al. 2007: see Chapter 2, section 5) but not all that is part of the English *separate* is also part of the English *break*: if I eat I also ingest but if I *break* a glass with a bat, I am not *separating* the glass (though the result is a material separation). Using lexicographic resources, we can see that *break* is defined with meanings that appear to be more general (*destroy, discontinue, disjunct, divide, fracture*) which, again, are not superordinate of *break*: as an example, divide does not include all breaking events, but it is possible that breaking events are also described by the verb *divide*. Inclusiveness does not seem to hold as it does for activities and discrete entities.

The subordinate level is just as challenging: positing subordinate from-meaning mappings for *break* means looking for levels of specificity which describe a complex event at a finer granularity, but which include the general *break*. For example, we could think of a subordinate of a breaking event such as the English *snap* (e.g., snapping a twig), which denotes an additional characteristic of a breaking event. According to the Oxford English Dictionary[3] , snap is "to break suddenly and (usually) with a sharp noise or report; to give way or part suddenly owing to strain

[2] For other action categories, such as PLACEMENT, other issues arise in finding a superordinate of the verbs sit, stand and lie (Newman 2009), while troponyms are possible (see Györi 2019 for a complete list).

[3] https://www.oed.com/search/dictionary/?scope=Entries&q=snap, last access 10/01/24

or tension". *Snap* appears to convey finer-grained information about how the change of state occurred, but if we use the troponymy test from above, we notice that snap is not simply *a manner of doing* a more general activity. The b. response in example (14) is not a satisfactory explanation to the question[4]: compared to example (9) one expects a clarification which the verb *break* cannot give – and cannot be a *basic way* interpretation.

(11) a. How did you break the pencil? Did you snap it?
b. ?No, I just broke it

Adding to this, we could test positive/negative responses with basic and subordinate candidates. In (12b) *snap* can be used as the specification of the breaking event as a valid answer to the question in (12a). Example (13) shows that we cannot exclude *snap* from the larger category of *break*, as the event in (13) can be described as both breaking and snapping.

(12) a. Did you break the pencil?
b. Yes, I snapped it
(13) a. Did you break the pencil?
b. ?No, I snapped it

Given these considerations, we can go back to the tripartite scenario for event categories outlined above and try to sketch a hypothesis.

Looking at Table 3 we can hypothesize that the three different categories of nouns, activity verbs and general verbs display different categorisation relations. Inclusiveness is the only relation possible for a hierarchical categorisation of discrete entities, which have a superordinate-basic-subordinate categorisation that is linguistically relevant in lexicalisations.

	EXAMPLE	**ENTITIES**	**RELATION BETWEEN ENTITIES**
NOUNS	Did you buy a bulldog? *No, I just got a dog	Dog = basic level Bulldog =subordinate level	Inclusiveness in the category
ACTIVITY VERBS	Did you devour the chips? I just ate them	Eat = basic way of doing some activity Devour = specific way of doing some activity	Inclusive and troponymic

[4] One can envision the B. response in a pragmatically marked context for someone wanting to cover how the breakage occur by using vagueness as a strategy.

GENERAL VERBS	Did you snap the pencil? *No I just broke it	Break = Basic level? Snap = Subordinate?	Not a basic way, not a troponymy Basic level?

Table 3. Tripartite distinction of noun categories, activity verbs category and general verbs category

The three-level categorisation is plausible for activity verbs, but other than inclusiveness, it is also based on troponymy. These two distinct relations for activity verbs (a basic level and a basic way of doing something) can be thought of as extending vertically (where the higher level includes the lower level) or horizontally (where one level can exist independently of the other level, as in the eat-devour example). Finally, according to the domain investigated here, general verbs do not have *a basic way of doing something*, and there seems to be a missing lexicalization for a superordinate level. This issue will be explored in Study 2, reported in Chapter 4.

4.2 Event Cognition

As humans, we talk about what we see, perceive and experience in the world. We share a biological and cognitive system which makes us go through very similar experiences. For this reason, it is claimed in theories of cognition that some core aspects of human perception and cognition are universal (see Norenzayan, Choi & Peng, 2007 for a discussion). According to this view, events and object representations we generate are all likely to be at least partially shared by individuals, regardless of culture and the language we acquire as children (Gleitman & Papafragou, 2005; Jackendoff, 1996). Within cognitive science studies, embodied theories claim that cognition is the result of the interaction between our body (and what it can experience given its physical characteristics) and the world (Glenberg & Kaschak, 2002; Gallese & Lakoff, 2005; Pulvermuller, 2005). However, as discussed in Section 2, events are complex conceptual items, whose understanding goes beyond perception and motor-schemas. In this work, we are interested in the way linguistic structures and lexical items contribute to event cognition, in how linguistic production and categorisation interact with linguistic construals.

The term *event cognition* refers to the processing of experiences into meaningful events. Event cognition examines the cognitive representations and processes used to understand events in the world as they are experienced, remembered and used to plan future actions (Kurby & Zacks, 2008; Radvansky & Zacks, 2011). As the world unfolds before us in a continuous stream, we can isolate meaningful units or *events.* This is the first step in event cognition, *event segmentation*, which is a prerequisite to interpreting the world, remembering experiences, understanding others and communicating about the world (Ünal, Yue Ji & Papafragou, 2021). In the sequential

processes of event cognition, it is particularly relevant to the purposes of this study how we come to interpret and understand a dynamic event happening in a certain environment, putting together sensory and world knowledge information: this initial stage is referred to as *event apprehension*. For the purposes of language production, language integration is the next relevant step, in which the event is talked about after a phase of message planning.

The linguistic issues investigated in event cognition span from semantic categorisation to the influence of aspects on eye movements. Often, experiments do not involve any linguistic task but are designed as non-verbal (that do not make use of language and do not require the participants to explicitly use their native language or target L2) in order to detect whether language differences are persistently found on cognition processes even when language is not in active use. Psycholinguistics has investigated the relationship between visual processing and language, especially using eye trackers.

The last decades have seen quick improvements in the usability of this technology and its availability to institutions. Although the field is relatively new, the course of research has found robust results in the vision-language relationship: starting with Cooper (1974) and continuing with Tanenhaus et al. (1995), it has explored the links between eye movements and the reception of spoken language. The study of sentence processing with visual inspection has become an established eye-tracking paradigm known as Visual World Paradigm (among the founding woks: Allopenna, Magnuson & Tanenhaus, 1998; Altmann & Kamide, 1999; Chambers, Tanenhaus, Eberhard, Filip & Carlson, 2002; Kamide, Scheepers & Altmann, 2003; and applications: Huettig, Rommers & Meyer, 2011; Benedettini, Bertinetto & Lenci, 2019). The premise of the Visual World Paradigm is that referents organised in a display are looked at depending on a linguistic input given during inspection. The scope is to find how looking behaviour provides an indication of our decisions concerning the processing of linguistic inputs and world referents (Cooper, 1974).

Later, studies extended to language production (Bock, Irwin, & Davidson, 2003; Griffin & Bock, 2000; Bunger et al., 2016; Bunger, Trueswell & Papafragou, 2012) – possibly facilitated by improvements in the technologies of eye trackers that begun allowing head-free (or at least chin-free) mounting. In the study of sentence production, researchers ask speakers to watch a set of stimuli and provide verbalisations (usually, a description of what is seen) while registering their eye movements to keep track of the areas of the visual stimuli that are attended: where, how much, how continuously. Research has shown that eye movements can be linked to the semantic components of the scene and the description – in particular, with respect to the matching between the order of fixations and the order of the elements in the sentence (Konopka, 2014; 2019).

It has been discussed here that there was considerable semantic diversity across languages in domains such as artefact names (Malt et al., 1999), spatial relations (Bohnemeyer et al., 2014; Levinson, 2003; Levinson et al., 2002; Pederson et al., 1998), numbers (Frank et al.,

2008; Gordon, 2004), colour terms (Berlin & Kay, 1969). When looking at cross-linguistic differences within theoretical frameworks and in accounts of meaning, the question arises about whether language is a window into the mind. If so, such diversity should be observed at the conceptual level and influence representation, also implying that speakers of different languages should perform differently on specific non-linguistic tasks matching the differences in their language. The debate in this domain is heated and polarised (Zlatev & Blomberg, 2021 for an overview). Although in Cognitive Linguistics it is generally assumed that linguistic form reflects conceptualisation, experimental studies have shown remarkable evidence of asymmetries in the performance of linguistic and non-linguistic tasks. Malt et al. (1999) found that speakers of English, Spanish, and Chinese differed in how they named a set of household containers (like bottles, jars, and glasses) while showing agreement in sorting these objects based on their physical similarity. Papafragou et al. (2008) found that English and Greek speakers described motion events differently, but eye tracking showed similar attentional patterns in looking at event pictures. As Homles and Wolff (2013) put it, pieces of evidence suggest that linguistic differences in lexicalisations are not invariably salient conceptually, arguing for a dissociation between the linguistic and the conceptual level, which might be due to the fact that language is shaped by historical, communicative and cultural forces (as, for example, captured in the Motivation & Sedimentation model – Zlatev, Jacobsson & Paju, 2021).

In the following sections, we introduce the theoretical basis of eye tracking in event cognition and linguistic studies and focus on the stages of event cognition for production, event apprehension and the planning and production phase in an attempt to discuss the connection between categorisation, construals and events.

4.2.1 Eye Tracking in Event Cognition

Speech as it occurs while we look at a visual display has less reworking and reshaping that the written means. The interest for cognitive scientists is in the features of a picture that constitute the first gist of information from which the observer forms an idea of the event to describe the event later.

Eye trackers measure when and where eyes move, and in what order parts of a stimuli are fixed. Due to the eye's physical characteristics, the span of high acuity is a small portion of the visual field (the fovea vs parafovea). This means that when we read a text or examine a scene, our eyes move so that the fovea is aligned with the spots from which we need to retrieve information. For this reason, eye tracking is exploited as a window into online cognitive processes, such as understanding texts and processing a video or image. To briefly introduce this methodology, the basic assumption must be presented: the eye-mind hypothesis (Just & Carpenter, 1984). The hypothesis states that there is a strong correlation between what is being fixated is indicative of

what is being processed - under specific conditions. This statement, in fact, is not to be generalised to all situations in the real world: we all know that we can look at a speaker and think of something else. We can look at a target and elaborate thoughts or verbal expressions unrelated to it, but we cannot process information about that target without first looking at it. The hypothesis emerged from studies in reading, which showed that the duration of gaze on individual words and the number of regressions reflected some difficulty in processing (e.g., Just & Carpenter, 1984). This hypothesis, of which the limits have been pointed out (Underwood & Everatt, 1992; Anderson, Bothell & Douglass, 2004) is still the underlying assumption in eye-tracking research.

Eye trackers in linguistic research have been used to examine spoken language comprehension and production during the inspection of pictures or video (or, with mobile/portable eye-tracker) for event cognition studies. The two aspects of investigating event cognition are processing (thus, language comprehension together with a visual stimulus) and planning (for the purpose of production, upon the view of a stimulus). As discussed throughout section 3, linguistic descriptions capture different aspects of events: if we see a boy pushing a trolly, the description may focus on the agent or the patient (active or passive sentences), the path or the manner (example 14):

(14) a. The boy is pushing a trolley
b. A trolley is pushed by a boy
c. A trolley is being pushed towards the stairs

Each speaker makes lexical and syntactic choices in the way they package information. Combinations, however, are not infinite, and there are recurrent patterns in languages. Comprehension and production studies focus on two complementary processes of event cognition: integrating the knowledge given by a linguistic construal into a visual stimulus and taking information from a visual stimulus to produce a specific construal. Under the eye-mind hypothesis, the eye tracking role is that in both comprehension and production, the gaze reveals the underlying processes. Research has shown that the gaze paths in inspecting a scene for comprehension purposes or production are linked to some aspects of prediction, meaning retrieval, foci of attention and privileged semantic roles, with cross-linguistic differences. As an example, prediction effects have been investigated in the Visual World Paradigm: when speakers listen to a sentence in which an agent performs an action on a patient while looking at a display with four pictures on a screen, participants anticipate the object at the onset of the verb, against phonological or semantic competitors. Following results from experiments in the Visual World Paradigm, it appears that visual search is guided by language, as a task effect because the task is verbal (participants listen to sentences). On the contrary, these effects seem to disappear when in

real-world or world-like scenes do not show language prediction effects that are found in the Visual World Paradigm (de Almeida et al., 2019; de Almeida & Manouilidou, 2015). In picture description tasks, instead, it was found that speakers fixate the semantic roles (e.g., agent, patient, instrument, path, etc…) in the order in which they are going to be mentioned (Altmann & Kamide, 2004; Bock, Irwin & Davidson, 2004).

Linguistic production and its planning have been tested experimentally with eye tracking, especially after it became established that eye-tracking is a tool to investigate shifts of attention and since it was possible to use eye trackers in a naturalistic way, with minimal impact from the machine for the speaker. Eye movements are taken as an indication of the cognitive processing of the events by looking at the areas in the stimuli that are fixated – how much, in what order and which points in the time course of message formulation. Evidence shows that speakers systematically direct their attention to the areas of the stimuli they are processing (e.g., Flecken et al., 2015; Konopka & Meyer, 2014; Konopka, 2019).

Different semantic-syntactic configurations may guide attention and affect memory tasks and categorisation tasks, and researchers focus on which linguistic aspects may guide our attention when perceiving world scenes upon request for verbalization. Flecken et al. (2014) found language differences in encoding the endpoint of motion events and attention patterns in German and Modern Standard Arabic speakers in both a linguistic task and a non-verbal task. German speakers attended to and linguistically encoded endpoints more frequently than Modern Standard Arabic speakers. This suggests that the grammatical aspectual systems of different languages can influence the extent to which endpoints of motion events are encoded and visually processed. Papafragou et al. 2008 conducted a study involving a non-verbal memory task to test the encoding of manner and path in motion events: native English and Greek participants. The allocation of visual attention was taken as an indication of cognitive salience, but no language-related preference was found. Can it be ruled out that participants used their language to perform the task? Even though the likelihood that participants would spell out complete sentences internally before or while pressing the button may be low, it cannot be concluded that participants did not use language in any form. Papafragou 2008 concludes that conceptual organisation is independent of language-specific encoding, while Flecken et al.'s (2014) experiment concludes that it is premature to affirm so with certainty.

4.2.2 Event Apprehension

Studies in event apprehension aim at understanding how information from the world is first received – thus, how long it takes to build a general semantic interpretation of a scene. Research in this field is established by decades of testing (Biederman, 1972; Biederman, Glass, & Stacey, 1973; Boyce, Pollatsek, & Rayner, 1989; Hollingworth & Henderson, 1998).

Levelt (1989) proposed a model where speech planning begins with conceptualisation, intended as a preverbal apprehension of the *gist* of a scene. Conceptualisation is followed by later linguistic processing and, eventually, articulation of the utterance (Levelt, 1989). How do we define what the gist is, and how many pieces of information compose the general interpretation? Henderson and Ferreira (2004) discuss this issue and the related terminology: *gist of a scene* is often used in the literature with a slightly different understanding of what it includes: one suggestion is that it establishes the identity of a scene, that is, the equivalent of its basic-level category (Tversky & Hemenway, 1983).

Griffin and Bock (2000) describe the process of event apprehension with an eye-tracking experiment where participants looked at drawings of actions (e.g. a girl spraying a boy with a garden hose, a woman shooting a man) that could be described in active or passive voice. The analysis revealed that there exists a phase when the gist of the scene is extracted, and the entities displayed are associated with their roles (e.g., agent for the girl and patient for the boy) and event representation is constructed. This stage is said to be dissociable from linguistic planning. Further eye-tracking studies using picture description tasks have shown that apprehension and linguistic planning processes may overlap temporally: initial attention shifts to event participants predict which thematic role will be mentioned first in the sentence (Gleitman et al., 2007). In other words, it has been shown that the paths of fixations during apprehension reflect the later ordering of the sentence. Follow-up research, however, found that the relationship between apprehension and planning depends on a series of factors – such as the properties of the stimuli (e.g., ease of object naming) or lexical primes (Konopka & Meyer, 2014).

Another aspect of scene viewing that influences apprehension is visual salience. It was shown that visual salience and *scene gist* are to be distinguished. Visual salience (e.g., colours in a picture that attract attention) captures the first fixation (Antes, 1974; Loftus & Mackworth, 1978), while from the first saccade onwards, in a scene viewing task, the gaze is generally directed appropriately to capture the semantic information in a scene, suggesting that the viewer have already oriented themselves in the scene and could direct the eyes towards semantically relevant areas (Oliva et al., 2003).

Apprehension is thus a flexible process, adapted to top-down factors and to specific circumstances (Gerwin & Flecken, 2016; Konopka & Meyer, 2014). It has also been claimed that only those features of scenes needed for the task at hand can be perceived in the first milliseconds of event apprehension, confirming the role played by that top-down factors. These pieces of evidence show that preparing to speak has different effects on how people allocate visual attention to components of a scene depending on the task at hand: if participants are asked to talk about a specific character or the relation between characters, their gaze will be directed differently to the scene (Papafragou, 2019).

Timings of event apprehension and information extraction have been measured in eye-tracking experiments in the past decades, manipulating the amount of time a participant is exposed to a stimulus. It was acknowledged early on that specific objects and their features require more fixations than some high-level representations of a scene (e.g., a situational context), which is accessed within a fraction of a single fixation – that is, within 30–50 ms – (Dobel et al., 2007 for a review). Fast information uptake also involves living and non-living entities: living entities can be distinguished with as short exposition timing as 20ms (Delorme, Richard, & Fabre-Thorpe, 2000). Dobel et al. (2007) tested event apprehension of coherent or incoherent scenes (as depicted in Figure 3) with groups of participants being exposed to the stimuli for 100ms, 200ms, 250ms or 300ms. Even with the shortest presentation time, participants could give some information about the characteristics of the semantic roles (e.g., human, animal, etc…) and whether the scene was coherent or not. Interestingly, with presentation durations higher than 100 ms, agents were identified better than recipients, and actions were identified less accurately than those of recipients or agents (performance varied between 19% for 100 ms and 46% for 300 ms). The reason for this is linked by Dobel et al. (2007) to the fact that objects as well are less recognised, and action recognition depends on the identification of the object involved. For simpler stimuli, such as objects, scene coherence is detected with only 30ms of stimulus presentation (Glanemann et al., 2016).

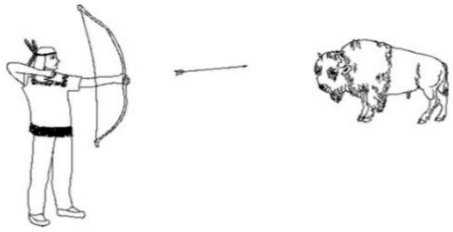


Fig. 1a. Coherent scene with agent placed left.

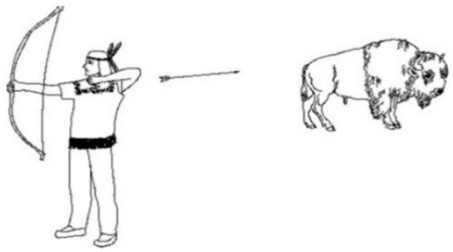


Figure 3. Coherent and incoherent scenes in Dobbel et al. (2007)

Gerwien and Flecken (2016) analysed first fixation locations and latencies constraining the amount of time participants (native English and native Spanish speakers) are exposed to the scenes (300ms, 500ms, 700ms), asking the participant to describe the scene after exposure. It emerged that the longer the time of exposure, the more specific action verbs and object naming were, with a language effect on the 500ms time exposure where Spanish speakers were significantly less specific than English speakers. The choice of languages in which the experiment was conducted, Spanish and German, follow the strand of research in (typologically) different languages (Talmy, 200; Norcliffe et al., 2015): as Spanish speakers tend to represent accidental events action, German speakers focus on the agent (Flecken et al., 2015). Moreover, the time manipulation showed that during language production under time pressure, speakers of different languages attend to slightly different aspects.

4.2.3 Message planning and sentence production

The production of a sentence is the last phase of a complex process: to utter a sentence upon the visualisation of an event, one needs to first go through the above-mentioned apprehension phase, understand what is being seen and decide what one wants to say, which has to be put in syntactic order, with lexical items, and physical organs have to be prepared. As discussed in the previous sections, deciding what to say is driven not only by the event itself but by our own conceptualisation. Thus, the interest in the phases that precede production is aimed at investigating the online processing of construal operation. The fact that different languages have different ways to describe events motivates cross-linguistic research in this area. In Levelt's

(1989) model of language production, there are two main steps involved: pre-verbal planning and sentence-level planning.

The preverbal message level is the phase in which the formulation of the message starts. It may include essential relational detail, how many characters or who did what to whom (Konopka, 2019). If we think of a transitive example, such as 'John breaks the glass', the pre-message phase consists of retrieving information about John and the glass, with their affordances, and their relationship. At sentence-level encoding, the appropriate lexical concepts are selected and given syntactical structure (Konopka, 2019). To do so, these pieces of information are assigned to words of different classes and semantic roles, according to the language spoken. As shown in section 3, given that different construals select different aspects of the events and vary cross-linguistically, these pieces of information retrieved from the same event may be put in different linguistic classes depending on the construal or the habitual patterns of the language spoken by the speaker. Sentence-level encoding starts only once a rough message-level representation is in place (Konopka & Meyer, 2014). Lastly, the production phase comprises the moment the sentence is uttered, with lexical items and phonological realisations (Dobel, Glanemann & Kreysa, 2011).

It is generally assumed that the two phases of preverbal planning and sentence planning work together. In fact, speakers rarely plan a full sentence and later verbalise it. Generally, speakers plan their messages and sentences incrementally (Levelt, 1989; Wheeldon, 2013). Research in this area has considered different aspects of these processes to put forward hypotheses about sentence production.

First, it is discussed in models of language production how, in building a sentence, the lexical items and syntactical structure are handled: is it first retrieved a lexical item, which triggers a syntactic form or is the other way around? In the first scenario (advanced, for example, by Bates & MacWhinney, 1982), the lexical retrieval comes with not only semantic but syntactic information, which serves to build the sentence. In the second scenario (advanced in Bock, 1986), syntax is accessed first, as an abstract structure and then completed with lexical items. For the purposes of this work, we can read these two positions under the Cognitive Linguistics claims (and also relevant here, Construction Grammar claims), that syntax and lexical semantics should not be taken as two parallel compartments. Syntax is, first, meaning: a Subject-Verb-Object (SVO) word order tells us at a very abstract level that the relation between two entities is such that S affects O. Verbs can also have general meanings – such as *do* or *make*. The question becomes whether, upon seeing an event (e.g., John breaks the vase) and planning a construal, speakers first decide on the relation between the participants (e.g. "someone acts on something" vs "something is being acted upon by someone") or first decides that a breaking action is taking place and then build the structure according to a specific perspective (e.g., "John breaks the vase"; "the vase is

broken"). A third way of investigating this issue, which is most relevant to this work, is taken by Konopka and Meyer (2014). To investigate production, they divide the process of formulation of a message into two parts: one is the encoding of individual participants of an event (non-relational encoding) and the other is the relationships between the participants (relational encoding). This distinction is relevant in verbalisation and non-verbalization, as it is related to the semantics of the event. Thus, it affects both the pre-message level and the sentence level.

A second aspect in the investigation of planning and production is the incrementality of planning. As mentioned above, there is considerable agreement on the fact that the preverbal and sentence-level messages interact during production (Konopka & Meyer, 2014). Individual differences exist in how the sentence is eventually produced, but generally speakers plan incrementally in small units (Konopka, 2019). Two models explain incrementality in sentence planning: on the one side the planning is viewed as progressing in small units of one concept at a time followed by one word at a time. This is known as Linear Incrementality (Gleitman et al., 2007), for which the small units planned correspond to the order of mention – which varies with different typological languages. Conversely, the gist of the scene and the relational information is planned first before linguistic encoding at the sentence level begins. This is known as Hierarchical Incrementality (as reviewed in Bock & Ferreira, 2014 and Konopka & Brown-Schmidt, 2014). These views imply that in the first case, the relational information (the event's action) is encoded after the subject has been recognised, planned, and verbalized. In the second case, relational information is processed early on, as this is necessary information to process the event participants, even those linguistically realised as the first linear elements in a sentence. Recent work has brought the attention to the role of the event in planning: the two processes depending on the task performed or the type of event inspected: different message-level and sentence-level processes may receive priority over others at different points in time (as shown in Konopka & Meyer, 2014; Norcliffe et al., 2015; Ganushchak, Konopka & Chen, 2018; Konopka, 2019; Takashima et al., 2020). Instead of positing two distinct ways of planning, there may be no one valid form incrementality determined by the architecture of the production system (e.g., Ferreira & Swets, 2002), but a flexible system which supports different strategies of planning under different conditions (see Konopka et al. 2018). In the case of an event that presents a human agent and an object, speakers of SVO languages with preferences for active sentences might be already oriented at processing the agent first and encode them in the subject position at the sentence level, which could induce a more linear planning. For example, Konopka & Meyer 2014 show that early eye movements differ between events where relational information is easy or difficult to encode.

5 Conclusion

This chapter introduced the theoretical background under which the research in this work is carried out. We stressed the fundamental aspects of cognition and language that are relevant to the study of event descriptions. For the way we understand events and manage to communicate about them, it is relevant to consider Cognitive Linguistics assumption about *embodiment*, which connects the mind and the language to what we, as humans, can perceive with our bodies and our senses, conceptualisation and meaning. It has been made explicit how this is linked to two key aspects in this work: categorisation and construal.

We discussed these two mechanisms. Categories contribute to giving meaning to events and structure experiences in the world. Construals provide the linguistic interpretation and subjectification of the events. Assuming that construals reflect conceptualisations and that lexical items form concepts, we investigate categorisation as it results from linguistic construals. Concepts get lexicalised into items, but the different languages exhibit different lexicalisations, lexical patterns, and form-meaning mappings: cross-linguistic aspects of lexicalisations across languages in a common domain will be the object of investigation of Study 1.

Cognitive Linguistics overcomes the classical theory based on logic and necessary and sufficient conditions to move to a theory of prototypes and exemplars based on the degree of category membership. The idea of basic level categories serves to organise concepts in levels of abstraction (superordinate – basic – subordinate) on a cline of generality-specificity with respect to the semantic components they pick up. Crucially, the basic level and other levels of abstraction seem to be reflected in lexicalisations, at least for discrete object categories. When moving to investigate events, it is still an open question whether they can be organised in a three-level model. We reviewed the literature about event categories to find that they have mainly been investigated as far as activities (such as EAT or RUN) are concerned in the lexical category of verbs. We proposed to take into consideration *general verbs*, which categorise a group of events which, unlike activities, do not have perceptual features in common. General verbs, such as *open* or *break*, capture a group of events that diverge in terms of action schemas performed. We show that if the linguistic realisations of activities by means of hyponyms, troponyms and hypernyms can be thought of as organised in a three-level hierarchy (with vertical but also horizontal relations), the same is not to be taken for granted in the case of general verbs. This aspect will be investigated in Study 2.

Lastly, we discussed event cognition, focusing on event apprehension and linguistic planning. We presented a brief review of research in scene descriptions showing how language affects eye movements in verbal description tasks, which is in turn interpreted as affecting planning processes, and how these aspects of cognition are investigated with eye-tracking. This step is preparatory for Study 3, which investigates the production of event descriptions. Next, we

move to reviewing the literature in the field of Linguistics on verbs that are used to describe breaking events – also referred to as *break* verbs.

Chapter 2. Events and the Breaking Domain

1. Introduction

In the domain of breaking events, the English verb *break* can be associated with the event of a vase shattered on the ground (as in intentional action or an effect of an external cause) or someone breaking someone else's arm, or as diverse actions as a cracked mirror, or a computer not working. These are all different actions, with different roles and participants, and different action schemas involved. However, they are referred to by the same verb. The verbal category of breaking is crucial in linguistic and behavioural development of children because it underlies action acquired and experienced early, being them related to the domain of object manipulation, and the understanding of the concept of cause-consequence. Research in language acquisition has shown that children start attending to linguistic and world categories (Bowerman, 2005). As discussed throughout Chapter 1, research has been engaged over the years to study categorisation and propose a model for the way we manage to structure the world's continuity and interpret it. Finding what aspects bring together a set of actions, as distinct as they are, to be referred to by one verb in a specific language is an open-ended question. In addition to these motivations for the continued interest in breaking events, we add that research in event categorisation has been mainly focused on motion, following up the pioneering work on linguistic typology (Talmy, 2000), neglecting other domains. Often, results from motion events are generalised to other domains, which may instead benefit from dedicated studies.

This chapter introduces the category under investigation – that of breaking events – and reviews the linguistic studies in this area of research. Studies on breaking events have been focused on cognitive aspects (psycholinguistics and neurolinguistics) and linguistic aspects (from syntax, argument structure, frequencies in corpora, and translation equivalences). Breaking events actions belong to the universal experience of humans and go a long way back in our history (Bowerman 2005): they are related to objects, which we daily deal with, and they are experienced early by children in their first encounters with basic elements of the world (such as the effects of gravity, force, access to contents).

We start framing the notion of event in linguistics and specifically discussing action verbs – the linguistic material mainly involved in describing events. With a brief discussion of aktionsart and verbs' classification, we stress the classes of verbs that concern breaking events (section 2).

We then present the breaking domain and its relevance in language acquisition (section 3). In section 4 we take into consideration the literature about semantics and syntactic components of verbs describing breaking events in different theoretical approaches. The interest in the category of breaking emerged from a tradition of studies that investigated syntactic patterns – such as Fillmore (1970) on the alternations of the verbs *hit* and *break* and Levin's (1993) with verb classes – or in event structures and continued in constructionist approaches, Frame Semantics, and Cognitive Linguistics studies. A cognitive-based interest was raised with a large cross-linguistic investigation on semantic categorisation (Majid et al., 2007), followed by many studies (Section 5) that investigated *cutting and breaking* events. The interest in cognitive studies has been driven to the cutting and breaking category (Majid et al., 2007) based on a conceptual selection of an area of investigation: Majid et al. (2007) discuss events of separation as events of change of state in which object integrity is materially separated into pieces. These events are furtherly distinguished into reversible and irreversible separations – the latter being the ones including events referred to with the English verbs *break, cut, rip, tear snap* (etc…). Since this study concluded that English distinguished two main clusters in the categorisation of irreversible separations – *cut* and *break* – this became an established domain for the investigation. Here, we operate a slightly different selection of domain (Section 6) to focus on a more restricted set of events.

2. Events, actions, and action verbs

The notion of *event* is a core element in cognitive sciences and linguistics as well – in particular, in analysing the meaning of (action) verbs. The term has been used in the philosophy of language and linguistics in all fields of inquiry – from Generative traditions onwards. Generative semantics have developed models for representing verb semantics by means of event structure components, in which the verb meaning is expressed in a propositional form, while Cognitive Linguistics mainly focus on the semantics of verbs and their role within the expressions referring to actions, assuming that construals reflect cognitive strategies. In Cognitive sciences (see Chapter 1, section 4), events are defined as a complex of participants, activities, and outcomes brought together by causal interrelatedness" (Elman, 2009: 572) or as segments of time conceived as meaningful (with a beginning and an end) by and observed (Zacks & Tversky, 2001).

In linguistics, the basic understanding of what events are derives from the philosophy-related works of Vendler (1967) or Davidson (1967). Events are described in contrast with *facts*, where the former are characterised by their relationship with time. In Event Semantics (since Davidson, 1967), events have played a primary role in understanding the links between the world

and linguistic forms. For Frawley (1992) and his work in Linguistic Semantics, it is fundamental to model the world's knowledge from *things* to *actions* to the *relation of things to actions. Relation* is the key term that motivates a verb-based study of events: being events complex conceptions unfolding in time and linking together participants, their relational nature appears as crucial. Verbs generally capture with lexicalisations this relational nature (Croft, 2012; Peelen, Romagno, & Caramazza, 2012; Shapiro, Moo, & Caramazza, 2006; Gentner, 1978).

In Vendler's *linguistics in philosophy* (1957), events, processes, and actions are all understood as occurring, taking place, beginning, lasting and ending. They "can be watched, heard, followed, and observed; they can be sudden, gradual, violent, or prolonged" (Vendler, 1967: 141). This differs, for example, from atemporal facts. In fact, Vendler argues that events, actions, and processes are distinguished from objects first and foremost in the dimension of time: only events unfold in time and are affected by the dimension of time. They are temporal entities, while facts are not in space and time (Vendler, 1967: 145). Another aspect that emerges as relevant in actions and events is that of *movement*: an activity or event that can be performed should involve some kinetic energy. At the conceptual level, movement is not a necessary and sufficient condition in the conception of events: in fact, we do not have concepts for meaningless movements, but we do have concepts for mental actions which lack a physical motion. In the following sections, we investigate these aspects that concern verbs and events, starting from the conceptualisation of actions from a neurological point of view to the linguistic encoding and, lastly, to the categories and semantics of action verbs.

2.1 Actions and conceptualisations

Before focusing our discussion on verbs and linguistic items, we briefly discuss the status of action concepts in neuroscience to highlight the tight link between participants of events in the creation of action concepts. As mentioned in the sections above, actions are basic elements of what has been defined as events, and they are generally expressed by verbs.

Our ability to recognise an event denotes a step forward from simply recognising a referent by means of perceptual capabilities. It requires making appropriate inferences between participating entities and contexts: relations between agents, patients, instruments, or targets and our cognitively mediated interpretation of them account for conceptualisations. Recent studies in the neurosciences have explored the neural basis for action concepts (Watson et al., 2013; Vigliocco et al., 2004; Pulvermüller, Shtyrov & Hauk, 2009) and if, on the one hand, actions are more complex than objects (relying on relational information) the fact that many human artefacts have physical features which are clues to how to interact with them (tools, in particular) point to

the fact that, at the conceptual level, the distinction between object and action may not be primary (Leshinskaya, Wurm & Caramazza, 2020). However, the organisation of action-related categories in the brain is still largely unknown. Wurm, Caramazza and Lingnau (2017) specifically address action categories in a non-verbalization experiment and whether they follow an organisation on the cortical surface similar to the category-specific organisation of object knowledge. The result of their study shows that actions in the lateral occipitotemporal cortex are represented on the dorsal area when they are based on people interactions (in proximity to person knowledge with actions, such as *teaching*) and on the lateral parietal area when the action involves objects interaction (in proximity to the area processing tools/artefacts, with actions such as *peel* or *cut*), suggesting a mutually dependent organisation of actions and objects. Huth et al. (2012), show that the brain maps object and action categories in a smooth, gradient continuum, despite the conceptual differences in objects and actions. Neuroimaging has thus identified two important loci for action concepts: part of the lateral temporal and lateral parietal cortex are proven to represent conceptual content about actions, but they also reflect object knowledge about tools (Leshinskaya, Wurm & Caramazza, 2020). Additionally, Watson et al. (2013) found a functional area in the (posterior) middle temporal gyrus which increases its activation when participants name actions that correspond to pictures or names of tools, showing that the neural representations of action concepts are intertwined with representations of objects (tools, particular). Representations about actions and tools are thus closely entangled in neural space, as expected under an account of action concepts composed of verbs and the objects they occur with – in linguistic terms, their thematic roles. Verb-selective responses are preserved in blind participants (Bedny, Dravida & Saxe, 2014) and range over different verb types (motion verbs in Bedny et al., 2012; Romagno & Caramazza, 2012; Bedny, Dravida & Saxe, 2013). It has been shown that exposure to verbs that refer to actions, or nouns that refer to tools can activate related motor areas in the brain (e.g. Preissl et al., 1995; Pulvermüller, Lutzenberger & Preissl, 1999). Panzner, Gaspers and Cimiano (2019) propose a model of language acquisition that views language constructions and action concepts as emerging simultaneously: concepts are evoked and first emerge upon encounter with a linguistic construction. The model is reconcilable with the mapping paradigm according to which novel words or constructions are mapped onto existing, priorly acquired or innate concepts. This view is related to grounded cognition based on embodiment (discussed in Chapter 1 as a tenet of Cognitive Linguistics) and supports a certain degree of linguistic relativity in its claim that lexical items form categories and influence our conceptualisations. If we accept the view delineated above of a concurrent emergence of linguistic structure and action concepts, we can see language as contributing to triggering the categorisation of the experience.

Research in neuroscience offers evidence of the relationship between objects, tools, or living entities. Action concepts are not retrievable from a specific area but interact with event

participants. This is relevant for the purpose of this work in two aspects. First, it highlights that verbs should be examined in close context with the participants of an event (section 2.3 addresses this issue). Second, since the processing of actions and verbs is entangled with event participants, we need to bear in mind that in an experimental approach aimed at understanding the processing of verbs, it is challenging to isolate the processing of actions alone (this issue is discussed in Chapter 5-Section 2).

2.2 Towards linguistic action encoding

The research interest is thus in how events relate to verbs, and so far, we have associated events with verbs. Pedersen and Bohnemeyer (2011) report that defining the verb as a lexical category with the semantic property of describing events is a tradition in Europe that dates as far back as the Greek grammarian Apollonius Dyscolus, even though explicit theories of event semantics would only be developed in the twentieth century. However, this relation is not straightforward: Vendler (1967) noted that the world category of discrete entities is different from actions and events but that there is no direct matching between a world category and a linguistic class. On the same lines, discrete entities are associated with nouns. Vendler exemplifies this idea with the nouns *fires* and *blizzards:* they are nouns, but unlike other nouns which denote objects (*tables, crystals, or cows*), the entities these nouns denote can "occur, begin, and end, can be sudden or prolonged, can be watched and observed - they are, in a word, events and not objects" (Vendler, 1967: 41). Thus, the equation we have made about verbs and events is a generalisation which should not be taken as universal. For the purpose of this work, verbs are the main carrier of action meaning as defined by the temporal and relational dimensions and most relevant for the category under investigation. For this reason, we centre the following review on verbs while still considering the role of event participants.

2.2.1 Aktionsart

The term *aktionsart* refers to the inherent qualities of a situation from the perspective of the events, rather than the speaker. As such, it is not relevant for verbal categories, but it is rather a semantic category: it refers to the ways a situation is carried out in its unfolding in time (Comrie, 1976). Following the Fregean tradition, the temporal development of events has been the focus of interest: the time an event takes to occur is its *durativity*. For example, in *John built a table*, there is a starting point and an ending point of the event; what is in the middle is the duration of the event. In a stative event, there is no beginning or starting point: *John is tall* is not linked to a temporal line. In the conceptual tradition, situations are categorised according to a logical

structure. Actions are durative (e.g., *read*) and changes are instantaneous (e.g., *break*). This conceptual categorisation determines its temporality.

Vendler (1957) proposed a four-way classification (summarised in Table 1), which has been a landmark in linguistics for the past decades:

1. States

   These events lack a change in time. We can mention 'John knows English' or 'John is tall'. The role of the grammatical subject is generally that of experiencer rather than agent.

2. Activities

   For activities, there is no predetermined duration in time (the action could go on indefinitely) but it usually composed of sub-actions or phases. These could be termed part-whole relations in Cognitive Linguistics. In Chapter 1-Section 2.5 we exemplified the sub-actions that compose an event of eating. Consider 'John is eating pasta': activities are characterised by the fact that at any given time during the event of John eating pasta, if John stops eating it could be said that John *has eaten*.

3. Achievements

   In achievements, there is an instantaneous change or transformation. Examples are 'John has arrived' or 'Sam recognised Laura'.

4. Accomplishments

   Accomplishments have a temporal duration and a limit. They indicate that an event has taken place through a certain amount of time and reached some completion. As an example, 'John built a house'.

Event	Telicity	Duration	Change	Example
STATE	-	+	-	*John knows English* *John is tall*
ACTIVITY	-	+	+	*Bill pushed the cart* *John is eating*
ACHIEVEMENT	+	-	+	*John has arrived*

				Sam recognised Laura.
ACCOMPLISHMENT	+	+	+	*Harry broke the toy*.

Table 1. Classification of lexical aspect in Vendler (1967)

Table 1 summarises their characteristics: states are the only class without a component of change, achievements are the only class without duration. Lastly, states and activities are atelic, while achievements and accomplishments are telic. The semantic interpretation of these classes is operated in terms of truth conditions of each predicate in a context. Dowty (1977; 1979), instead, relates temporal classes to the conceptual ontology. He adds an operator to the formal representation of the verbs (Table 2) to capture their entailments.

Operator	***Aspectual class***
Do	Activity
Become	Achievement
Cause-Become	Accomplishment

Table 2. Classification of lexical aspect in Dowty (1979)

This view can also be retraced in Chafe (1970). His original discussion is proclaimed by himself as tentative. However, his insights are relevant for the purpose of this work and for the studies it prompted. He proposes a first distinction between *states* and *non-states*. Non-states are distinguished from states because they are "happening" events – they answer the questions 'what happened, what is happening'. Within *nonstates*, Chafe distinguishes processes ('The wood dried', 'the rope broke'), actions ('Mary sings'), and action-changes ('Mary breaks the plate'). We see here that Dowty's *cause-become* is similar to Chafe's *process-action*.

Category	Question	Example
STATE		The fruit is dry The vase is broken The door is open
PROCESS	What happened to X?	The fruit dried The vase broke The door opened
ACTION		Mary sings Mary runs
PROCESS-ACTION	What did Y do to X?	Mary dried the fruit Mary broke the vase Mary opened the door

Table 3. Chafe (1970) classification of states and non-states

The perspective taken by Comrie (1976) can further contribute to this discussion. Concerning the lexical aspect, he points out three major characteristics of what he refers to as *situations.* The first is *durativity* and the second is its opposing concept of *punctuality*. To this, he adds the idea of *semelfactive*: actions like *cough* or *sneeze*, which are generally said to be punctual, can be categorised as very short events that end with the entity finishing the action in the same state as when the action begun. One cough is a semelfactive event with *some* durativity: if we imagine seeing a one-coughing event in slow motion, we understand better that is has its evolution through time (Comrie, 1976: 43). This event is different from an iteration of coughs – for which the durativity is more evident. Thus, Comrie challenges punctuality in its strict conception if it is related to real-world events and how we can investigate them with technologies (e.g., seeing them in slow-motion). In these terms, we can think about the single *cough* as an event made of sub-events that must happen in a certain succession (involving air passing in and out of our body and movements of vocal cords, lungs, epiglottis, and rib muscles). This way of thinking about a single cough is akin to the description of the activity of eating given above. However, the coughing process is not consciously performed *as* an activity and is irrelevant from an actional point of view – we do not really decide how to cough, and we do not see the subevents as they happen. As Comrie suggests (1976), despite the challenge to the notion of punctuality in world events, it remains a notion that seems relevant linguistically, as captured by verbs. To bring the argument further, it could be argued that the linguistic relevance is not arbitrary but reflective of something in the world that is relevant to our culture and/or perceptions: this means that it does not matter that in physics, all events have *some* durativity, as far as what is relevant for our lives, events can be conceptualised as punctual.

The second aspect is telicity, and it is relevant to the fact that Comrie (1976) discusses it in terms of events and verbs separately. A telic situation ('John is making a chair') is different from an atelic situation ('John is singing') and verbs cannot be given labels of being telic or atelic as their meaning changes with respect to telicity ('John is singing a song' becomes a telic situation): "situations are not described by verbs alone, but rather by the verb together with its arguments (subject and objects)" (Comrie, 1976: 45). It is crucial that subjects are included here as well as objects: it points towards a wholistic account of events, defined by all participants, as well as actions. This view is supported by experimental research and neuroimaging, which shows how our retrieval of action concepts involves the participants who can afford those actions (as discussed in section 2.1).

Last, he distinguishes states from dynamic situations. A state is durative (it is acknowledged in its existence during a period of time), while dynamic situations can both be durative or punctual or semelfactive.

The different perspectives on aktionsart presented in this section show common aims: capturing aspects of events in their temporal dimensions (with durativity), their effects on participants (such as changes or transformations) and their relations (such as causality). It seems relevant to investigate event descriptions and point out the mismatch between the world's events and the linguistic realisations: examples such as punctuality show that we operate interpretations of the world to create meaning and categories. Lastly, it is relevant to consider verbs together with the participants of an event. As observed by Chafe (1970), to some degree anecdotally, the relevance of the internal coherence in the verb-patient configuration is reflected in the fact that idioms are often realised as verb-patient configurations (as in 'break the ice').

2.3 Action verbs

This discussion leads to considering verbs and how they describe the classes mentioned above. As we read in Table 3, the same lexical item (or verb root) can have different aktionsart – and be part of different classes. There is not an absolute relationship between a verb form, and the class it can appear in: an example from Dowty (1979) is that the concept evoked by the linguistic form *dry* has some characteristics which can be expressed by visual properties such as "lack of water content", "wrinkles" etc…, which can be expressed as a state or a process ('the fruit is dry' vs 'the fruit dried*').* In the case of *break*, it belongs both to processes and process-actions. This is because the aktionsart is not a property of the verb itself but rather of how the verb describe an event. And verbs are flexible enough, thanks to linguistic devices (such as tenses and aspect, but also verb-

external items such as modifiers) and non-linguistic devices (such as pragmatics or context) to denote different aktionsart. Concerning modifiers, consider examples (1-2).

(1) He has been running the whole afternoon.
(2) He ran a mile today. (Dowty 1979)

Run is a typical activity verb (composed of sub-events, a specific motion of the feet, that does not require a culmination) but when modified by an adverb or another modifier restricting or limiting the activity (e.g., 'the whole afternoon', 'a mile'), the action described becomes an accomplishment. The modifier *a mile* defines the terminal point of the process and makes it telic (Dowty, 1979: 60).

Moreno-Torres (1994) reasons on the temporal aspects of situations, saying that they are necessarily linked to change: if there is no temporal evolution there cannot be change, and vice versa, if there has been change, there has been a temporal evolution. Thus, states can be considered a-temporal situation that captures the world as it is, without the component of time. A stative verb describes an entity with respect to a property that is not changing. Anything else is a change. However, as *change* is a term used in the field of breaking events, we need to distinguish *change* as used here – which applies to anything that is not a state and takes the perspective of any type of action involving some sort of modification (if you run, you change your initial state) – and *change of state*, which further specifies that an action has some sort of effects in the final state of the action (when you stop running, you return to your original state; when you break a vase, the object is in a different final state).

A further distinction is made between verbs denoting activities and general verbs (Moneglia, 1998). This distinction partly retraces Vendler's classification: activity verbs fall into the activity class, while general verbs merge accomplishments and achievements. The distinction is operated on an approach that is defined 'a la Wittgenstein' and uses explicit instantiations of the actions that underlie the use of a verb. General verbs refer to events that are very different from each other if we consider visual perception properties: so different that they do not seem connected to each other. General verbs do not identify a specific activity in their semantic extension (Moneglia, 2005). The "a la Wittgenstein approach" refers to the idea of family resemblance as described by Wittgenstein (1953) with the game example (see Chapter 1, section 2). If we look for similarities of different referents, which we call *games*, we only find a net of similarities that are intertwined but not necessarily and in sufficient conditions for which they are all called *games*. Moneglia adapts this reasoning to *general* verbs. Activity verbs refer to events occurring with a set of visually recognisable action schemas – such as *eat* (ingesting food, putting it into the mouth, chewing and swallowing), *run* (moving the legs one after the other in a specific manner), *sing (*emitting sounds from the mouth in a musical way), while general verbs are not.

Two main properties are discussed in Moneglia concerning differences between activity verbs and general verbs. First, the former, but not the latter, can be used without a theme ('Mary eats' vs *'Mary opens'). Second, the former, but not the latter, have a gestalt configuration (the parts that compose the event are not as relevant as the whole: consider eat and the subevents take food to the mouth, chew, swallow). Consider instead the case of *open:*

- Open the door
- Open one's eyes
- Open an umbrella
- Open a bottle
- Open a can of tuna

All these are events occurring with the performance of different actions, and they do not have part-whole configurations. They are not performed with the same succession of actions (as in the case of *eat*): even considering the first of this list of events, requires different actions – for some doors we grasp a handle and push it down, for others we execute a sliding motion. Some of these events are similar because they invoke the frame of access to content; others can be linked to this frame because they refer to *access to the external world*. There remains that these two interpretations do not capture the entire meaning potential of *open*, and the referents of open are a class of events that do not seem to share visual properties. The *Wittgensteinian* question emerges here: how can you explain what *open* is?

The solution proposed by Wittgenstein ("This and similar things are what we call games" – Wittgenstein 1953) is adopted by Moneglia (1993; 1998) to discuss actions. However, this solution is still vague and operates on risky territory: boundaries of similarity in a continuous world and conceptual-linguistic levels. Concerning boundaries of similarity, as discussed in Chapter 1, it has been seen as problematic to find sharp boundaries based on similarity in categorisation theory. Opening a kitchen door is similar to opening a front door. Is it similar to opening a train door or an aeroplane door? The categorisation must operate on higher level conditions, more akin to the superordinate level (Chapter 1-Section 2.3). Concerning the distinction between conceptual and linguistic levels, we need to recognise that what I see and conceptualise as similar according to my sensory experience might be different from the form-meaning mappings in languages. The events that *open* describes are not necessarily described in other languages by an equivalent verb with a one-to-one relationship. As different languages segment realities in different ways, we should at least take into consideration that the similarities between 'games' are not objective nor necessarily imposed by the world but conceptual, dynamic and flexible.

To sum up, the concept of *event* has a crucial role in understanding the meaning of action verbs and in forming concepts about actions. Having reviewed different definitions of events and different classifications that consider the events' semantics and the linguistic items that describe them, we find the recurring points that seem to be firmly relevant for linguistic investigation. First, it is evident that the fundamental characteristics of events are their relational nature and their temporal dimension. The temporal and relational nature of events is confirmed by neurolinguistic and cognitive studies. Regarding actions, the most central linguistic category seems to be that of verbs – but this should not be taken as a necessary and motivated condition, and it should be kept in mind that verbs interact with the action's participants. However, exploring the relation between events and action verbs (however partial) further emphasises the role of flexibility in verbs, which denote different aktionsart and categories. The distinction between activity verbs and general verbs is considered relevant for the purposes of this work. It will be referred to in the following sections and addressed in this work's empirical and experimental chapters.

3. The breaking domain

Break most prototypically denotes a fragmentation whose consequence is the non-functionality of the affected patient. By extension, it denotes only non-functional events, even if the fragmentation is minimal (if a cup is cracked and cannot hold a liquid, the cup is said to be broken) or non-visible (a washing machine can be broken even if externally intact). Different languages have different ways in which events are labelled or referred to with predicates. As mentioned in section 2.2, Mandarin has different categories than English *break.* It is thus necessary to establish the grounds for linguistic investigation of a conceptual domain. If the world unfolds continuously and our conceptualisations form events, we, as researchers, will have to operate a segmentation of reality to capture a domain for analysis. This operation is an approximation but it a necessary step to investigate and aim for reproducibility of a study.

The main feature that has been considered in the literature concerning the *breaking* category is that it relates to object manipulation actions and that it produces a material separation. This is valid for the English *cut*, as 'cutting a cake' results in a separation of pieces as well as *breaking a vase*. For this reason, *cut* and *break* have been discussed together in the domain of CUTTING AND BREAKING. This is a common trait that relates to perception and to how these events share characteristics. However, the verbs *cut* and *break* have been discussed together since

Fillmore (1970) and later in Levin (1993). Moreover, it has been shown in the literature how children overextend events that result into a material separation in pieces, and this gives ground to discuss this as a conceptual domain that is psychologically relevant. However, as discussed in Chapter 1-Section 4, events are not defined in perceptually relevant terms alone but are built upon intersubjective and sedimented concepts built on higher-level operations, such as goals or function.

Certain events of *opening* create a separation in the object (opening a container by removing a lid), but these are not breaking events. *Opening a container* and *breaking a vase* are distinguished in the literature (Majid et al. 2007, see section 4 for a detailed discussion) as event types by the fact that the first is reversible (one can easily put the lid back on and reverse the action that has just been performed) while the second is not: putting together the pieces of a vase is possible (requires materials, expertise) but not in the same reversible way as with the container. Thus, breaking and cutting events can be taken as a separate domain for their saliency in the irreversible separation, as opposed to other types of separations (e.g., *open*). The loss of the original integrity of an object may be determined by different scopes or results: when cutting, one may do it purposefully to satisfy a distribution scope, as in 'cutting a cake'.

Bowerman (2005) describes English *break* as prototypically related to the material destruction of rigid objects but it is also used for some flexible objects (e.g., a rope) while sounding unnatural with others: a sheet of paper, a blanket or, as suggested in the title of her seminal work, *a cooked noodle*[5]. Flexible objects are usually referred to in the event of destruction with the verb *tear* (or *rip*), which is characterized by an added manner implication: bit-by-bit separation (Bowerman, 2005). The same object, for example a fabric bag, can be described in an event of tearing if the action is gradual or breaking if the action occurs in a sudden movement – Bowerman exemplifies this with the sentence "the heavy groceries broke the shopping bag". It is evident from this example that the use of a verb to describe an event goes beyond visual perception or sensory information (the flexible material): in this case, the bags have a functionality which is lost because of the disintegration of the bags. Bowerman reports that English *cut* distinguishes from *break* the object manipulations which take place thanks to a sharp instrument. This distinction is typical of other languages, but the boundaries between cutting and breaking are not universal. For example, Sranan (Essegbey, 2007) distinguishes between a clean break (koti), regardless of the instrument, and a messy, scattered break as that we can imagine is produced by a pot broken with a hammer, (*broko* 'break a hard, brittle object'). *Cut*, instead, prototypically evokes a knife or a scissor, but any sharp edge instrument may be used: cookie cutter or piece of broken glass (Bowerman, 2005). In addition, it is necessary for a cutting event to be caused by a force and not be brought about by

[5] Referencing Bowerman's seminal paper "Why can't you "open" a nut or "break" a cooked noodle? Learning covert object categories in action word meanings" (Bowerman 2005)

a blow on the instrument. From this, it emerges that *cut* necessarily requires an agent who volitionally initiates the action. Instead, agentivity implications are more nuanced when considering breaking events: the event can also be described inchoatively as in 'the vase was broken'. Further syntactic and semantic discussion is given in section 3. A last point to consider when introducing breaking events is the aspect of goal and functions when the material destruction does not involve an external, visible separation but functional damage. Irreversible separations are, in fact, typically combined with the idea of non-functionality (as in the grocery bag example above).

3.1 Break verbs in language acquisition

World categories and linguistic categories are acquired early and simultaneously by children: one does not prevail over the other, and they both work together to form conceptual structures. In the following section, language acquisition studies are reported to elaborate on this aspect of the emergence of concepts.

Bowerman (2005) reports a conversation between an almost 8-year-old child and an adult, talking about objects contained in a box – broken toys and fabric scraps. The child asks what is in the box, and the interlocutor replies that there are other broken things. The child says: "I wouldn't call felt broken, I would call it ripped". Knowing the usage of English verbs, we know that this is what we generally do: we use *tear* and *rip* for fabric and soft/flexible material and break for rigid objects. The questions that arise are how children eventually reach this knowledge and why is it that soft, flexible objects are ripped, and hard objects are broken? The answers to these questions relate to what has been dealt with so far about categorisation, events, and linguistic mapping.

Children acquiring the characteristic aspects of events must be aware of world categorisation combined with language categorisation. In the case of the breaking domain, cross-linguistic research reports that K'iche' Mayan speakers must be aware of categories that English children are instead ignoring, but in acquisition breaking is overextended following language-specific patterns (Bowerman, 2005). Pye (1996) investigated the category of breaking in English, Mandarin, and K'iche' Maya to find that English children overextend *break* far more than Mandarin and K'iche' Mayan speakers. For example, English children extend *break* to tearing paper (56%) while this did not happen in their data for Mandarin children. Bowerman (2005) explored the motivations for overextension: The likelihood for English children to overextend *break* is related to the coherence of the associated *covert object categories*. The acquisition of fine-grained

differences seems to occur sooner than general categories such as those like *break* or *open*. Erkelens (2009) report that Dutch children rarely make errors with more specific verbs, such as between single-blade cutting (usually with a knife) and double-blade cutting (usually with scissors). General categories comprise perceptually diverse objects, and more experience is needed with all of them before learning how to discern them. Both children and adults often overextend *break* to flexible objects or to reversible separations (e.g., 'opening a can'), while *cut* may be overextended to actions involving instruments with no thin linear edge (e.g. crushing ice). Overextensions unveil what the attention is on: saying *cut the ice* instead of *crush the ice* shifts the focus on the separation result rather than the manner in which the action comes about. Overextensions seem a relevant part of language acquisition and, later, *usage*, as a testament to the dynamicity of language and flexibility of categories: in the view of semantic taken here, *crushing ice* is a collocation but there is not intrinsic value in the verb or the event that makes it necessary to denote such event.

The covert object categories associated with *break* and *cut* encompass a vast range of perceptually and functionally diverse objects and, therefore, necessarily take children more time to refine their categorisation (Bowerman, 2005). In Pye et al.'s (1996) language production task, 25% of the English-speaking children said *cut* for dividing play dough with a pencil and none of the adults uttered the same. An average of 45% said *break* across five actions of tearing paper but not adults. This means that, perceptually, there is something similar in the two actions (dividing the play dough with a pencil and cutting something with a knife), but language operates a conceptual structure that goes beyond the visual aspect. Within the category of action verbs, general verbs are a high-frequency class of verbs that extend to actions that are perceptually very different. For this reason, it takes longer to categorise all objects that can be associated with a general verb, such as *break*. General verbs are also more central to a child's daily life: *open* or *break* are verbs consistently present in the first attempts at manipulating objects and are thus learned early. In Pie et al., children and adults were asked for their intuitions about the meaning of *cut* and *break* when shown stimuli made up of series of events that ranged over types of actions (cutting with hands or cutting a ruler), instruments (hands, rulers, scissor, a string, a pencil) and results of breaking and cutting. The study features objects such as Play-Doh, peanuts, crackers, paper and dental floss. The results showed that English-speaking children and adults used *break* to refer to the "prototypical action of breaking a toothpick by hand" and that a considerable percentage of children used *break* to describe the events that involved paper (in contrast to answers given by adults, who used *tear*).

Goldberg (2010) discusses how overextensions are learned and categorisation is perfected in the acquisition phase. Goldberg (2006) argues for indirect negative evidence available for children to acquire the correct categories in language. From a constructional point of

view, Goldberg considers the principle that no two constructions are semantically and pragmatically identical: assuming that every construction is unique, there must be contexts in which a certain construction is preferred over others with a close meaning but different implications. The amount of information children are exposed to allows them to infer which is the preferred construction among two competing forms. If the same situation is repeated several times, the child's inference becomes stronger and eventually leads them to unlearn an over-generalisation. To give an example, the ditransitive construction alternates with the prepositional paraphrase based on the information structure of the sentence. When the recipient is non-focused, and the transferred entity is focused, the ditransitive construction is the most preferred; in the reversed situation, the prepositional paraphrase is the most preferred.

(3) a. Mary told Jane a secret.
b. Mary told a secret to Jane.
c. Mary whispered a secret to Jane. (Goldberg 2010)

With verbs that can occur in both constructions, as *give*, children exploit this pragmatic difference. In the case of *whisper* from (4c) above, the recipient is not focused, and the child expects it to be well-formed. When the child encounters the prepositional paraphrase instead, they will infer that *whisper* is not a possibility in the ditransitive construction and unlearn the overgeneralisation. Goldberg's conclusion is that children make use of an enormous amount of stored data to first overgeneralise and then, thanks to recurrent indirect negative evidence, narrow down the generalisation (Goldberg, 1993; Goldberg, 2019). Nonetheless, Goldberg (2010) believes children exploit multiple sources of evidence to acquire language and multiple cognitive mechanisms. This constructionist view is reconcilable with the studies on semantic categorisation and the acquisition of action verbs investigated by Bowerman and Pye, as they both point against universal language acquisition theories. Pye, Loeb and Pao (1996) claimed that in the acquisition of verbal categories children first form a "universal concept of breaking", but the acquisition is not universal, as cross-linguistic comparison shows (e.g., English-speaking children used the general verb *break* for a wider range of actions than Chinese and K'iche speakers). Considering their results, Pye, Loeb and Pao (1996) argue against referential semantics and (following Labov 1973) feature analysis. Features are a set of necessary and sufficient conditions for which linguistic items have a certain meaning, but "in reality, each feature can be overridden by contextual factors" (Pye, Loeb & Pao, 1996).

Hsiao (2015) provides a corpus analysis combined with a categorisation task, aiming at investigating the role of FORCE evoked when native speakers categorise cutting verbs and how

the idiosyncratic components (FORCE) of verb meanings constrain the combination of verbs and semantic roles in Mandarin. Mandarin does not have a pure binary distinction of breaking and cutting events. It relies on using complex predicates (e.g., bipolar verbs). Features relevant to cutting verbs in Chinese are instrument, manner and features of the affected object (softness vs hardness, flexible vs rigid). Their experimental task was performed with 36 participants who were asked to first categorise verbs (10 cutting-related verbs and 5 *break* verbs, plus fillers) based on their intuition and provide their sorting criteria. After, participants were asked to select which verbs are related to strength (*force*) in a descending order and to provide a list of everyday tools that they associated to the performance of the action described by the verb. It resulted that 60% of the participants sorted verbs into 4 categories: force, separation, precision actions (e.g., engrave vs carve), and instrument. The results of the ranking test showed that all selected verbs are related to FORCE but differ in degree. It emerges that force, result and instrument are features affecting the way native speakers categorise verbs, with a group of verbs (*kǎn* 'chop', *duò* 'chop' and *pī* 'hack') implying the events of separation with physical strength, whereas another group (*diāo,* 'engrave' and *kē,* 'engrave; carve') are defined by control of force.

All in all, experiments in language acquisition suggest that when children start speaking, they have already acquired linguistic categories even if not worked out completely. Thus, children acquiring different languages will show different overextensions. Experiments from the literature show that children sometimes acquire a fine-grain difference (e.g., cut with a single or double blade) before general categories such as break. The general verbs are learned, but their categorisation is not completely mastered. Bowerman (2005) observes that children broadly generalise the use of an action verb when they register that a word is used with perceptually diverse objects but act in a more conservative way when two closely similar objects are used with different verbs (as in the case of Dutch and cutting with a single or double blade). Children employ and develop categorisation skills during this process, including noticing and comparing objects in the world and constantly updating their knowledge. As the world they are experiencing is coded differently in different languages, they will add to their knowledge of linguistic categorisation.

4. Linguistic studies on cutting and breaking

Form-meaning relations have been investigated in the breaking domain with respect to the verbs' semantics and argument structures they participate in. The early accounts in this domain discussed *break* in contrast with *hit*, starting with the seminal work by Fillmore (1970) *The grammar of hitting and breaking*. Hale and Kayser (1987) later contributed a syntactic account of

middle constructions, and Levin (1993) provided ground-breaking work in the field with verb classes. In terms of languages investigated, the number of contrastive linguistic studies and variety of languages is remarkable: we have already mentioned English, and the list continues with French (Bouveret & Sweetser, 2009) and Spanish (Spalek 2015), Arabic (Hassan 2015), Greek (Stathi 2023), Turkish (Stathi 2023), Polish (Goddard & Wierzbicka, 2009), Japanese (Huisman, Van Hout & Majid, 2021; Fuji, Radetzky & Sweetser, 2013; Goddard & Wierzbicka, 2009), Korean (Kwon 2016). More than this, Majid et al. (2007) cognitive-linguistic study took into consideration 28 languages: one Altaic language (Turkish), two austronesian (Biak, Kilivila), one Caribbean (Tiriyo´), one Dravidian (Tamil), seven indoeuropean (Dutch, English, German, Hindi, Punjabi, Spanish, Swedish), two mayan (Tzeltal, Yukatek), three niger-congo (Ewe, Jalonke, Likpe) one Otomanguean (Otomi) one Pama-Nyungan (Kuuk Thaayorre), one Sino-Tibetan (Mandarin), one Tai (Lao), one West Papuan Phylum (Tidore), one Witotoan (Miraña), one Creole (Sranan), two isolate (Chontal, Japanese) and two isolate Papuan (Touo, Yélî Dyne). As the interest in this semantic domain for the Western tradition of linguistics studies started from English, the literature review will be dominated by English-based considerations – being one language in our investigation sample as well. In what follows we first review the classical linguistic studies in the breaking domain (occasionally including neighbouring categories).

4.1 Fillmore (1970)

In analysing verbs of breaking and hitting, Fillmore delimits its analysis to their 'basic, non-transferred meanings' (p. 12). The classification is not merely semantic but rather syntactic or grammar driven. The first distinction made is in the grammatical structures in which these verbs are used according to the verbs constraints on the NP's they select:

(4) a. The stick broke
    b. John broke the stick (with a rock)
    c. A rock broke the stick
    d. John hit the tree
    e. A rock hit the tree

Break shows transitive and intransitive uses, which *hit* does not. In the first part of his analysis, Fillmore individuates three separate meanings for *break* and two for *hit*:

- Break-1:

Example: *the stick broke*

Definition: intransitive verb, asserts of its subject the particular deformation or change of state associated with the meaning of the verb

- Break-2:

Example: *John broke the stick (with a rock)*

Definition: used for asserting the same change-of-state mentioned in connection with BREAK-1, but it is of its direct object. It assigns an agentive role to its subject. There may be an instrument specified (e.g., the rock)

- Break-3:

Example: *a rock breaks the stick*

Definition: it accepts inanimate subjects and does not occur with an instrument.

- Hit-1:

Example: *John hits the tree*

Definition: it is parallel to break-2. It assigns an agentive role to the subject

- Hit-2:

Example: *A rock hit the tree*

Definition: it is parallel to break- 3. It assigns an instrumental role to the inanimate subject

Hit does not show the type of break-1 meaning, which is the inchoative use. While this analysis takes syntactic facts that are valid and evident, Fillmore rejects this way of examining grammatical structures and meaning and divides verbs in classes according to a set of properties: *break* verbs undergo a change of state, *hit* verbs are involved with a surface contact. Crucially, with hit verbs, there is no change in the patients due to the physical contact. However, we know thanks to our experience in the world that a violent contact – e.g., between a bat and glass window – produces a change of state. This event can in fact be described as a hitting event or a breaking event with two different construals: thus, *hit* and *break* are conceptually linked as verbs via our world knowledge and our ability to predict the consequence of actions.

Continuing with Fillmore's observations, the derivation of stative adjectives is possible with *break* verbs but not with *hit* verbs.

(5) a. The window was broken/bent/shattered
b. *The window was hit/struck/slapped

Examples in (5a) are ambiguous in ways (5b) are not: they could either be passive sentences or description of states. Another difference concerns object vs place (6).

(6) a. I hit his leg vs I hit him on the leg
b. I broke his leg vs *I broke him on the leg

Hit can have a place in the object position (*on the leg*), but not *break* (see section 4.4).
Some idiosyncrasies emerge from break-type verbs: Fillmore notes that not every change-of-state verb can be used intransitively – *smash* and *cut* cannot. Verbs of the break class, which is the change-of-state class (e.g. *crack, smash, crumple*) occur in anti-causative sentences, while verbs of the hit class (e.g. *kick, batter, kiss*) do not (examples 7a-b). Hit but not break verbss are said to occur in the conative construction (examples 8a.b).

(7) a. The bicycle broke
b. *the car hit
(8) a. John hit at the bicycle
b.*John broke at the bicycle

Fillmore also mentions semantic properties: patients of the verb *break* must be rigid objects. Fillmore notes that the verb presupposes that world objects referred to by the nouns in a sentence have specific physical properties, and this aspect is also hard to generalize across verb classes. For example, bend requires an object to be rigid but allows a degree of flexibility. This aspect is what Comrie captures in discussing lexical aspects (section 2.3).

4.2 Levin (1993) and following work

Levin (1993) is a seminal work on English verb classes and many later works took it as a reference for research on verbs, alternations and argument structures (even in approaches theoretically distant e.g., Goldberg 2002, 2006). Verbs have been divided into classes according to their syntactic behaviour in a more detailed or coarse-grained manner in various influential works

(Dowty 1979, Van Valin and LaPolla 1997, Levin and Rappaport-Hovav 2005, Croft 2012) which had the common aim of finding grammatically relevant meaning components. Theoretically, Levin's work is nested into the idea that "the ideal lexical entry for a word should minimise the information provided for that word" (Levin, 1993: 11).

Levin (1993) states that "the nature of lexical knowledge confirms that various aspects of the syntactic behaviour of verbs are tied to their meaning" (Levin 1993: 5). This work identifies argument structure alternations (in diathesis, transitivity, and morphological features) which appear to be sensitive to verb meaning, proposing verb classes based on the verbs' shared syntactic alternations. Verb classes are thus based on testing the verbs in the alternations, using introspection to assess their acceptability in the alternations. The verbs *break, cut, hit* and *touch* are used by Levin in her introduction, building on Fillmore's research, to demonstrate the nature of lexical knowledge. These are all verbs used in transitive constructions but differ in the way they participate in other constructions: middle alternation, conative construction and body-part possessor ascension alternation. Starting from the latter, *break* does not participate in the body part possessor ascension alternation (9), as also noted by Fillmore (section 4.1). *Cut* and *break*, but not hit or *touch*, participate in the middle alternation (example 10) (Levin, 1993: 6). Finally, *cut* and *hit*, but not *break* or *touch*, participate in the conative alternation[6] (example 11)

(9) a. Jan {cut/touched/hit/broke} John's leg.
b. Jan {cut/touched/hit/*broke} John on the leg. (Levin 1993, p. 7)

(10) a. The bread {cuts/breaks/*touches/*hits} easily.
b. The door hits easily
c. Terry touched the cat → *Cats touch easily. (Levin 1993: 6)

(11) Jan {hit/cut/*broke/*touched} at the bread

The tests on each verb's alternating behaviour leads to discuss the aspects of the verbs' meaning that constrain them syntactically. The body part possessor alternation requires that the verb involve some contact between two entities – a component entailed by cut, hit, and touch, but supposedly not break. Considering the middle alternation, instead, it is characterised by the agent's omission. The construction requires a change of state, which is involved in *cut* and *break*. Levin (1993) observes that this is reflected in the zero-related nouns for these verbs: *a cut* and *a break* are visible results of the event, while *a touch* and *a hit* refer to the occurring of the events

[6] The conative construction is set apart by the transitive variant by the fact that there is no entailment that the action denoted by the verb was completed, only the attempt at some result. Thus (9) means "Margaret tried to cut/hit the bread."

but do not leave visible results. The conative alternation involves a motion and contact component, which Levin claims for *hit* and *cut*, but not for *break* or *touch*. Equally, verbs with only a motion component (e.g. *walk, run*) do not participate in it either. Table 4 summarises the distribution of constructions and verbs.

	Touch	**Hit**	**Cut**	**Break**
Conative Construction	(1)	+	+	-
Body-Part Possessor Ascension	+	+	+	-
Middle Construction	-	-	+	+

Table 4: Levin (1993) argument structure constructions for 4 verb classes: touch, hit, cut, break.

According to Levin (1993), the four patterns of syntactic behaviour in Table 4 correspond to four different verbs classes that share aspects of meaning. According to this claim, it is possible to predict syntactic behaviour based on semantics. Verb classes identified by Levin (1993 : 28) which follow this argument structure behaviour are the following:

> **Break Verbs**: break, chip, crack, crash, crush, fracture, rip, shatter, smash, snap, splinter, split, tear.
>
> **Cut Verbs**: chip, clip, cut, hack, hew, saw, scrape, scratch, slash, snip.

The semantic features they share are summed up in Table X. All verbs except *break* involve a certain degree of contact, with *hit* and *cut* involving both motion and contact. *Cut* involves a "separation in its material integrity" in the words of Hale and Keyser (1986). *Cut* inherently involves an instrument moving in contact with an object to bring about a change of state. *Break* is a pure verb of change of state and *Touch* is a pure verb of contact.

	Contact	Contact by motion	Change of state
Break	-	-	X
Cut	X	X	X
Hit	X	X	-
Touch	X	-	-

Table 5. Adapted from Levin (1993) – 4 verb classes and their semantic features.

The picture that emerges is that a verb's behaviour arises from the interaction of its meaning and general principles of grammar: "the lexical knowledge of a speaker of a language must include knowledge of the meaning of individual verbs, the meaning components that determine the syntactic behaviour of verbs, and the general principles that determine behaviour from verb meaning" (Levin, 1993: 11).

Another relevant alternation for breaking events is the Instrument Alternation. The interest for this alternation is to be found in the semantic implication of instruments in the events that verbs refer to: a breaking event can occur with an instrument and in our experience, it is not rare to have encountered a situation of a breaking event occurring with a variety of tools. This denotes agentivity but may or may not be intentional. In any case, the result indicated by the *break* verbs and the instrument should be coherent. The intransitive variant of this alternation is used to indicate an instrument's suitability for carrying out the action named by the verb. Examples with *break* in (12) are compared to the verb *cut* in (13a-b), which are both felicitous sentences.

(12) a. *This hammer won't break (on the interpretation where the hammer breaks something)

b. This hammer won't break the window.

(13)a. This scissor won't cut

b. This scissor won't cut the paper (Levin 1993: 38)

This alternation seems to shift the semantic focus: in *cut*, the fact the scissor cannot produce a change of state in the paper defines the scissors as unsuitable for the performance of the action denoted by the verb – irrespective of the theme. With break, this cannot be said: *break* cannot define the properties of its subject for semantic reasons, which are also link to the actual possibilities in the world. *Break* is not a verb that denote a physical action but just a result, which is not dependent on the grammatical subject alone. This alternation shows two things that will be

used to lead the empirical investigation and that had been captured by Comrie (1976): that semantic roles operate together with the verb in describing an event.

Levin comments about *break* verbs (1993: 241), saying that they refer to actions that bring about a change in the "material integrity" of an object and that, unlike the cut verbs, they are pure result verbs, and their meaning provides no information about how the result came about. The constructions and alternations they list (partially reported here), however, show that verbs in verb classes do not subscribe to the same alternations in all cases. Moreover, some verbs participate in more than one class: *break* is also part of the *hurt verbs* and *destroy verbs.* Destruction verbs are discussed by Levin (1993) as "they might seem similar to the *break* verbs […], which also appear to describe the destruction of entities; but the *break* verbs describe specifics of the resulting physical state of an entity (e.g., whether something is broken, splintered, cracked, and so on) rather than simply describing the fact that it is totally destroyed". Moreover, Levin adds that destroy verbs also differ from the *break* verbs in their syntactic behaviour – they are not found in the causative/inchoative alternation, nor in other causative alternations. This aspect is taken as indicative of the fuzziness of category boundaries and the flexibility of verb meanings.

Although definitions of the term *event* can be contextual and change in different theoretical frameworks, an event can be partitioned in at least an initial state and a final state. Simplifying, we can say that performing an action means intervening in the initial state. The final state may be different from the initial state or not. In breaking events, talking about the manner means focusing on the action performed to achieve the final state, while talking about results means focusing on the final state. The following works in the area of break and hit verb classes have focused on the two main aspects that these verbs contribute to event descriptions: the result and the manner. The difference between *break* and *hit* has been traditionally accounted for as a distinction between a result and a manner verb since Fillmore (1970) and in particular starting with Rappaport Hovav and Levin's *complementary hypothesis* (Rappaport-Hovav, 2010; Levin & Rappaport-Hovav, 2013). The complementary hypothesis states that a single verb root cannot encode both a manner and result: if a verb encodes the result, the manner is never specified, and vice versa. Both components can be encoded with complex verbs only, given that each root specifies a single trait (i.e., the English 'sweep clean'). Counterexamples to this hypothesis found by Beavers and Koontz-Garbooden (2012; 2020) include verbs of manner of killing (e.g. drown, hang, electrocute, crucify), manner of cooking (e.g. sauté, roast, fry, stew, braise), or ditransitive ballistic motion (e.g. throw, toss).

Whether a verb root contributes meaning about manner or result is generally tested with semantic and syntactic diagnostics. We summarised (Table 6) a selection of diagnostic tests used in manner-result complementarity studies with examples of the verbs of interest for this work.

	DIAGNOSTIC	EXAMPLE
TEST 1	Object deletion	*John breaks John hits
TEST 2	Denial of change ("but nothing is different about it")	*John broke the chair, but nothing is different about it John hit the chair, but nothing is different about it
TEST 3	resultative constructions (in pieces, open, in half)	John broke the chair in half.
TEST 4	Different agents (human, instrument, weather elements)	*The bat broke the vase. The bat hit the vase
TEST 5	"But I didn't move a muscle"	*I broke the vase, but I didn't move a muscle
TEST 6	take-time test. Punctual vs durative. It took 5 minutes vs for 5 minutes.	*I broke the vase for 5 minutes. I hit the vase for 5 minutes

Table 6. Summary of diagnostic tests.

We can consider diagnostic tests 2, 4 and 5. Diagnostic test 2 relies upon a crucial semantic characteristic of result verbs – that the result means there has been a change in the initial state. Thus, denying a change should diagnose a non-result verb. Examples in (14) show this test with result and manner verbs.

(14) a. *John broke the vase, but nothing is different about it

b. John hit the window, but nothing is different about it

Diagnostic test 4 (Beaver & Koontz-Garboden, 2017) is designed according to the argument that manner verbs restrict the types of subjects the sentence can have according to manner of action (examples in 29). Result verbs, instead, require no specific action of their subjects (indeed, only a result). This point is read here as a semantic argument: for a manner to be performed, a subject is required to be able to perform such action, which in turn requires agency. It seems redundant to operate with a concept of manner defined by the necessity of a human agent and devising a diagnostic for manner verbs based on subject restrictions. According to the view taken in this work, this point is hardly seen as a syntactic restriction: rather than the verb 'allowing' for a specific selection, it is the world categories (e.g., human agents) and what is possible in the world that is realised linguistically.

(15) a. The earthquake broke/destroyed the vase.

b. The hammer broke/destroyed the vase.

c. #The wind/mop wiped/swept the floor.

Diagnostic test 5 considers a fundamental component of actions – that is, movement. An action involves some performance, and Beavers & Koontz Garboden (2012, 2020) argue that in a causative, agentive sentence, movement cannot be denied if there is a result:

(16) a. *John broke a vase, but he didn't move a muscle.
b. John broke a vase, but he didn't hit it.

The argument is that if a verb only encodes the result, and nothing about the action performed, it should be possible to deny the action occurred (example 16a). Can someone cause something without performing any action? According to Beavers and Koontz-Garboden (2020) an example might be negligence (e.g., not doing maintenance of a car, leading to the car breaking – 2020, p. 175). Thus, result verbs should be compatible with this diagnostic. Irrespective of the felicitous use of 'but he didn't move a muscle' with a result verb in a negligence or non-negligence context, even if what is lexicalised is causation and not any sort of action, causation entails that an action has occurred. Sentence (3b) is, in fact, possible because it only denies the type of action.

These diagnostics help take different perspectives on the breaking category. From these tests, we understand that in *break* is used when (i) there is a change in state that cannot be denied, (ii) this change underlies an action being performed (which, in turn, cannot be denied) and (iii) the causer, agentive or non-agentive *affords* the action.

Moreover, manner and result remain useful terms to focus on different lexicalisations and components of the events. There are frequent correlations between the classification of verbs into manner or result and other verb classification criteria: manner verbs (e.g. hit) are usually activities in Vendler sense. In contrast, result verbs are always telic, and tendentially accomplishments (e.g. *clean*). Result verbs (e.g. *clean, cook, break*) tend to be much more general than manner verbs (e.g. *wipe, fry*). Thus, manner and result can be seen as strategies for categorisation and linked to the discussion on event categories (Chapter 1, section 4.1): it might be motivated to see manners and 'ways of doing something' (such as activities) as having a different hierarchical categorisation than results.

However, contextuality and terminology are two major issues. Manner and result are not described or defined as terms *per se* in this body of literature. Rather, they are used with their self-explanatory meaning, which can differ according to theoretical frameworks or types of analysis. *Hit* and *break* are discussed together because they occasionally describe the same event, in which one denotes a manner, and one denotes a result. Taken in other contexts, *hit* is as much general as *break*: hitting a vase can occur with different tools (a hand, a bat), in a repeated blow or a single

blow, from a distance or close. Besides, knowing that someone hit a vase does not give us a clear idea of what has happened – in terms of how it happened and what effects had on the patient. The only information is that there has been some contact. On the contrary, with *break* I know that there is a result and that a patient was *affected* (see section 4.4.) but I do not know the manner or the type of contact. The status of manner verb may seem contextual (in this case, when compared to *break*). Evidence for *hit*'s not full status of manner verb comes from Levin (to appear): this paper examines *hitting verbs* in directed motion event descriptions but states that the verb *hit* is not found in directed motion event descriptions. As directed motion event descriptions encode a hitting movement that occurs with a motion (e.g., the truck rumbled into the driveway), *hit* cannot contribute a specific manner. In this context, *hit* is neutral with respect to a manner.

4.3 Event structure

Stemming from the studies reviewed in sections 4.1 and 4.2, the *event structural* approach to verb meaning "takes verb meanings to be defined in terms of structured classifications of the events they describe" (Beaver and Koontz-Garboden, 2020: 6). It has been extensively used in lexical semantics (among others: Zacks & Tversky, 2001; Dowty, 1979; Jackendoff, 1990; Winskel, 1982, Marantz, 1997; Van Valin & LaPolla, 1997; Davis & Koenig, 2000). The event structure representation of the verb *break* involves subevents: an act, a cause, and a result (16), to be read as 'X acts in such a manner to cause Y to become broken'.

(17) [[X ACT] CAUSE [Become [Y <BROKEN>]]]
X BECOME BROKEN

About this approach to verb meaning, we propose two considerations. The first is cross-linguistic. This model is language-specific and cannot be generalised to other languages. For instance, the same model would work for Italian *rompere*, while in Swedish, which uses a predicate formed by a light verb and an adjectival particle (*ha sönder,* Eng. 'break') the two event components are split into two items that compose the predicate: CAUSE is realised by 'ha' and BECOME BROKEN is realised by sönder (Viberg 2020). The second concerns semantic theory and investigation: as Gärdenfors (2021) puts it, in this kind of analysis the even structure is represented using linguistic material, that is also reproposing the same verb under investigation. The event structural model thus cannot help investigate the meaning of the verb. However, it captures higher-level implications: that break describes a causal event performed agentively and that there is a change of state. The [ACT] could correspond to the force vectors that bring the contact, discussed in the

section above, necessary in causative breaking events. BECOME corresponds to the result and CAUSE is the mapping from force vector to result vector (Gärdenfors, 2021).

An interesting approach to the event structure of breaking events is taken by Mckoon and Love (2011), who set out to experimentally investigate event structures in terms of how speakers understand semantic boundaries. They start from the idea that the complexity of lexical decompositions (syntactic structures in which a verb occurs are controlled by a part of the verb's meaning) of verbal meaning has observable effects on online processing times. Verbs *hit* and *break* can both describe the same event as they both involve the fact that an agent can be responsible for applying force. Due to the syntactical differences (as described in Fillmore 1970) and their event structure, it is hypothesised that break is more complex than hit and requires longer processing times. The reasoning is that if event structures are encoded in the lexicon as parts of verb meanings, they should affect online operations – such as lexical decision response times. They perform three experiments. In experiment 1, they chose 16 pairs of verbs (one hit and one *break* verb) matched on seven factors known to affect lexical decision response times: frequency, subjective ratings of frequency, ratings of imageability, length in letters and syllables, and number of different lexicographic meanings. The stimuli were presented one at a time, and participants had to press a key if the item was a word and a different key if it was a nonword. Response times were 28 ms shorter for hit than *break* verbs, a difference significant by participants and items with analyses of variance. It supports the hypothesis that lexical decompositions, here implemented as event templates, are a part of verbs' lexical representations. In experiment 2, reading times for transitive *hit* and *break* sentences were measured to look for processing differences with a yes/no acceptability task. The hypothesis was that reading times for *break* sentences would be longer because they require processing of the CAUSE TO BREAK template (and the fact that a change of state occurs, contrary to *hit* verbs): *chipping* the tiles should take longer than b*anging* the tiles. Results showed that judgments were significantly faster for the *hit* transitive sentences. In the last experiment, they focused on the moment of processing difference (earlier or later the sentence) with a self-paced reading task where participants were presented with words once at a time and were asked to press a key to read the following word until the sentence stopped making sense. The experiment showed that response times for sentences with hit verbs were faster than those with *break* verbs ($F1(1,21) = 7.3$ and $F2(1,10) = 9.2$), with the difference occurring at the verb slot: response times were 7% shorter for the *hit* than the *break* verbs. A difference was also found at the final word of the sentence (a 10% difference). These differences are consistent with the hypotheses that the lexical representations of verbs reflect their event-template complexity and that processing times for sentences are, in part, determined by this complexity.

4.4 Constructional approaches

The construction grammar approach aims to overcome the verb-centred analysis typical of lexical-projectionist approaches, according to which argument structure constructions revolve around the verb. Constructionist approaches focus on the syntactic-semantic interface in argument structure constructions – which semantic values a specific syntactic arrangement has. We review three relevant studies for our purpose: Piunno and Mereu (2019), Romain (2021) and Kemmerer (2003).

Romain's (2021) approach focuses on the role of the theme in the constructional behaviour of the verb. The study is a corpus analysis on the causative alternation with a selection of verbs – among which verbs in the breaking domain: *break, crack, crush, snap, shatter* and *tear*. The analysis considered the degree of attraction of the verb to the theme and the differences in the pairing of verb-theme in relation to the construction (causative or inchoative). Results first showed that occurrences of *break* were higher in the causative construction than in the intransitive causative construction, meaning that *break* is most commonly denoting agentive events. *Break* was found with various themes (239 in 641 occurrences). Ultimately, this work proves that the causative and anti-causative alternation is only partially an alternation: not all verb-themes combinations alternate. Verb-argument configurations can appear in the causative and anti-causative alternation in a variety of forms, at least as far as physical objects are concerned. Some restrictions can be found in the abstract breaking that can occur, such as the example in (18):

(18) a. John broke the law
b. *The law broke
c. ? The law is broken

The reason for this syntactic behaviour is to be found in semantic implications concerning the relation between the verb and the argument. A physical object can be broken by an agent or be subjected to other forces or events that break it, or it may look like it broke spontaneously. When breaking the law, the law is not belonging to the participants of the communicative context. The law continues to exist for the others in a society. To draw a quick comparison, Italian does not use the general verb *rompere* for law, but the Italian anti-causative construction is still not accepted with the equivalent collocation for 'breaking the law':

(19) a. Luca ha infranto la legge
Luca broke the law
b. *La legge si è infranta

The law broke

A causing agent is needed depending on the semantic entailment of the theme – whether it is affected by the change in its entire existence or not (Spalek 2015). The notion of affectedness is further explored in section (7.1). Romain (2020) concludes that knowing a verb's meaning is insufficient to predict its occurrence in constructions and that constructional meaning should be described with a focus on the theme rather than the verb alone.

Piunno and Mereu (2019) conducted a construction-based study contrasting English and Italian, aiming to show that argument structure is not a fixed feature of verbs but is built based on linguistic and situational context. The focus is on semantically related classes of verbs which share argument structures, but also show structural differences in the domain breaking and hitting. They collect data from the Paisà Corpus (a maximum of 500 occurrences per verb) and select concrete uses of the verbs only. The verbs chosen share frame elements. Concrete occurrences are 49.6% of the total occurrences. The types of constructions found in corpora are reported below, taken from Piunno and Mereu (2019)[7]:

- **Bi-argumental transitive construction**. It is highly schematic and largely productive. Different degrees of agentive control over the action are allowed:
  - *L'acqua ha rotto la finestra* ('The water broke the glass', e.g., in the event of a flood)
  - *La mamma ha rotto il bicchiere* ('Mom broke the glass')
- **Oblique to subject alternation**. In this construction the subject is an instrument
  - *Una pietra ruppe le finestre* ('A rock broke the windows')
- **Passive constructions**. Here the authors find that the passive constructs differ in terms of aspectual properties: some of them denote a dynamic event and are expressed as verb, others refer to ongoing states, in the adjectival form:
  - *Tornare al momento in cui il vaso fu rotto* ('Going back to the moment the vase was broken')
  - *Mi ritrovo in strada e il mio motorino è rotto* ('I find myself in the streets and my motor-scooter is broken')
- **Tri-argumental and quadri-argumental transitive constructions**. Transitive constructions with arguments for semantic roles other than the patient:
  - *Un minatore ruppe il naso ad un ranger* ('A miner broke a ranger's nose')
  - *Il leone gli ruppe l'armatura con i fendenti degli artigli* ('The lion broke the armour with the cutting blow of this claws')

[7] Examples adapted.

This configuration, compared to the biargumental constructions are the constructional schematization of the dimension of construal of specificity described in Chapter 1-section 3.1.

- **Anticausative Si-construction**. In this construction the agent is missing, and subject/object is lacking animacy, intentionality and control over the event.
  - *Quando il motore si ruppe per la seconda volta* ('When the engine broke for the second time')

The different syntactic constructions are distributed differently among the verbs despite the fact the five verbs share semantic frames. Looking at the distributions, it emerges that the anti-causative construction is the most frequent for *rompere* (where the agent is not expressed nor understood). This is an interesting cross-linguistic finding, considering that Italian has the opposite trend of English, as found in Romain (2020) and the results from the annotation of the English *break*.

Kemmerer (2003) puts the focus on one argument structure alternation that involves *hit* and *break*. We discuss here the notion of affectedness in breaking events, as this relates to the semantic of breaking. The literature has shown that breaking events have been studied in contrast to hitting events. Beavers and Koontz-Garboden (2020) go as far as saying that *hit* and *break* have *similar meanings*: "physical impingement of some entity by some other entity" (Koontz-Garboden, 2020: 3). If here we can hardly agree on this statement, it is undeniable where it comes from: *break* is involved with contact, which produces a change in the state of the object that is encoded by the verb. This notion is captured syntactically by the studies on the body-parts ascension alternation:

(20) a. The cat bit John on the hand
b. *The cat bit John's hand
(21) a. Sam John broke Bill on the hand
b. *John broke Bill's hand

As we can see in examples (20-21) this alternation provides two grammatical constructions for the same event. In the first, a locative specifies a body part where the action has taken place, while in the second, a genitive expresses the possession of said body part. The reason why *break* verbs do not participate in this construction has been linked to the fact that *break* do not involve contact (see Levin, 1993 discussed in section 4.2). Kemmerer (2003) shows that *break* does not necessarily encode contact making two examples (22a-b). However, they are not concrete uses of

the verb *break*, rather metaphorical, and so do not directly relate to productive uses we have discussed so far.

(22) a. He swings at curve balls that break toward him rather than away from him
b. The fighter planes broke formation and returned to the base

Kemmerer adds to the notion of contact that of *affectedness.* In the body-part possessor ascension there are different construals of the body-part affectedness: according to the affectedness principle, entities expressed as the direct object NP are interpreted as being most affected by the action (Gropen et al., 1991). As a result, two alternating constructions changing which entity is in the direct object position change the affectedness of the entity: in 22a the focus is on the person and in 22b is on the body part. Kemmerer argues that the ascension construction conveys the impression that the person as a sentient being is being affected because a particular part of their body comes into contact with another entity. Dowty (2001) comments that in Levin (1993) there is no mention of a change in meaning involved in the body-part part ascension alternation: as far as Levin (1993) argues, the semantics of alternation is the same. Moreover, it is not explained why each alternation should be associated with specific semantic features. Dowty (2001) proposes to relate the body-part possessor ascension to other partitive constructions, calling PART-to-WHOLE SPREAD the phenomenon for which there is a relation between a patient and the part that is a physical sub-part of the patient. The part-to-whole spread entailments can have a long chain of extension (Dowty, 2001: 183), as in example 23.

(23) a. Mary touched the toenail on John's big toe. > Mary touched John's big toe. > Mary touched John's foot > Mary touched John's body > Mary touched John.

His hypothesis is explained with the following schema:

X a' Y on **z'**
Bill hit Tom on the leg

x a' z' is true and x is in physical contact with z	→	*Bill hits the leg* is true and *Bill* is in physical contact with *the leg*
z is a physical proper subpart of y	→	*The leg* is a physical subpart of *Tom*

x a y is true also (because of part-to-whole spread relation)	→	*Bill hits Tom* is also true
the meaning of a' involves motion from x to z.	→	The meaning of *hit* involves motion from *Bill* to *the leg*.

Table 7: schema of the part-whole spread entailments, adapted from Dowty 2001, pp 183-184

Turning to *break* verbs, 'Mary broke John's arm' does not entail 'Mary broke John'. Relating to the schema in Table 7, breakage of a part does not affect the whole patient—as Kemmerer suggests, the notion of *affectedness* is relevant here.

Within this schema, predictions are that verbs that do not allow part-to-whole relations do not participate in the construction. It can be added that *break* does not occur with animate patients in a causative/transitive construction (animate X verb animate Y): 'Mary broke Bill' is not conventional. It follows that an event that affects a part cannot entail the whole if the whole cannot be affected. These considerations can be related to what is discussed in Spalek (2015) regarding the affectedness of a whole for the metaphorical extension of the *break* verbs (see section 4.4). Following Dowty (2001), change of state denoted by the verb *break* also entails that there is some loss of functionality (e.g., a broken glass cannot be used to drink). When the function served by the part affected is different and separable from that served by the whole, there cannot be a transfer of entailment from part to whole: breaking a person's arm does not extend to 'breaking' the person entirely. Looking at other change-of-state verbs, like cutting verbs, part-whole relations do not entail that the new state invalidates a certain function. The events of c*utting an arm* or *a cake do not focus on the item's functionality*. However, in this case, the patient matters: for example, cutting a cake involves a redistribution purpose, while cutting an arm involves harming a person.

Thus, *break* seems to be ruled out from the body-part ascension construction not for reasons of contact, rather for part-whole relations and affectedness. This leads to the conclusion that contact does not seem relevant for *break,* which in turn leads to reevaluating its supposed similarity to hit. This argument is further examined by taking the perspective of the manner-result components of events and how it is linked to *hit* and *break*.

4.5 FrameNet

FrameNet (Baker, Fillmore & Lowe, 1998; Baker, Fillmore & Cronin, 2003), drawing from Frame Semantics, offers a classification of parts of speech (not only verbs) according to semantic similarity, rather than shared argument structure constructions.

In Frame Semantics (Fillmore, 1976; 1985) word meanings are understood in terms of complex structures of experiential knowledge. Verbs are grouped together when they evoke the same semantic frame, that is, when they describe the same types of situations involving the same number and types of participants (Fillmore 1976, 1985). These frame classes do not predict syntactic behaviour but represent semantic relations. The verb *break* in Framenet is linked to the Frames reported in Table 8. Searching individual frames, we find the verbs connected to it (Table 9).

Verb	**Frame**	**Lexical Entry**	**Frame description**
break apart	Cause_to_fragment	forcibly separate into pieces	An Agent suddenly and often violently separates the Whole_patient into two or more smaller Pieces, resulting in the Whole_patient no longer existing as such. […] the fragmentation is necessarily judged as injurious to the original Whole_patient.
Break apart	Breaking_apart	break into fragments	A Whole breaks apart into Pieces, resulting in the loss of the Whole (and in most cases, no piece that has a separate function).
Break down	Cause_to_fragment	Dismantle	An Agent suddenly and often violently separates the Whole_patient into two or more smaller Pieces, resulting in the Whole_patient no longer existing as such. […] the fragmentation is necessarily judged as injurious to the original Whole_patient.
Break down	Breaking_apart	come apart into small fragments	A Whole breaks apart into Pieces, resulting in the loss of the Whole (and in most cases, no piece that has a separate function).
Break down	Losing_it	fail to maintain emotional	A Protagonist goes from a default state of mental and emotional equilibrium into a temporary state of disequilibrium, often a state

		control and become upset.	of anger. The Protagonist's behavior is portrayed as deeply affected. An Explanation for the change of state may be given. Likewise, Actions that the Protagonist takes as a result of their new emotional state may be expressed.
Break out	Catching_fire	A fire suddenly catches	-
Break up	Cause_to_fragment	cause to separate into pieces	(see above for description)
Break	Cause_harm	to cause something to fracture	An Agent or a Cause injures a Victim. The Body_part of the Victim which is most directly affected may also be mentioned in the place of the Victim. In such cases, the Victim is often indicated as a genitive modifier of the Body_part.
Break	Compliance	fail to observe (a law, regulation, or agreement)	This frame concerns Acts and State_of_affairs for which Protagonists are responsible and which either follow or violate some set of rules or Norms.
Break	Experience_ bodily_harm	to fracture something (usually a bone), resulting in physical pain	An Experiencer is involved in a bodily injury to a Body_part (In some cases, no Body_part need be indicated). Often an Injuring_entity on which the Experiencer injures themselves is mentioned.
Break	Cause_to_fragment	cause to suddenly separate into pieces in a destructive manner.	(see above for description)

Break	Render_non_functional	make non-functional	An Agent affects an Artefact so that it is no longer capable of performing its inherent function.
Break.v	Breaking_off	to fall off of or separate from	A single Subpart breaks off from a Whole.
Break.v	Breaking_apart	separate into pieces as a result of a blow, shock, or strain.	(see above for description)

Table 8: Frames and descriptions from FrameNet for the verb break

Frame classes are general (e.g., cause harm) or more specific (e.g., break apart), resulting in differences in the verbs that share the frame. The Cause_harm frame features 47 verbs, while the Breaking_apart frame is linked to 8 verbs. The frames retrieved in Framenet capture the aspects of the verb *break* both in the concrete and metaphorical domains (e.g., Compliance; Losing it). Crucially, it takes constructions of verbs with prepositions into consideration (e.g., break apart, break down). The Frame approach leads to a different categorisation of *break* if compared to Levin (1993). It discusses the meaning contributed by the verb to an event that is made of participants identified by active and passive roles (actor, agent, experiencer, victim) and perspectives (e.g., causative or anti-causative).

Frame	Verbs in Frame Class
Cause_harm	Bash, batter, bayonette, beat up, beat, belt, biff, bludgeon, boil, break, bruise, buffet, burn, butt, cane, chop, claw, clout, club, crack, crush, cudgel, cuff, cut, drug, elbow, electrocute, flagellate, flog, fracture, gash, hammer, hit, horsewhip, hurt, impale, injure, jab, kick, knee, knife, knock, lash, main, maul, mutilate, pelt, poison, pummel, punch, run through, slap, slice, smack, smash, spear, squash, stag, sting, stone, strike, swipe, thwack, torture, transfix, twist, welt, whip, wound

Compliance	Abide, Adhere, breach, break, by-pass, circumvent, comply, conform, contravene, disobey, flout, follow, honour, keep, obey, observe, transgress, violate
Experience_bodily_harm	Abrade, break, bruise, burn, cut, graze, hit, hurt, injure, jam, pull, scrape, smack, sprain, strain, stub, sunburn, tear, twist
Cause_to_fragment	Break Apart, break down, break up, break, chip, cleave, dissect, dissolve, fragment, fracture, rend, rip up, rip, rive, shutter, shiver, shred, sliver, smash, snap, splinter, split, take apart, tear up, tear
Render_nonfunctional	Break, Disable, Hose, Incapacitate, Knock out
Breaking_off	Break, Chip, Snap
Breaking_apart	Break apart, Break down, break, crumble, fragment, shatter, snap, splinter

Table 9. FrameNet's Frame classes

The classes that emerge are, therefore, different compared to Levin's verbs classes. It enhances the dynamic nature of verbs, which appear in different classes, as a single verb can evoke different frames. From FrameNet, we take the salient frames evoked by the English verb break:

- a separation in pieces at more general or more specific levels
- non-functionality, with respect to an object's inherent function
- causing of bodily harm: it mentioned that *break* refers to breaking of bones in the body.

These frames will guide the semantic investigation of categorisation in Study 1 and 2.

5. Semantic categorisation of cutting and breaking

One of the most thorough studies in breaking events was conducted at the Max Planck Institute for Psycholinguistics in Nijmegen: the researchers' effort was aimed at understanding whether the domain of cutting and breaking events was conceptualized as coherent by speakers of different languages. This study had a significant impact on cognitive linguistics and in the field of research of categorisation and semantics. Follow-up studies contributed to the growth in the understanding of the cutting and breaking domain, by providing new insights and testing again

the original stimuli. Experiments have also been carried out in other theoretical frameworks. This body of literature is reviewed in the following sections.

5.2 Majid et al. (2007)

A large-scale cross-linguistic study of the *cutting and breaking* domain was published in a special issue of Cognitive Linguistics in 2007. It included 28 languages from 13 language families:

- one Altaic language (Turkish)
- two austronesian (Biak, Kilivila)
- one Caribbean (Tiriyo´)
- one Dravidian (Tamil)
- seven indoeuropean (Dutch, English, German, Hindi, Punjabi, Spanish, Swedish)
- two mayan (Tzeltal, Yukatek)
- three niger-congo (Ewe, Jalonke, Likpe)
- one Otomanguean (Otomi) one Pama-Nyungan (Kuuk Thaayorre)
- one Sino-Tibetan (Mandarin)
- one Tai (Lao)
- one West Papuan Phylum (Tidore)
- one Witotoan (Miraña)
- one Creole (Sranan)
- two isolate (Chontal, Japanese)
- two isolate Papuan (Touo, Yélî Dyne)

The data collected in this study revealed that there was consensus on how languages distinguished, using different verbs, *cutting and breaking* events from other events that shared visual properties with them. It emerged that distinguishing between reversible and irreversible separations in the object's material integrity is an aspect generally shared by different languages' verbs. Languages tend to refer with different verbs to irreversible events, such as 'tearing a piece of paper' or 'cutting a cake', from reversible events, such as 'opening a box' or 'removing the cap from a pen'. Once this was established, the set of irreversible separation events was put under focus in order to find out which were the most frequently shared characteristics of an action that was associated with a verb. It also emerged that the next shared dimension across languages is what distinguishes exactly cutting from breaking events – that is, how predictable the separation is going to be (I am more in control of the separation when I *cut*, rather than when I *break*). It was

shown how languages rely on different clusters of action-verb associations in their internal categorisation. For example, English distinguishes a *cut* cluster and a *break* cluster, while Swedish distinguishes a *break* cluster from *tear*, cut with a single blade, cut with a two-bladed instrument and a *snap* cluster.

The aim of this work was to define and delimit the categories of cutting and breaking, focusing on the general concept of events of separation with a description task comprising 61 video clips[8] showing physical events of separation. To select a set of events representing separation that could least be biased by English and thus return a valid, comprehensive picture of the linguistic variety, the stimuli were designed according to language acquisition studies, with reference to errors and overextensions (reviewed in Section 3.1 of this chapter): it was taken that children learning a language overextend on conceptual grounds, rather than linguistic. The actions were also non-conventional ways of cutting or breaking. Participants described the videos with a sentence, such as 'X cut the carrot'. The verb was the focus of analysis and particles (prepositions, adjectives) were not mentioned or coded in the analysis. Agents, instruments, and type of performance of the action changed. The annotation procedure per language was carried out by native-speaker linguists of the language being annotated or linguist experts in the target language. In total, 24 researchers were involved. Participants per language varied in number, from 1 to 7. For coding and analysis reasons, descriptions were taken as if the participants had given similarity judgments, in a binary mode: if two scenes were described at least by one participant with the same verb, they were taken as similar and coded with a 1 on a similarity matrix (whereas 0 was coded for non-similar scenes). For each language, a matrix was compiled based on pairs of action scenes (61 video clips taken in pairs make a total of 1830 pairs in each matrix). The authors ran two correspondence analyses to visualize a semantic mapping. The first correspondence analysis was run on all videos and served to isolate cutting and breaking categories from other separation events, such as opening and detaching. The second correspondence analysis was run on cutting and breaking scenes. Three main dimensions were extracted from the analysis:

- Predictability of the locus of separation
- Tearing events
- Snapping vs Smashing.

Cluster analysis brings evidence that cutting and breaking events mainly extend based on high or low predictability of the locus of separation: slicing a carrot is highly predictable and karate-chop a carrot is less predictable. This distinction is also referred to in follow-up studies (Fuji, Radetzky

[8] The clips are available for download at the following address:

& Sweetser, 2013) as messy (*break*) versus clean (*cut*) separation. The object modified by *cut* undergoes a clean separation, made with a dedicated instrument (e.g., a knife, a fork). A breaking event results in the object being fragmented into an unpredictable number of irregular pieces. Tearing events appear to cluster together separated from the other cutting and breaking events, applied to flexible objects like paper and fabric. Ten languages out of twenty-eight used only one verb for these events. Lastly, the third dimension distinguishes snapping events from smashing events. The snapping cluster comprises events in which a one-dimensional rigid object is separated into two pieces by the application of pressure to both ends (clips 25, 19, 57, 5), while the smashing cluster is made up of events in which a rigid object is fragmented into many pieces by a blow, e.g., with a hammer. Some languages observe this distinction and cannot overlook it, while in others, there is no such difference. English has this distinction, but smashing and snapping events can both be described with the general verb *break* – this would appear to be a basic-superordinate levels distinction (as discussed in Chapter 1, Section 4.1).

The languages in this study highlighted that some distinctions are possible but not observed in every language, while there are generally macro-categories with a more universal-like behaviour, such as the difference between reversible and non-reversible separation events, the most typical one. This approach has brought useful insights which were fundamental to the understanding of breaking events in particular, and event categorisation in general. As it is expected, the results are tightly connected to the specific scenes used to gather descriptions. The scenes are necessarily limited in the objects used, the setting, the instruments. The interpretation of the correspondence analysis is to be related to the stimuli. The dimensions that emerge are readings of the results based on the universe of actions that are shown to participants. The first dimension individuated in Majid et al. 2007 may also be influenced by this aspect: the predictability of the locus of separation is a dimension extracted from the scenes, which features instruments of varying sizes used in conventional or non-conventional ways. In these scenes, instruments play an important role in showing how, e.g., a carrot or a piece of fabric are separated in pieces, which may create more or less messy separations. Looking at the scenes, we can see that they are not linked to all frames that emerge in FrameNet for English: there are not cause_harm events or non-functionality events. For example, the contexts did not provide situations in which something was rendered non-functional and the actor in the video showed it could not use that object in its intended way anymore. Compromission of functionality is a central aspect of breaking events, that may or may not be linked to the physical, visible disintegration of an object. It is the case of technological devices, as in example (24):

(24) a. Julia broke my phone (the phone was dropped but the exterior is intact)
b. Julia broke her computer

As suggested by Sweetser and Bouveret (2009), the extension of breaking to electronic devices when they are rendered non-functional without an external physical deterioration is not accepted by all speakers. They report (although without a support or corpus data or experimental data) that English speakers might not accept 'I broke the phone' but they would accept 'the phone broke.

The set of video scenes representing separation in the material integrity of objects are all agentive and characterized by a selection of instruments and themes used in different combinations in the actions. Two aspects that are worth further discussion are the following:

1. The scenes display unusual events
2. There are no links to the functionality of the object: an object is visibly manipulated to separate parts and the objects chosen are themselves generally meant to be acted upon to change its state (e.g., carrots) o in contexts where there is no hint of functionality (such as cutting a piece of fabric that is stretched between two tables).

Concerning unusual separations, Bohnemeyer (2007), describe what constitutes the unusual:

- a mismatch between the state change or the instrument used to perform that change
- the existence of more canonical alternatives for the event to take place

In Bohnemeyer (2007) and Majid et al. (2007), the underlying idea was that collecting video descriptions of unusual events provides a way to explore the meaning potentials of verbs. Meaning potentials refer to the way speakers exploit specific verbs or constructions to describe events they have never described. In this way, one can tap into individual differences among the participants in the choice of construal. In addition, unusual event elicits more inter-speaker variation because the event seems new to the speaker, who has to categorise it on the spot the event and find the most appropriate construal to describe it. The following examples (25a-d) from Bohnemeyer (2007) show how an unusual event can be described[9].

[9] The clip is part of the "cutting and break video-clips" material - Bohnemeyer (2001)

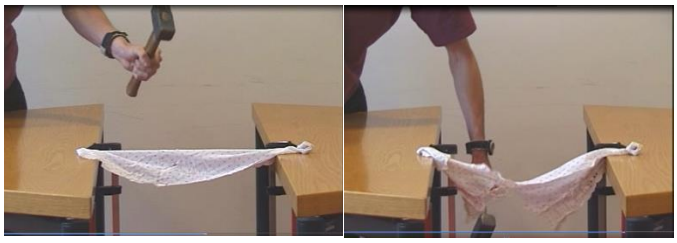


Figure 1. Clip showing an unusual event, part of the "cutting and break video-clips" material - Bohnemeyer (2001), used as stimuli in Majid et al. 2007.

(25) a. He hit the shirt with a mallet → fails to encode state change

b. ?He slashed the shirt with a mallet → the verb strongly implicates a bladed instrument rather than a mallet

c. ?He cut the shirt (in half/two parts) with a mallet → the verbs entails a bladed instrument

d. ?He broke the shirt w/ a mallet → *break* implicates a (semi-)rigid theme (Bohnemeyer 2007)

Speakers are used to describe events that recur typically in their lives, and there is a larger consensus on which verb to use to describe this sort of event. The verb comes to be associated with certain aspects of an event: its theme and specific properties of the theme, the instrument used, and the force used. When the event is unusual, the speaker must adapt their knowledge of a verb to a different circumstance. As an effect, agreement between speakers will be lower: different speakers might find certain aspects of the unusual event better represented by different verbs. Language-specific verb lexicalisations influence the results of such tasks, thus having different degrees of agreement on how to describe an event depending on the language. Cross-linguistic investigation helps look at these differences and tap into the verb's semantics. Is there a reason why we do not use the verb *break* with fabric material? The answers given by Bohnemeyer (2007), in a broader picture (not limited to *break* and fabric), are that a certain action is no part of the culture or that a certain action is not an efficient way to achieve a specific result state. Finally, he simply suggests lexicalisations do not have language-internal motivations (against structuralists' claims).

Taking together the discussion in this section, we bring to attention the methodological strength of the study in using video stimuli to investigate events, thus enabling the comparison of various languages through videos without language interference in cognitive activities. The scenes are carefully construed to investigate a specific domain but leave the potential for exploring the domain further with a different set of events. The two aspects that emerge from the study as involved in the conceptualisation of breaking events are:

1. It is a widely shared feature of languages to use different verbs for reversible separations (opening a container) and irreversible separations (disrupting the integrity of something so that it cannot be placed back in the same way you put a lid back on a container).
2. Since the consequence of a physical, irreversible modification of an object is its resulting non-functionality, these two frames are entangled in the semantics of the verb break.

5.3 Follow-up studies

Research following Majid et al. (2007) study tried to contribute by suggesting other aspects to be considered in the investigation of cutting and breaking events. Fuji, Radetzky and Sweetser (2013), with their Frame Semantic approach, mention how functionality is a frame that is part of breaking events and should then be taken into consideration. This aspect was probably beyond the original study, which focussed on the conceptual domain of physical separations, independent from their consequences on functionality. However, the aspect of functionality is central in the use of *break* verbs in context, which has not been investigated yet. It is evident how separations and functionality are linked: when I break a cup, it means the cup cannot be used for its original function anymore. However, the verb *break* (focussing on English for the moment) is used for events in which an object is rendered non-functional even if there is no (visible) separation. For example, when we say that our washing machine is broken but we see it in its integrity. In these cases, we are still dealing with a physical, concrete event – not a metaphorical one – but not necessarily a separation. These cases are referred to as non-actual separations in Devylder & Zlatev (2020): "while many would agree to treat [she broke my heart] as metaphorical, few would do so for [My computer broke], […]. A broken computer implies loss of functionality, which may often be the result of physical breaking, *but the result of an event and the event itself should not be conflated*" (my emphasis). In Fuji, Radetzky and Sweetser (2013), the non-functionality frame of the breaking category is only an extension for the English *break*, while the separation frame is the central one. This conclusion is reached by comparing English to other languages: Japanese *kowasu* (main translation: *break, damage*) was not used for separation events not linked to functionality. Dividing a cookie into two parts can be described in English as "breaking the cookie in two" but could not be described by kowasu, as in this scenario, the cookie still has its functionality – more than that, it serves the purpose of sharing. Kowasu could be used, instead, if cookies in a box were damaged and were not functional anymore as a present. English *break* conflates these two events.

Sweetser and Bouveret (2009) evidence how French has three verbs – rompre, casser, briser – that alternate in covering areas that in English are covered by *break* and how the

difference in the categorisation is captured using frames. The combination of different frames accounts for the crucial difference in the choice between words belonging to the same conceptual domain, as in the case of *break* verbs. In their approach they also recognize that the verb is not the only carrier of the meaning capable of evoking frames when an event is described. In addition, Fuji, Radetzky & Sweetser (2013) suggest further dimensions to consider in the semantic categorisation based on an English-Japanese comparison and multi-frame semantics. They find that the dimensions emerged in Majid et al. 2007 are aspects of varying frames of separation which involve different force-dynamics, different sounds connected to the action, different resulting visual patterns of separation. They suggest other parameters that should be taken into consideration as they seem to influence the use of *break* verbs in breaking events:

1. Loss of functionality. The stimuli used in Majid et al. 2007 do not involve functional objects, while Fuji, Radetzky and Sweetser (2013) highlight how relevant this aspect is in the concept of breaking.
2. Forceful bending and breakage. For example, o*ru* (most common English equivalent: *break*) involves forceful bending, but not necessarily a breaking result. This parameter is to be analysed in association with the type of patient entity: we know with experience that the bending of a flexible item would not result in breakage (we can think of a piece of fabric), unlike objects with a certain degree of flexibility (e.g., a pencil or a twig). English *break* is neutral in regard to impact and bending, while, for example, Japanese is ambiguous in describing the event *of breaking a finger and bending a finger,* which can be described with the verb 'yubi'.
3. Overall or partial loss of integrity with focus on detached pieces or on degree of affectedness of the whole. As examples, *smash* which features an overall loss of integrity, while *chip* determines a partial loss with focus on the detachment of a piece, or the Japanese tigiru, which encode a full detachment.
4. Functional divisions: some verbs denote a type of separation that carries a functional meaning: English slice. Moreover, the examples by Fuji et al. show that the semantic component captured by this parameter is better expressed by multiword expressions (going beyond the verb level): break a cookie *in two*, cut a cake *into pieces.* This is a function that, in English, group together cutting and breaking categories.
5. Intentionality and agentivity. In English this factor is relevant for *break* and *tear*, while cut verbs always needs an agent. In Japanese, unintentional action vs intentional action correlated with transitive/intransitive form used.

These considerations are partially captured by Frame classes as well (1-3) but also add aspects that have been discussed in syntactic approaches (5). We find number 4 an additional input in the investigation of breaking events that will be considered in Study 1 and 2.

5.4 Categorisation of closely related languages and codability

Relevant for the purposes of this work is also Majid et al. (2008) which furtherly investigated four genetically related languages (English, German, Swedish, and Dutch) with the same Majid et al. (2007) data. On assumptions that related languages must be more similar, it was instead revealed that they present interesting differences in the historical development of cognate words and their contemporary usage which resulted in a different semantic categorisation of this semantic domain. They considered cognate verbs in these languages: the historical perspective allowed them to notice and recall how verbs do not have an intrinsic meaning and how semantic categories change over time and space. The focus is again on verbs, regardless of the role prepositions have in defining an event. According to this research, these four Germanic languages cluster cutting and breaking events differently, despite being genetically related. The methodology was as follows: native speakers described some video clips (same clips as the study described above), but it is not reported how many participants per language. The consistency of responses was calculated using Simpson's diversity index, and it showed that Swedish speakers gave the most consistent descriptions, varying less in the verbs used in describing the videos. Results showed that the four Germanic languages cluster cutting and breaking events in four different ways, reported in the table below (Table 10). This historical perspective also allows us to notice and recall how verbs do not have an intrinsic meaning and how semantic categories change over time and space.

	Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**
English	Breaking (comprising smash, snap, tear)	Cutting (with knives or scissors, or in chopping events)			
German	Break (comprising chopping events)	Cut	Tear		

Dutch	Break (comprises chop, smash, snap)	Cut with a single blade instrument	Cut with a double-bladed instrument	tear	
Swedish	Break (Hugga)	Snap (Bryta)	Cut (skära)	Cut (Klippa)	Tear (Slita)

Table 10. Cutting and breaking clusters in Germanic languages (Majid et al. 2008)

This study used Simpson's diversity index to account for the internal coherence of participants' descriptions of the video stimuli. This index is a measure of how similar the responses to the description task as, which takes into consideration both the number of different verbs and their distribution (e.g., 10 responses of which 9 verb A and 1 verb B or 10 responses of which 5 verbs A and 5 verbs B will have different indices). The results are interesting in respect of the present work for two reasons: they show that closely related languages contribute to categorisation research with distinctions worth investigating and pave the way to include in a study the particles that constitute lexical units and that should be accounted for.

5.4 Stathi (2023)

Stathi (2023) investigates typological patterns in the verbalization of events (putting and taking; dressing and undressing; eating and drinking; cutting and breaking; hitting and kicking; opening and closing; washing; motion events) and objects (drinking vessels; sitting objects; clothing items). The languages investigated are English, German, Turkish, and Greek, on the grounds of their typological distinctions in satellite-framed (German and English) and verb-framed (Turkish and Greek). The focus is on the concept of *granularity*, the level of detail at which different languages verbalise objects and events. The objective is to find systematicity in this aspect of linguistic expression across different languages. In Stathi (2023), granularity, is characterized by two essential properties:

- the fact the degree of detail that one can express is a cline: one can be more or less granular in a description about the same entity
- the relativity of granularity: the level of details is not an objective fact, and whether a lexical item is general or specific depends on what it is contrasted to.

Granularity in language is investigated in lexical items – thus, according to the lexicalisations that different languages have. Generality and specificity in lexicalization are discussed in terms of *hierarchies* and hierarchical relations between lexical items, arguing that events are organised in hierarchies as well as objects. These aspects relate to what we have discussed in Chapter 1-section 4.1 and are relatable to what we discuss in terms of the dimension of construals (Chapter 1-section 3.1).

The categorisation of cutting and breaking resulting from this work is in line with Majid et al. (2008).

Stathi (2023) notes that occasionally the events are not construed with *break* verbs (rather, e.g., as a hitting or opening event), more typically so in the two Germanic languages than in Turkish and Greek. More lexical diversity is found in the Germanic languages than in Turkish and Greek, but in all languages, there is one dominant verb occurring in the description of the clips – cut in the case of English, followed by a *break*. Lexical diversity was analysed with Type-Token ratio (TTR) and the mean number of verbs per stimulus in each language, showing that Germanic languages have a higher TTR and mean. Statistical analysis with ANOVA confirms a main effect of Language, an effect for Language Type (thus, satellite-framed vs verb-framed) and follow-up t-tests show the significant difference between most language pairs, except for English and German.

6. Selecting the domain for investigation

Taking together the discussion on categorisation and the existing research on breaking events, our study aims to provide further data for event categorisation studies, proposing the use of a different set of (already existing) ecological scenes (to reduce the *native language bias* in stimuli creation) and taking into consideration Italian, whose linguistic categorisation of *break* verbs has not been discussed before. In what follows, we proceed in three steps: first selecting an investigation domain to contrast event categorisation in three languages, then looking at the linguistic referents and third analysing participant's construals.

6.1 Operational definition of breaking events

We circumscribe the events of analysis for practical purposes with an operational definition:

> An object's change of shape from a State A of physical integrity to a State B of physical separation
> in pieces operated with hands or instruments in a disruptive way, where the object loses its original functionality.

The *breaking* category should not be equated to events named by the English verb *break*. For example, according to the list above, *opening a can of tuna* would be a breaking event because it involves the change of state of an integral object which loses its functionality (=holding a content). However, in English this is generally referred to as *opening* rather than *breaking* – since the purpose of the action is to access a content and once the box is empty, its functionality (holding the content so that it does not leak) is not relevant anymore.

6.2 Languages under study

A sample of three languages suitable for a contrastive analysis is investigated in this work: English, Swedish and Italian. The languages are related, in being Indo-European languages and belonging to the Germanic and Romance families. English and Swedish, the two Germanic languages are thus genetically closely related. However, it is shown in Majid et al. (2008) that closely related languages show remarkable semantic differences, even in cognate words. Moreover, English has been in contact with a Romance language for long periods, which left traces, especially in vocabulary (Baugh & Cable, 1993; Van Gelderen, 2014). Typologically, Swedish and English are satellite-framed languages while Italian is a verb-framed language (Talmy, 2000). However, English is often taken as an example of a mixed-framed language for the verb-framed constructions with verbs of Latin origin such as *enter* and *exit*. Although the verb-framed and satellite-framed typological distinction is drawn upon motion events (describing how the path of motion and manner of motion are expressed), it has been used to discuss change of state events as well as other domains (as in Stathi 2023). Thus, the investigation of these languages Each language's verbal characteristics are introduced below, with respect to the domain of interest here.

6.2.1 English

English, which we have introduced throughout the literature review, is one of the most investigated languages with respect to the breaking category. Studies have focused on simple verb roots (*break, snap*), but in the English verbal system, phrasal verbs have a relevant role. Phrasal

verbs are composed of a root and a particle (typically a proposition or an adverb) and are conventionally recognized as lexical units (reported as dictionary entries).

6.2.2 Italian

Italian was introduced in section 4.4 regarding construction and the study conducted by Piunno and Mereu (2019). In Italian, a verb-framed language, verb roots mainly express breaking events. The most general verb for the breaking category is *rompere*. Intentionality or unintentionality of the breaking event is distinguished in Italian with the causative construction ('*Giulia ha rotto il piatto'* – Eng 'Giulia breaks the vase') and the anti-causative si-construction ('*il piatto si è rotto'* – Eng. 'the plate broke') where the *si* particle is reflexive. Piunno and Mereu (2019) highlighted how the anti-causative si-construction is the most used in Italian, in opposition to the trend of the English *break.*

6.2.3 Swedish

Swedish was introduced through the study on the categorisation of closely related languages (Majid et al. 2008). Cutting and breaking categories have been explored at length by Viberg (1985, 2007, 2020), other than Majid and colleagues (2007), throughout the years. Viberg (1985)[10] Swedish *ha sönder* have a prototypical meaning that combines the notions of separation (conceptualized as "disconnection") and non-functionality. In its purest form, this state of affairs is referred to as the second term in the antonymous pair of adjectives *hel* (Eng. 'whole') and *trasig* (Eng. 'broken'). Viberg (2007) used a FrameNet approach, and here, separation was linked to fragmentation with the frame 'cause to fragment'. Viberg (2020) suggested an open issue in the study of breaking events after decades of interest by different researchers: there is still a need to focus interest in a systematic way on the particles that complement the Swedish verbs. As discussed in section 5, Majid et al. (2007, 2008) studies in semantic categorisation showed that Swedish has more specialized clusters in the cutting and breaking domain, compared to other Germanic languages (Dutch, English, and German). Swedish does not have a single-root general verb for breaking events as English does but uses as a general verb a predicate composed of a light verb and an adjectival particle:

ha/ta/göra + sönder
have/take/make + in pieces

[10] Viberg has carried out extensive research in semantics on verbs across different *semantic fields* (e.g., perception, Viberg 1983 or posture verbs, Viberg 2013).

The predicate changes in constructions (Table 11): the causative construction uses *ha sönder*, the inchoative construction uses the verb *gå* (Eng. 'go'), and the stative construction uses the verb *vara* (Eng. 'to be').

Construction	Predicate	Example
Causative	Ha sönder	Johan har sönder burken *Johan has in pieces the vase* [Johan breaks the vase]
Inchoative	Gå sönder	Burken gick sönder *The vase went in pieces* [the vase broke]
Stative	Vara sönder	Burken är sönder *The vase is in pieces* [the vase in broken]

Table 11. Overview of the general break verb in Swedish

The particle sönder can be combined with other verbs than those included in the table. *Slå sönder* is a typical verb-particle combination that translates as 'smash'. Its root has the meaning of 'hit' and it thus combines the performance of the action (by hitting) with the result (sönder). An alternative to these predicates is the verb *bryta,* close to the English 'snap'. Viberg's studies focus on the themes used with each *break* verb/construction and the way they are found in translation in a Swedish-English parallel corpus and find that *bryta* is mainly used to refer to breaking of bones – a meaning that in English is comprised in *break.* It has little literal usage. Breaking bones refers to two types of Body-Harm depending on the agentivity: intentionally inflicting bodily injury (thus, inflicting body harm – Inflict_Body_Harm) and accidentally hurting oneself (Experience_Body_Harm). Examples (29-30) from Viberg (2020) refer to accidentally breaking a body part, and example refers to inflicting bodily harm.

(26) Jag bröt armen en gång som barn.
'I broke the arm once as child.'
(27) Om du bryter nyckelbenet på någon.
'If you break the collarbone on someone.'

Alternative verbs that may be used in breaking contexts are knäcka 'crack', krossa 'crush', spräcka 'crack' Coherently with bone breaking, *bryta* is also used for breaking of long thin objects by

bending, as in 'bryta av en pinne' (Eng. 'break (off) a stick'). Seen in comparison to English, *break* is more dominant as a correspondence of bryta than the opposite way around. About the degree of generality, the predicate *ha sönder* is more general than *bryta*. Lastly, it should be noticed that in the Swedish Academy Dictionaries (Svenska Akademiens Ordböcker) *ha/ta/göra sönder* is not listed as a lexical entry, but as a collocation.

6.3 Cross-linguistics considerations

Differences in the categorisation of events emerge based on the languages being compared. Comparison of languages is functional to typological research and to investigate the boundaries between universal-like shared features (shared conceptualisations among humans) and language-specific features, but it also contributes to the understanding of individual languages. By comparing linguistic features, there emerge insights that could not be detected with introspection, intuition, or data collection. The choice of languages to compare must be motivated according to one's research aims. As far as this research is concerned, the languages (English, Swedish and Italian) are chosen based on typological reasons, feasibility of the research and gaps in the literature. First, as far as Italian is concerned, it has not been thoroughly investigated yet in the domain of breaking events. Moreover, the result reported for Swedish seems partial: it has not thoroughly taken into consideration the role that the particle *sönder* plays in the breaking domain, which is a distinctive feature of Swedish. As a result, Majid et al. (2007) on Germanic languages, return a picture that could be complemented with further research which includes more explicitly how verbs are paired with particles. Lastly, English and Swedish have been found quite different with regards to breaking events, laying on the opposite sides of a cluster model (Majid et al., 2007): the former having the lowest number of clusters and the latter having the highest. As these results are based on a specific set of events, a different set of events could provide more data on these languages' clusters.

English, Swedish, and Italian belong to different language families (Germanic vs Romance) but the contact these two Germanic languages had with Romance languages has been extensive, especially in the case of English. In fact, concerning breaking events, it has been observed that in a comparison of Germanic language's cognates verbs there are major differences with English. Comparing the three languages in the way they refer to breaking events, it stands out that Italian and English have (general) verb roots, *rompere/break* which Swedish does not. Swedish uses general verbs paired with the particle sönder (lit. 'asunder', in pieces). The same particle is also used with more specific verbs of irreversible separation (e.g. *riva sönder*, 'tear in pieces') or other

verbs (*falla sönder*, lit. 'fall to pieces', 'collapse'). Besides the general verbs, the three languages have a range of specific verbs that are indicated by the main dictionaries[11] as synonyms to "break", reported in table 12:

	English	**Italian**	**Swedish**
General	Break	Rompere	Ha/Ta/Göra sönder
Other verbs	fracture	spezzare	förstöra
	bust	spaccare	slå sönder
	shatter	frantumare	bryta
	smash	guastare	knäcka
	snap	dividere	krossa
		interrompere	

Table 12. Examples of lexical variety in the break category in our sample languages

7. Conclusions and next steps

Events and their linguistic descriptions have undoubtedly attracted significant attention over the years and provided evidence of cross-linguistic shared structures, typological differences, and lexicalization patterns (among others: Stathi, 2023; Talmy, 2000; Malt et al., 2008; Slobin et al., 2014). This chapter set off from providing a characterization of events and action verbs in the literature according to the ways they have been classified in different works. This discussion highlighted some relevant characteristics of verbs and events:

(i) the relevance of the dimension of time in events
(ii) their links for force dynamics
(iii) the mismatch between lexical categories and events
(iv) the importance of considering semantic roles along with verbs.

We then presented the focus of investigation of this work, the breaking category, bringing to the attention of acquisition studies, which inspired a continuous interest in cross-linguistic

[11] English: Cambridge English Dictionary, https://dictionary.cambridge.org/
Italian: Dizionario Italiano Hoepli, https://www.grandidizionari.it/Dizionario_Italiano.aspx?idD=1
Swedish: Svenska Akademiens Ordböcker, https://svenska.se/

exploration. We reviewed the linguistics analysis carried out on verbs expressing the breaking category aimed at discussing the most characterizing aspects of its linguistic expressions. We provided an overview of semantic and syntactic aspects that characterize verbs expressing the breaking category: From Fillmore's *grammar of hitting and breaking* to Levin's verb classes, event structures, constructional approaches and the FrameNet perspective. This review highlighted that verbs expressing the breaking category encode the change of state of an action, described as a loss of material integrity and functionality of the object. The verb *break* is associated to the notions of affectedness of the object and result. Affectedness is distinguished from the notion of contact, which is instead relevant for the verb *hit. Hit* and *break* are also often compared as the encoding manner and results of breaking events. If we presented a critical point about this strict distinction, we also find manner and result two useful terms to characterize events: regardless of the linguistic expressions, these two components define the unfolding of events in the world. We then reviewed studies which explored semantic categorisation, and experimental approaches, to find that one parameter that involved in verbs that describe breaking events is that of the dimensionality of the object.

To carve an area of investigation and enhance the possibility of coherent contrastive analysis, we delimited the scope of our research to the breaking category according to a working definition (section 6.1). Additionally, we provided a description of our sample languages with an overview of the general verbs that describe the breaking category. This review sets the stage to address the gaps individuated in the literature and address them in the following chapters:

- Sample of languages: Italian was not among the languages included in Majid et al. (2007) or other cross-linguistic studies. Italian verbs of breaking have not thoroughly been addressed in a cognitive linguistics approach.
- Contrastive difference between English, Italian and Swedish: Swedish's predicates of the breaking category are typologically different from those of English and Italian. Swedish uses a predicate structure of a light verb or manner verbs followed by an adjectival particle, *sönder*, which encodes a result (ha sönder, hamra sönder). Thus, Swedish packs the manner and result components of a breaking event in a more efficient way compared to the English and Italian verb roots (*break, rompere*).
- Sample of events: research in categorisation has focused on a sample of events with specific characteristics (e.g., unusual events). Including everyday actions in the sample, and a different set of objects and tool, could provide further insights.
- A research agenda which combines methodologies: the research reviewed spaces from introspection, corpora and different types of elicitation (experimental or quasi-

experimental). A combination of methods of investigation and data collection could provide a more comprehensive view.

In the following chapters, we propose two empirical studies and one experimental study. We start with a corpus data investigation (Study 1) based on the parameters and characteristics of breaking events which are resulting from the literature review. Corpora contribute to this field of study by providing a wide overview of how action verbs that describe breaking events are used, with the context of occurrence and the themes they occur with. In this study, we explore typical configurations of verb-patients and the level of equivalence of a group of verbs in English and Italian and assess their generality-specificity cline. Corpus data is annotated manually with an innovative annotation schema.

The second study (Study 2) uses video stimuli to elicit written descriptions. Using videos is a proven efficient method to investigate events and actions (Majid et al., 2007; Stathi, 2023) since it captures the temporal unfolding of the events (compared, for example, to image stimuli), which, we discuss here, is a defining characteristic of events. Moreover, using video stimuli, it is possible to control the participants of the event (in this case, agents, objects, and tools). The aim of this study is to investigate the categorisation of breaking events, with a possible theoretical impact in categorisation theory, specifically on event categorisation, in English, Swedish, and Italian. This study will also provide the material to compare the resulting categorisation of the breaking domain with other previous studies (Majid et al. 2007 and Stathi 2023).

Lastly, in Study 3, we perform a video eye-tracking with spoken descriptions experiment to investigate the cognitive aspects of cross-linguistic differences in categorisation during sentence planning, with particular focus on manner videos and result videos.

Chapter 3. Cross-linguistic Equivalences and Semantic Boundaries[12]

1. Introduction

Verbs are the linguistic device most often involved in the description of events in the world. How the mapping between languages and events takes place has been discussed and debated for centuries, especially in view of linguistic diversity: different languages use different linguistic material to refer to events, have different lexicalisations, and event construals (Talmy, 2000). Linguistic typology and contrastive linguistics have shown this aspect of linguistic variation, highlighting different large-scale patterns that language families have in referring to events and finer-grained distinctions within closer languages.

In discussing categories throughout Chapter 1, we highlighted their inherent fuzziness: category membership is not sharp and is affected by function (Labov, 1973) or individual differences, as in the example of drinking vessels (Section 2.1). The tasks of naming physical objects that display some continuities, as well as naming activities (such as manners of human gaits: Slobin et al., 2014), have been shown to have grey areas, with a lack of coherence of participants' responses and varying degrees of variability. Assuming that lexical items evoke categories, it remains to be investigated what is the fuzziness in each linguistic item's meaning potential – thus, what is the fuzziness of linguistic categories. As well as fuzziness, Chapter 1 highlighted the role of typicality in the prototype theory of categorisation. We further ask what is the most typical semantic implication of a linguistic item. At first glance, it is very challenging to think intuitively of all uses of a verb to delineate a comprehensive definition of a verb: we highlighted in Chapter 1 that definitions with truth-conditions do not capture form-meaning relations. Even harder it may seem to discriminate one verb from another when they have a synonymy relation and are used to describe the same set of events. We may ask ourselves what might distinguish *smash* from *crush* and an immediate answer might be that we use *smash* to talk about a violent destruction into pieces and *crush* to denote a compression of some kind. This is itself informative, and a type of information given by dictionaries. However, both verbs are used

[12] This chapter reports and elaborate material published as Cacioli & Vernillo (2023)

for a variety of situations interchangeably: if I step on a flower, I may say that I *smashed* but also *crushed* the flowers.

This study sets out to investigate, from a usage-based perspective, a selection of breaking events and a selection of verbs in Italian and English when both can be included in the breaking category definition provided in Chapter 2, Section 6.1. The study brings together the world's perspective on breaking events (by using videos) and the linguistic perspective.

We start by assessing a set of breaking events in a grounded perspective, taking instantiations of breaking events from a multilanguage video repository by searching for the scenes described by the English *break* and the Italian *rompere* (the most general verbs to refer to breaking events in these two languages). This world-to-word approach is complemented by a corpus investigation. An extensive manual annotation investigates specific parameters that seem to characterize breaking events according to the literature. Through this data and annotation, we first look for the information about action schemas a verb is carrying and most often associated with (*typicality* of the category), and which are the verbs' semantic boundaries within a language (*fuzziness* of the category). This investigation in a contrastive view allows us to uncover the differences between a pair of non-distant languages in the verbs that refer to a shared set of scenes and which are generally regarded as translation equivalents. We then assess the extent of their overlapping and, consequently, their translation equivalence status. In the broader picture, this case study provides a contrastive view on a semantic domain with possible application in translation studies and L2 teaching.

2. Operationalize the study of semantic boundaries

It was stressed throughout Chapter 1 and Chapter 2 how events can be described with general verbs which are highly schematic, and how general verbs extend to events that are perceptually different from each other. To start investigating the relationship between events, categorisation, and linguistic items, we use a multilingual video repository, Imagact (www.imagact.it), which features videos depicting actions with a linguistic annotation in 16 languages. This resource is annotated with more than one verb per action, showing all the alternative ways one can describe an action. This is specifically important to investigate the relationship between a general verb (in our case, *rompere* and *break*) and other verbs belonging to the semantic domain of breaking events. This resource helps disambiguate events and provides a common ground for cross-linguistic research. We gather the events annotated by English *break* and Italian *rompere* and examine the features that bring them together under in the *break* or *rompere* categories. We check the other verbs that conform to our definition of the breaking category and that are used in the

Imagact repository to describe breaking events. This operations returns a list of verbs in the breaking category at different levels of specificity. To investigate fuzziness and category boundaries, we select a group of verbs per language that share semantic implications. Next, we use corpora to retrieve a larger number of contexts of naturally occurring usage. We combine the two resources in order to obtain pieces of information that complement each other: on one side we compare a selection of events in the world, on the other side we take a distance from referents and focus on language use. Taken separately, both resources have their limitations:

- Breaking events in Imagact are limited and only account for a small number of the possible events in the world.
- Linguistic data on corpora is generally not explicit enough to understand the action schemas underlying the events and other event components which have been shown to be relevant in verb categorisation.

As an example of this last point, Figure 1 shows screenshots of two events taken from Imagact. They show two events involving the same theme/patient (an egg) on which the agent performs an action. Both events can be described with the same construal (example 1), but only the event on the right can be described with the construal in example 2 (with English and Italian construals being the direct translations of each other). The descriptions in example 1 do not inform us of the type of action carried out in these events: we know that the egg has been acted upon so that it has now lost its integrity, but it could happen in either way of those depicted in Figure 1.

Figure 1: On the left, Imagact scene 'smash an egg with a hand'; on the right, depiction of 'break and egg'.

(1) a. Luca rompe l'uovo

b. John breaks the egg

(2) a. Luca schiaccia l'uovo

b. John crushes the egg

However, we may have an intuitive, inter-subjective knowledge about which one is the most typical, based on our experience in the world. And, in fact, when speaking or writing in context, we do not need to specify how breaking an egg takes place (e.g., in recipes it is not specified that you should not break the egg by hitting it with a palm of your hand!). The linguistic context in corpora provides clues to understand the event underlying the use of a verb. Additionally, it allows to consider the interaction of verbs and themes.

In this study, we operationalize the investigation of the verb's meaning by looking for the components of the events that are relevant for breaking events and are inferable in linguistic contexts. We first present the Imagact resource and the analysis of the events, and then describe the corpus investigation, annotation, and results.

3. Imagact

IMAGACT is a multimodal and multilingual ontology that uses short videos to visually depict action concepts with linguistic annotations (Moneglia et al., 2018; Panunzi et al., 2014). The resource shows 1010 videos annotated in 16 languages, with verbs and short descriptions in context (a minimal sentence, as shown in Figure 1). This resource serves different aims, among which are contrastive studies and L2 acquisition studies (among others: Gärtig-Bressan, 2021). It was conceived for the study of meaning in the expressions of actions: the visual means is a way to avoid the vagueness of semantic meta-definitions, which is especially an issue when discussing actions. The semantic information provided by the ontology allows to investigate the verbs variety within the scope of an action and the cross-linguistic differences. Each scene represents an action concept and is associated with one action verb, or more (Figure 2). In turn, each action verb is linked to more than one prototypical scene, which is retrievable by clicking on the verb (Figure 3). Action verbs sharing a common referent (or a subset of action scenes) are said to be locally equivalent verbs (Moneglia, 1998) (Figure 4). The resource can be explored in three different ways: a gallery of actions, a bilingual dictionary and the compare function (as shown in Figure 4). This resource's strengths lie in the pictorial means, the ease of cross-linguistic comparison and the different possibilities for research. In fact, the user (a researcher as well as a language learner) can access the resource by looking up one verb in a source language and checking the verbs it is annotated with in other target languages, extending then the research to other events described

by the verbs in the target language, and, in turn, to other verbs in the source language. This procedure results in an empirically based comparison devoid of language-driven definitions.

Figure 2. Imagact scene 'break the box', associated with verbs break, smash, destroy, and bust.

Figure 3. Imagact scenes for smash

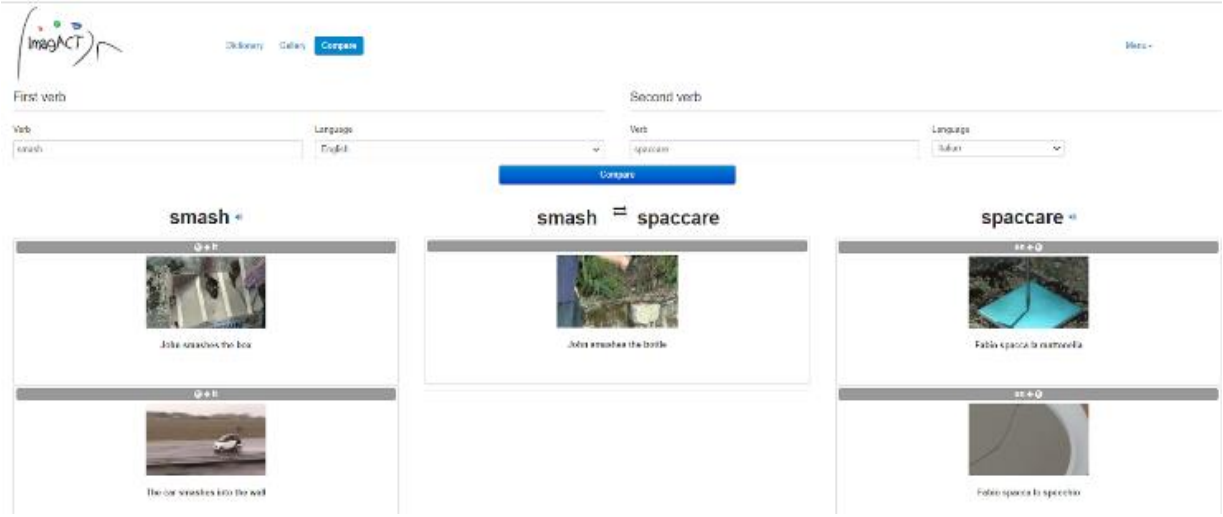


Figure 4. Local Equivalences for smash and spaccare in Imagact

3.1 Local equivalence

Imagact captures the different verbs that can be used to describe one single event. In the conception of the resource, the term *local equivalence* highlights the difference among the actions, making explicit the differential of each concept prototype with respect to the others (Panunzi et al., 2014). The concept of local equivalence does not rely on the concept of synonymy: verbs are *locally equivalent* when they can be used to refer to the same event (or classes of events) while having different meanings (Moneglia, 1997; 1998; 1999).

We can exemplify local equivalences with some practical examples in the breaking domain – and we can do so without resorting to translations. Figure 3 shows the local equivalences in the scene 'break the box'. The verbs *smash, break, destroy*, and *bust* can all be used to describe this event, but each of them contributes differently to the description. Local equivalences can also be investigated across languages. The scene in Figure 3 can be described by the Greek verbs διαλύω (dialyo) and σπάω (spao). Looking up these verbs we notice that they extend to different scenes. For example, we can compare the English *break* and σπάω: we see that they share most of the scenes (Figure 5). The centre column of Figure 5 shows the actions that can be referred to with both these verbs, while on the left and the right columns are the events described, respectively, with the Greek verb but not the English verb and with the English verb but not the Greek verb.

This example can be connected to the categorisation theory revised in Chapter 1: the Greek verb seems to extend to more events than the English verb – we can thus argue for it to be more *general* (as confirmed by Sathi 2023). This example also illustrates the motivation for using Imagact in this work: despite not being a resource designed under Cognitive Linguistics assumptions, it serves well for the purpose of this work. As we have illustrated, it provides a means of linguistic comparison in event descriptions as a first step into investigation. Moreover, what is here described as local equivalences, corresponds to the possible different *construals* of events.

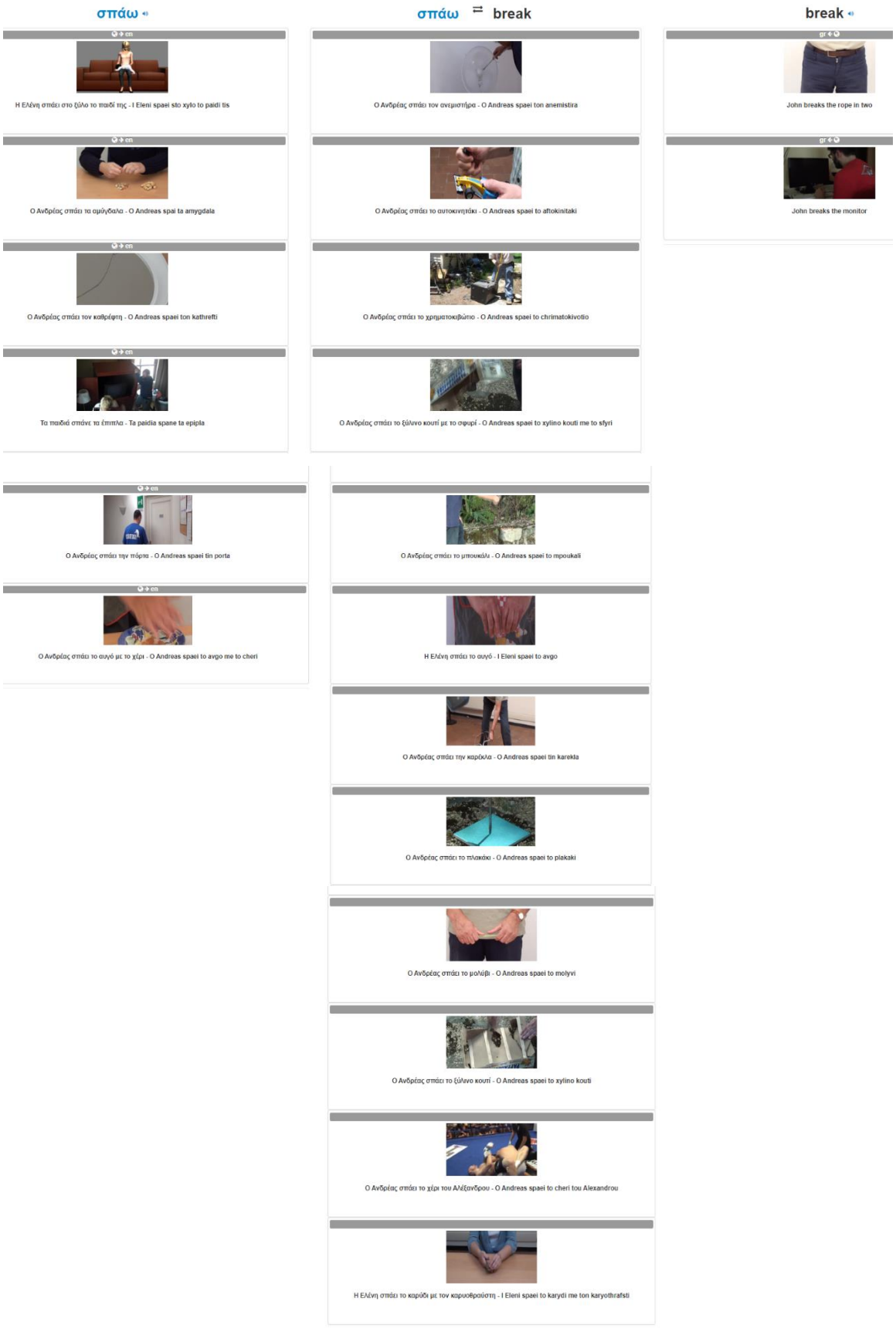


Figure 5. Compare function in Imagact. Verbs Break (English) and σπάω ([spao] Greek).

4 Breaking events in Imagact

Taking the main verbs that refer to the category BREAK in Italian and English we can look at their extension in the Imagact video scenes and compare them (Table 1). The videos tagged with *break* and *rompere* are varied with respect to action schemas and properties. All properties in the videos are described in Table 2.

Concerning Table 1, *rompere* and *break* extend to the same video scenes except for two scenes. The first is the monitor scene (John breaks the monitor), which is not tagged with *rompere* ('break') in Italian in Imagact. The second is the scene where a mirror is hit with a tool and it cracks, which is described with *rompere* but not with *break*. However, for some speakers, Italian *rompere* can be used in the monitor scene and English *break* can be used for the mirror.

We see how Italian and English have a general *break* verb, *rompere* and *break* respectively, covering a variety of breaking events. They only inform on the result – a fragmentation or a non-functionality (with or without exterior fragmentation).

4.1 Locally equivalent verbs

Looking at the alternative verbs in the scenes tagged with *break* extends the view on other ways of referring to the same actions. Verbs locally equivalent to *break* that emerge from Imagact are *bust, smash, snap crack, damage, destroy, open, open up*. The equivalent verbs may be used to describe the same scene but shift the focus on different aspects of the action. Local equivalent verbs of *rompere* are *staccare, sfasciare, spaccare, spezzare, schiacciare*, *danneggiare, forzare* and *aprire.* The equivalent verbs searched, in turn, in the repository reveal other scenes in which they are applicable but are not equivalent to *break*. In this section we sketch a profile of the meaning of each verb that is provided by Imagact according to the action it describes. When English and Italian verbs are cognates, we describe them together.

Video	Rompere	Break
	Fabio rompe l'automobilina • Rompere • Danneggiare • Staccare	Mary breaks the wheel off the toy • Break
	Fabio rompe il ventilatore • Rompere • Danneggiare	John breaks the fan • Break • Bust • Damage
	Fabio rompe la mattonella • Rompere • Spaccare • Spezzare	John breaks the tile • Break • Bust
	Fabio rompe la bottiglia • Rompere • Spaccare	John breaks the bottle • Break • Smash • Bust
	Fabio si rompe un braccio • Rompere	John breaks Simon's arm • Break • Bust
	Fabio rompe la cassaforte con il piede di porco • Aprire • Rompere • Forzare	John breaks the container open • Break • Open • Open up
	Fabio rompe la cassetta • Rompere • Distruggere • Sfasciare	John breaks the box • Break • Smash • Destroy • Bust
	Fabio rompe il fondo della cassetta • Rompere • Sfondare • Danneggiare	John breaks the bottom of the box • Break • Bust

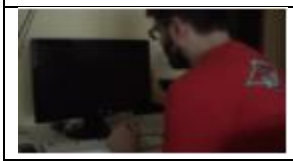
	Fabio rompe la sedia • Rompere • Danneggiare • Staccare	John breaks the chair • Break • Bust
	Fabio rompe l'uovo • aprire • rompere	Mary cracks the egg • Crack • Break • Open • Open up
	Maria rompere la noce • Aprire • Rompere • Schiacciare	Mary cracks the nut • Crack • Break • Open • Open up
	Fabio rompe la matita • Rompere • Spezzare	John breaks the pencil • Break • Snap • Bust
	Fabio rompe la corda • Rompere • Spezzare	John breaks the rope • Break • Bust
	Fabio apre la noce • Aprire • Rompere • Schiacciare	John cracks the walnut • Crack • Break • Open • Open up
	Fabio danneggia lo schermo • Danneggiare • Rovinare	John damages the monitor • Break • Damage
	Fabio rompe lo specchio • Rompere • Spaccare	The mirror cracks • Crack

Table 1. Screenshots of the Imagact events with sentence description and locally equivalent verbs in Italian and English

	SCENE	ACTION QUALITIES	OBJECT TYPE
1	smash bottle on a wall	OBJECT IN PIECES / CAN'T BE USED	CONTAINER
2	Smash wooden box with hammer	OBJECT IN PIECES / CAN'T BE USED	OPEN CONTAINER
3	divide rope in two by pulling the extremes	OBJECT IN TWO PIECES	FLEXIBLE ROPE
4	divide pencil in two	OBJECT IN TWO DISCERNABLE PARTS	STIFF 1D OBJECT
5	Break chair separating legs from the back	OBJECT IN TWO DISCERNABLE PARTS	PIECE OF FURNITURE
6	Breaking tile with tool	OBJECT IN PIECES, UNPREDICTABLE	SURFACE OBJECT
7	crack a mirror with tool	SURFACE OF OBJECT IN PIECES	SURFACE OBJECT
8	Wrestling-breaking arm	NON-VISIBLE SEPARATION IN BODY PART	BODY PART
9	instrument to make a fan non-functional	OBJECT NON-FUNCTIONAL, NON-VISIBLE SEPARATION	MECHANICAL DEVICE
10	Computer display not working	OBJECT NON-FUNCTIONAL, NON-VISIBLE SEPARATION	ELECTRONIC DEVICE
11	Create hole in a wooden box with a single hammer blow	OJBECT WITH ONE FRACTURE	OPEN CONTAINER
12	Breaking a wheel off a toy car	PART OF OBJECT DETACHED FROM SOURCE	MANIPULABLE TOY
13	branch from plant	PART OF OBJECT DETACHED FROM SOURCE	PLANT
14	Force with a tool to open a safe	OBJECT IN TWO PIECES TO ACCESS CONTENT	CLOSED CONTAINER
15	Peel peanuts	EXTERIOR OF OBJECT IN PIECES TO ACCESS A CONTENT	PRODUCT WITH CONTENT
16	Crack walnuts	EXTERIOR OF OBJECT IN PIECES TO ACCESS A CONTENT	PRODUCT WITH CONTENT
17	Crack an egg	EXTERIOR OF OBJECT IN PIECES TO ACCESS A CONTENT	PRODUCT WITH CONTENT

Table 2. Break and Rompere Imagact events described by two parameters: action quality and object type.

4.1.1 Aprire and open

Aprire and *open* share the semantic extension captured by the Imagact action scenes. In Imagact, there are 28 scenes annotated with *aprire*, of which 24 are annotated with *open* as well. *Aprire* and *open* are high-frequency general verbs and their extension is at a high level of abstraction. They refer to scenes denoting access to a container (e.g., a box), a content (e.g., nuts, book), or a

space (e.g., a room) and the expansion of a folded or crumpled material (e.g., *aprire l'ombrello / open an umbrella*). Many action schemas are involved: motions, forces, and contact:

- Pushing and pulling: *aprire la porta* / *open the door*, *aprire la finestra* / *open the window*
- Turning: *aprire il rubinetto* / *open the tap*
- Lifting: *aprire la scatola* / *open the box*, *aprire gli occhi* / *open one's eyes*
- Stretching: *aprire la mano* / *open a hand*

The actions can be performed with hands or tools, in repetition or in single strikes.

4.1.2 Forzare

Forzare is only used in one Imagact scene. The English cognate *force* is not annotated in Imagact. It is semantically characterized by the type of action – performed by exerting a force on an object – without lexicalizing the result.

4.1.3 Danneggiare and *damage*

Danneggiare is found in 6 scenes, 5 of which are shared with *rompere*. The sixth scene displays a man leaving a dent on the door of a car by pounding with a hammer ('Fabio danneggia la macchina' / 'John damages the car'). The extension shows that this verb is used when an object is altered in a negative way, so that it results either non-functional or affected in a non-reversible way (in the sense of Majid et al. 2007). The material loss of integrity is not a requisite of the event for it to be described with *danneggiare*.

4.1.4 Distruggere and *Destroy*

Distruggere and *destroy* extend to the same three scenes in Imagact:

- 'L'esplosione distrugge l'edificio' / 'The explosion destroys the building'
- 'i teppisti distruggono il mobilio' / 'People destroy the furniture'
- 'Fabio distrugge la cassetta di legno' / 'John destroys the box'.

This domain is certainly close to BREAK, but it appears in Imagact that their differential is in the size and extent of the object's change of state.

4.1.5 Spaccare

The Italian *spaccare* ('smash') extends to three scenes ('Fabio spacca la mattonella', 'Fabio spacca lo specchio', 'Fabio spacca la bottiglia'), which are all included in the extension of *break*. They depict separation in pieces obtained with an impact with force.

4.1.6 Spezzare

Spezzare ('snap') is annotated in three scenes ('Fabio spezza la mattonella', 'Fabio spezza la matita', 'Fabio spezza la corda'), all of which are *rompere* events. It occurs with one-dimensional and two-dimensional items. The actions are performed with hands in the case of one-dimensional

items, while they occur with a tool in the two-dimensional one. *Spezzare* and *spaccare* are equivalent in one scene ('Fabio *spezza/spacca* la mattonella').

4.1.7 Staccare

Staccare ('detach') is annotated in seven scenes (Figure 8), which refer to actions with a focus on surface contact, in particular, the removal of the contact between two objects. In case the removal occurs with the use of a certain degree of force, is non-reversible and compromises the functionality of the object, it is equivalent to *rompere*.

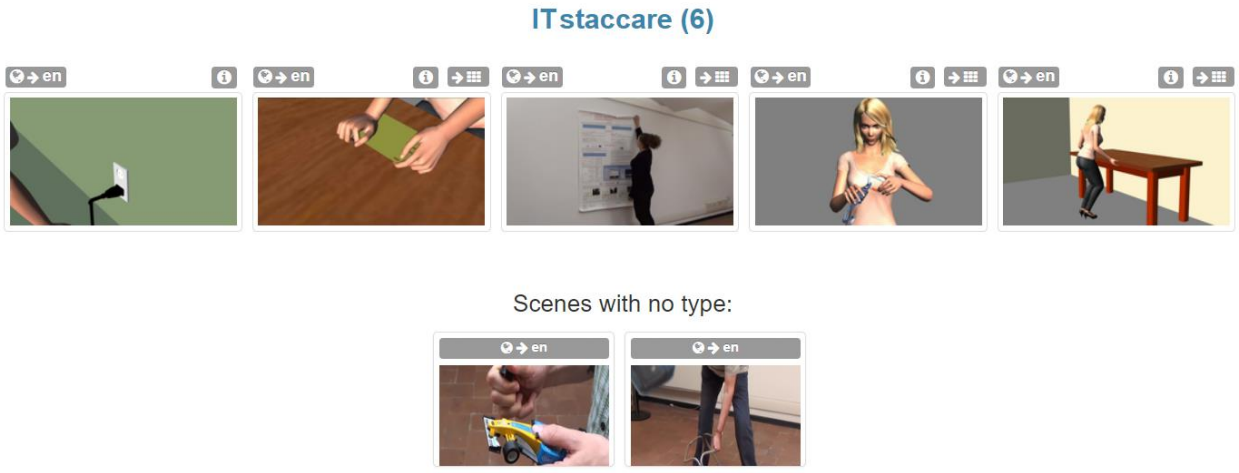


Figure 8. Imagact scenes for the verb staccare (Eng. 'detach')

3.1.8 Sfasciare

Sfasciare recurs in two scenes, one which is a *rompere* scene. The scenes depict actions of rough disruption of an object or multiple objects.

3.1.9 Schiacciare

Schiacciare is annotated in 12 scenes, which display various object modifications. Most of them do not involve material loss of integrity (Figure 9), rather different degrees of reshaping (e.g., squeezing a stuffed animal) and reversible modifications (e.g., pressing a button). They all involve contact and a pressing force, which may or may not induce different degrees of result – a reshaping, as in the case of the bottle scene (reducing the volume of a plastic bottle applying a force on both ends), reversible modifications (pressing down on a stuffed animal, which changes its shape temporarily) or the loss of integrity (pressing down with a fork on a boiled potato).

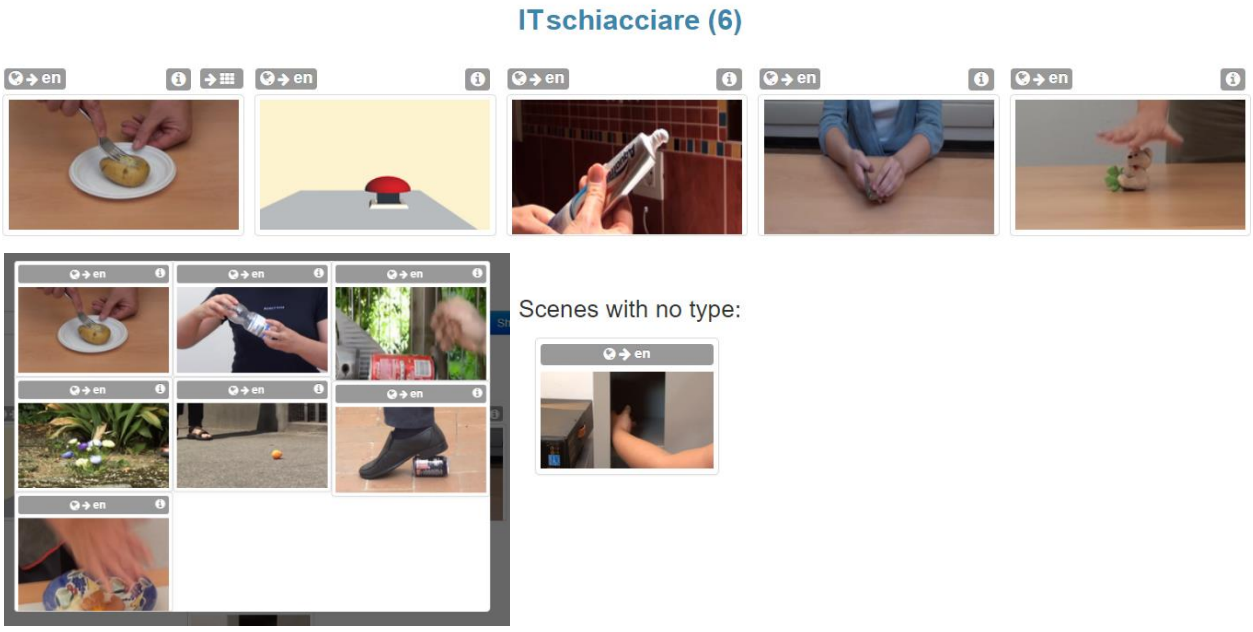


Figure 9. Screenshot of Imagact scenes for schiacciare (Eng. 'crush').

The focus of *schiacciare* seems to be on the action rather than the result: in fact, the result of this action is highly dependent on the object and its morphology, but in all cases the performance of the action includes exerting a pressing force. The objects are varied in shape and consistency and are acted upon with body parts (hands or foot) or tools (e.g., a fork on the smashed potato or a can in a pressing machine). As mentioned, there is not always an irreversible change of state: if the action performance is stopped (that is, the pressing is stopped), the object can go back to its initial state (when allowed by its physical morphology). Further scene comparison shows that *schiacciare* may also be involved with a pressing from above, applied vertically on the object. Figure 9 shows the differential: the scene on the right is annotated with *stringere* ('squeeze') and the one on the left with *schiacciare* ('crush').

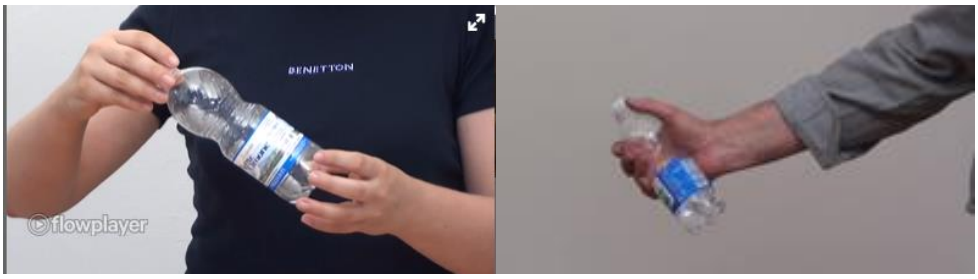


Figure 10: Two Imagact scenes showing different lexicalisations for the orientation of force application: crush (on the left) and squeeze (on the right).

3.1.10 Smash

In English, a greater extension to a variety of scenes is found with *smash* (N=10). The scenes are all linked to material destruction or disruptive object manipulation, even if it does not result in a material loss of integrity but in a disruptive reshaping (e.g., 'smash a can with a foot'). *Smash* is also locally equivalent with *hit*.

3.1.11 Snap

Other than with snapping a pencil, *snap* is used for the scene of snapping a branch, when a fracture with consequent separation occurs between the trunk of a tree/plant and a branch. *Snap* is applicable to the fracture of a branch itself (as with the pencil). The objects with which it is found in Imagact are all extended on one dimension (e.g., long, and thin objects).

3.1.12 Crack

Crack is used for 'crack and egg', 'crack a nut' and 'crack a mirror' scenes. In 'crack a mirror' the semantic component captured by the verb is that of the creation of separation lines in a two-dimensional material (e.g., a tile, a mirror) compromising the integrity of the object. This result state can be achieved with a body part or with an instrument. With eggs and nuts, the action occurs with three-dimensional objects and with the specific aim of accessing content – but there remains in the result of the action the component of fragmentation by the creation of separation lines. Given this characteristic, the verb is used with rigid materials, which can physically result in such a state.

3.1.13 Crush

Crush denotes a material modification of the object with or without a clear separation in parts. There is a reduction in volume and disruption of the original shape (depending on the modified object) due to the application of a force. The change is irreversible in all scenes (e.g., 'crush a potato') with a variety of objects (from soft to rigid). As mentioned above, *crush* can be compared to *squeeze*: the two verbs are distinguished by the action scheme performed on the object in different spatial configurations. *Squeeze* is generally reversible, and the pressure is applied on two sides at the same time. In *crush,* the pressure is applied on an object against a surface.

3.1.14 Contrastive considerations

After this cursory analysis of the verbs in the way their meaning emerges from Imagact, we can attempt drawing translation equivalent pairs between Italian and English (Table 3). Most lexical items are close to each other and refer to the same (or very close to each other) set of scenes. English *crack* does not have a direct translation that captures the semantics we described. The

same can be said for the Italian *sfasciare* which is often translated as 'smash'. As mentioned above, the Italian *forzare* has a cognate in English (the verb *force*) which is not annotated in Imagact. Italian *staccare* ('detach') does not describe a breaking event focused on non-functionality or result in pieces. It is a locally equivalent verb to a break scene but evoke a different domain. In English, the same scene is not construed as evoking this domain.

Italian	**English**
Rompere	Break
Spezzare	Snap
Spaccare	Smash
Aprire	Open
Destroy	Distruggere
Danneggiare	Damage
Schiacciare	Crush
Staccare	-
Forzare	-
Sfasciare	-
-	Crack

Table 3. English-Italian correspondence pairs

In this section, we used IMAGACT as a tool to compare *break* verbs on the action schemas and event components (effectors, objects, materials, scope) they refer to within each language and in contrastive view. Only concrete senses of the verbs have been taken into consideration. Imagact provides a way to select verbs used to encode a specific semantic domain of action concepts: using local equivalence it is possible to select other than the general *break* and *rompere* verbs that can be used to denote a breaking event. What emerges is the way two different languages segment the possible realm of real-world events and encode them in language with different verbs. *Rompere* and *break* are the general verbs that refer to these events and cover a vast range of visually different actions, exploiting different motor schemas, results, and different object modifications. In their concrete sense, the two verbs are equivalent in most of the action types and scenes. However, the other verbs that are locally equivalent with *break* and *rompere* are shown to contribute to an event description with different components. Moreover, they do not always have straightforward translations. The investigation carried out in this section was led by the question of how different verbs select different aspects of reality when we talk about events, which was operated by means of visual scene inspection. The contrastive analysis also provided pairs of English-Italian translation equivalents, based on the reference to specific actions. In what follows, the closest three of these pairs will be investigated in corpora, with a usage-based perspective: *break-rompere; smash-spaccare; snap-spezzare*. The aim is to investigate the semantic boundaries of *break* verbs within a language and their equivalents across languages.

5 Corpus study

We now analyse corpora occurrences of *break* verbs in English and Italian, aiming to uncover semantic boundaries between general and specific *break* verbs. We propose a methodology for a qualitative analysis of the corpus occurrences.

The first problem to address in a typological or contrastive study is to find equivalences in the elements to be compared. Corpus-based approaches allow a systematic comparison of the objects of interest, analysable in both qualitative and quantitative ways. In a corpus-based approach, one can use a parallel corpus or two comparable corpora. Using a parallel corpus allows to highlight the major translation strategies and the recurrent collocations. Attention should be paid to the general tendencies and not to the attestation of a hapax, as it may be misleading. In parallel corpora, unidirectional or bidirectional, the researcher can find translations of same-source texts (e.g., from Italian into English and from English into Italian): having the same texts available enhances comparability across languages and makes it possible to identify translation correspondences. This approach is used by Viberg (2021) for *cutting and breaking* verbs. Parallel corpora are interrogated to look for translation equivalences and, more importantly, examples of non-equivalence of two near-synonyms (examples reported in paragraph 5.1). As an example of recurrent translation, in the field of drugs and medicines it is frequent to find the following:

(1) Prima dell' impiego, togliere il rivestimento del blister e rimuovere la compressa orodispersibile senza **romperla** (Translation: *Before using, carefully peel open the blister and remove the dose of orodispersible tablet without* ***crushing*** *it*)

The verb *rompere* is frequently translated with *crush* in this context, evoking the frame CAUSE_TO_FRAGMENT. In English, *crush* is the preferred translation when referring to tablets – *crush* evokes the FrameNet frame GRINDING. However, there is not always loss of functionality (a tablet can be purposefully reduced in pieces for easier ingestion). *Break* is also used in the sense of division per distribution and purposeful breaking:

(2) we must **break the tendency** to import American incongruities into Europe in this area as in other areas → Dobbiamo inoltre **porre fine alla tendenza** di importare (Opus 2 – Europarl3)

The correspondences between languages are further confirmed with bilingual dictionaries and a world-to-word approach (Malt et al. 2010).

With comparable corpora the researcher can use original texts (selecting the same text types or genres) and compare the languages based on the context of occurrence, constructions, or other parameters relevant for research questions. In this study, we chose to use comparable corpora to investigate semantic boundaries. The first motivation is that we assessed translation equivalence using Imagact – in an attempt to reconnect the events to their linguistic description. The second motivation is that original texts from comparable corpora are the most appropriate here: they provide naturally occurring data or various genre, avoiding authenticity issues and translation effects. The occurrences cannot be compared on a one-to-one basis but have more ecological validity.

5.1 Verb selection

The verbs investigated here are, as discussed in section 4.2, the English *break, smash, snap* and the Italian *rompere, spaccare, spezzare,* going from general to specific (*break* and *rompere* as the more general; *spaccare, spezzare* and *snap, smash* as more specific). It should be noted that the Italian verbs select specifically breaking events, while English *snap* and *smash* are polysemous and are used in other events as well. In dictionaries, the material disintegration is their central meaning, but occurrences of non-breaking events were found in our corpus selection. These were excluded from the analyses and the sample compensated with the extraction of the missing occurrences to keep the samples balanced across languages.

From an intra-linguistic perspective, the selected verbs are cohesive sets which can be used to refer to the same events – in Moneglia's (1998) terminology they are locally equivalent in general-specific pairs: that is, *break-snap, break-smash, rompere-spezzare, rompere-spaccare*. In one video-scene, *rompere, spaccare* and *spezzare* are all locally equivalent. According to this premise, the semantic boundaries between the Italian set of verbs should be fuzzier, as there exists at least one event in which they can be used for description purposes. From a contrastive view and in terms of near-synonyms (Enghels, 2018), we chose to compare these verbs because of their pair-wise translation correspondence (*break-rompere*, *smash-spaccare*, *snap-spezzare*) in bilingual dictionaries.

5.2 Method and procedure

We extracted a random sample of occurrences from ItTenten2016 for Italian and EnTenten2015 for English. The Tentens are a family of text corpora built upon linguistically valuable content crawled from the web (using the Spiderling tool: Suchomel and Pomikálek, 2012). This is a large family of corpora containing more than 10 billion words per language and available in more than 40 languages. The texts are lemmatized, and part-of-speech tagged, and the same procedure of data crawling is followed for all languages, so that they can be perfectly compared with one another, regardless of the chosen languages (Jakubíček et al., 2013).

Of the sample, we kept the first 500 occurrences per verb (N verbs=6; N total occurrences: 3000), which referred to breaking events and where the target word was used as a predicate. Nominal (e.g., 'Displaying snaps of some of history's most elegant women') or adjectival uses (e.g., 'Steam spewed from the broken engine') were discarded. We rejected the occurrences which did not refer to breaking events (e.g., 'The dog was snapping at the air'), could not be interpreted in an actional context (e.g., 'It is faith, blood, mud, fire, water, speed, smash and build'). Metaphorical and idiomatic uses were kept in the sample and annotated separately. The total numbers of concrete occurrences are reported in Table 4.

Italian data were annotated separately by two linguist annotators (as reported in Cacioli & Vernillo 2023), who then cross-checked to resolve conflicts. English data were annotated by two native-speakers annotators for the parameters of Force, Degree of Modification and Result State on a random sample of 200 occurrences. The inter-annotator agreement (Armstein, 2007) was calculated on the coding of these parameters because it relies on a combination of inferring (in particular for Force), which is sensible to subjectivity, and close reading of what is explicitly mentioned or not mentioned in the occurrence's context. If the value for these parameters could not be inferred and was not mentioned, it was coded as unspecified, leaving room for different inferences and individual differences among annotators. Dimensionality and Flexibility were coded classifying the mentioned patients according to their physical characteristics, with lower risks of annotators biases. Cohen's Kappa was calculated between two native Italian annotators and two native speakers English annotators, resulting in good to strong agreement. Cohen's Kappa values for Italian are 0.81 for Force, 0.89 for Degree of Modification and 0.81 for Result State. For English, Force 0.65; Degree of Modification 0.78; Result State: 0.73.

Rompere	**Spezzare**	**Spaccare**	**Break**	**Snap**	**Smash**	**Total**
160	139	164	96	80	300	939

Table 4. Total number of occurrences describing concrete instances of breaking events.

5.3 Annotation schema and parameters

The annotation procedure was based on the visual inspection of 25 prototypical action scenes from Imagact, conducted according to the identification of salient action properties, and aspects of breaking actions discussed in Chapter 2. The visual inspection resulted in a matrix of action concepts and their visual and actional characteristics. Then, we moved to the manual annotation of corpora occurrences. The first step of the manual annotation is the individuation of the concrete occurrences. The occurrences were tagged as concrete if and only if the verb referred to a physical action (e.g., 'As we piled into [the metro], smashed like sardines' vs 'The bowl slips from her hands and smashes on the floor'). We selected 5 parameters: Force, Degree of Modification of the object, Result State of the object, Dimensionality and Rigidity of the object.

The first parameter, **Force**, refers to the type of force schema employed to affect the physical integrity of a patient entity. We only considered the force dynamics that seem to operate within the breaking domain according to the visual inspection of scenes and the literature review and we thus focused on contact forces, having considered the role of contact and affectedness in breaking events (Chapter 2 – section 4.4). We considered four types of contact forces. The second parameter, **Degree of Modification**, refers to the degree of change occurring in the object acted upon. The third parameter, **Result State** refers to the partial or total disintegration in the final state of the affected object. The fourth parameter, **Dimensionality**, refers to the object's shape as a relevant component in human cognition (specifically, with reference to actual or fictive object manipulation; see Matlock, 2004; Matsumoto, 1996; Langacker, 1986) and for breaking events (Majid et al., 2007; Bouveret & Sweetser, 2009). The fifth parameter, **Rigidity,** refers to the object's physical properties in terms of softness, brittleness and flexibility as described in the literature (Bouveret & Sweetser, 2009; Fuji, Radetsky & Sweetser, 2013). However, the full array of materials was too heterogeneous for the purpose of our annotation and only two values are coded here: *flexible* vs *non-flexible*. Table 5 reports the annotation schema as reported in Cacioli and Vernillo (2023).

We annotated each corpus occurrence with the five parameters described above. Being naturally occurring sentences gathered from a corpus, information on the action and the visual state of the result is not directly accessible in the verb or in the close vicinity. Thus, we investigated the broader context as provided in Sketchengine (up to entire paragraphs). In the event of underspecified parameters in the broader context were annotated as *unspecified* (as in *si può essere rotto questo e quest'altro* (itTenTen, "It is possible that this and that broke"). This tag is also

used when the context was vague enough so that more than one value of a parameter could be plausible or could be inferred (in *una volta che avremo rotto l'occhiello avremo in mano un'esca veramente inutilizzabile* (itTenTen, "once we break the eyelet we will have a completely unusable lure" = more than one of the values for *force* could produce this result) or when the modified theme is not mentioned (e.g., *un pazzo stava rompendo tutto*, itTenTen, "A crazy man was breaking everything"). This vagueness is not surprising, but inherent to languages: the verbs we investigate are general verbs (Moneglia, 1998) and leave linguistically unspecified much of the actional or perceptual information.

The distribution of the verbs with the possible values they can take (according to Imagact) is shown in Table 6. We notice that r*ompere* and *break* are tagged in videos that show breaking events occurring with the greater variety of the values we identified for our parameters, and with the possibility of no modification/no change (e.g., with electronical devices). The more specific Italian verbs are associated to fragmentation events, while *smash* is also used for change of state by reshaping that do not result in material disintegration (e.g., flatten a can with a foot) and *snap* also refers to an event resulting in a change of state by separation (snapping a twig from a plant stem, with both parts retaining their identity). S*pezzare* and *snap* are similar in the way they imply forces, as neither of them is here used for disruption by compulsion, but they differ in the object's dimensionalities they occur with *snap* is here used with 1D objects, while *spezzare* is both used with 1D and 2D objects.

Parameter	**Value**	**Description**
Force	Compulsion	The force applied is delivered with an impulse and may result in an impact between the object and the agent or

		between the object and the surface with which the contact occurs, e.g., *she broke the vase with a hammer* or *il pugno ha spaccato una delle sue lenti* (itTenTen, "the punch smashed one of her lenses")
	Tension	Force is exerted on the far ends of an object, e.g., *he snapped the pencil with his hands* or *spezzare il rosmarino con le mani* (ItTenTen, "breaking rosemary twigs with hands")
	Pressure	Prolonged application of force perpendicularly to the surface of an object per unit area, e.g., *she sat on the branch, and it snapped* (EnTenTen) or *l'uccello si è posato sul ramo e si è rotto* (itTenTen, "the bird sits on the branch and the branch broke")
	Counterforce	Two forces colliding face-to-face, e.g., 'the particles smash together' (EnTenTen)
Degree of Modification	Overall modification	The initial state of the object is affected in its entirety (e.g., *I broke a vase in pieces/ho rotto il vaso in mille pezzi*)
	Partial modification	The initial state of the object is only partially affected (e.g., *the car was smashed on the passenger side* or *il tubetto tende a spaccarsi sui bordi* (itTenTen, "the tube tends to crack on the edges")
	Absent external modification	The object is rendered non-functional, but it is externally intact (e.g., *the computer broke/il computer si è rotto* when the device is externally intact, but it does not work).
Result State	*Fragmentation*	An object fragments or breaks into pieces and it is no longer recognizable as a whole (e.g., *I broke the glass bottle, un vaso che cade si rompe* (itTenTen, "a vase breaks when it falls")
	Separation	An object is separated into parts which have their own identity and function (e.g., '*a Paolo si rompe lo spallaccio dello zaino* (itTenTen, "Paolo's backpack strap breaks")
	Reshaping	The object is deformed (e.g., *he smashed the plastic bottle and threw it in the garbage bin* (enTenTen), *i rulli rompono l'acino* (itTenTen, "the rollers break the grape")
	No change	The object has no visible or external modification (same as described above with 'no loss of integrity')
Dimensionality	*One-dimensional object*	The object has the length as its most prominent characteristic (e.g., *a stick/un bastoncino*)
	Two-dimensional Object	An object that has length and height but less prominent depth (e.g., *a tile/una mattonella*)
	Three-dimensional object	An object that is roughly equidimensional in length, width and height (e.g., *a ball/una palla*)
Rigidity	*Flexible*	The object can be bent or is elastic (e.g., *a rope/una corda*)
	Non-Flexible	The object cannot be bent or deformed in any way without breaking it (e.g., *a pencil/una matita*)

Table 5. Annotation schema from Cacioli & Vernillo 2023

	Force	Modification	Result	Dimension	Rigidity
Romper e	all forces	all modifications	Fragmentation Separation No change	all dimensions	Flexible Non-flexible
Spaccare	Compulsion Pressure	Overall Partial	Fragmentation	2D, 3D	Non-flexible
Spezzare	Pressure Tension	Overall	Fragmentation	1D, 2D	Flexible Non-flexible
Break	all forces	all modifications	Fragmentation Separation No change	all dimensions	Flexible Non-flexible
Smash	Compulsion Pressure	Overall	Fragmentation Reshaping	3D	Flexible Non-flexible
Snap	Tension	Overall Partial	Fragmentation Separation	1D	Flexible Non-flexible

Table 6. Summary of parameters per verb in Imagact scenes

5.4 Results

Each parameter is investigated with respect to how its values are distributed within and across languages (stacked bar graphs in Figure 11). Looking at Force, the two general verbs, have similar frequencies for all forces, but it stands out that half of the occurrences are annotated as unspecified (*rompere* 53%; *break* 64%): as mentioned before, these are general verbs with inherent vagueness with respect to actional schemas, paired with the fact that in writing about events for real life purposes (with shared knowledge among the participants to the conversation or between reader-writer) a lot of information is implied and inferred. *Compulsion* is the dominant force for *smash* (88%) and *spaccare* (58%), while almost non-existent in *snap* (1%) and *spezzare* (8). *Smash* is the only verb having *counterforce. Snap* is most frequently associated with tension (54%), as well as *spezzare* (38%, despite being annotated for almost half of its occurrences as unspecified). Looking at Degree of Modification, the most frequent value is *overall integrity* across all verbs. As expected, *break* and *rompere* are the only verbs annotated with the values *no loss of integrity* (respectively, the 10% and the 13%) for occurrences in which non-functionality is involved rather than external, visible separation, deterioration, or loss of integrity. However, *smash* has also a small percentage of *no loss of integrity* (6%), attributable to *hit*-like occurrences of the verb (contact without change of state), which were not excluded from the sample because it could be ruled out completely that a result breakage was not occurring. There are a few occurrences in which the object is partially affected (e.g., "il contenitore è rotto nella parte

superiore" – Eng. 'the container is broken on the top'; "I broke [the glass] on one side, then slid the pane over so I could get my arm through without cutting myself to pieces").

For Result State, Fragmentation is the most frequent type of result change for all verbs. As shown in Figure 11, *spezzare* (100%) and *snap* (94%) are strongly connected to *fragmentation* while *reshaping* is the least frequent value (1% in *rompere* and a 2% in *smash*), followed by separation (5%, *snap* 4%, and *rompere* 1%). Rigid objects are the most recurring in our sample (Figure 6): *Snap* and *spezzare* are the two verbs with the stronger domination of flexible objects (respectively, 30% and 24%; compare with: *rompere* 6%, *spaccare* 2%, *break* 1%, and *smash* 1%). The most frequent dimensionality is 1D items for *snap* (81%) and *spezzare* (58%): *snap* is found mainly with body bones (arm, neck, foot, leg), cables, twigs and branches, while *spezzare* is found in our data with a more a heterogeneous subset of entities: hair, nails, teeth, muscles, leaves, bread. *Spaccare* and *smash* are instead strongly associated with 3D items (respectively, 58% and 63%). Tables 7 and 8 show the aggregated data for verbs and parameters, with the emergence of some clusters.

The typical configuration of *Rompere* and *break* alternate between unspecified force with overall fragmentation and no loss of integrity in the object. *Spaccare* and *smash* have similar configurations. *Spaccare* is found with overall fragmentation and compulsion force, with non-flexible items (25.61% of the times with 3-dimensional objects and 14.02% with 2-dimensional objects), while *smash* recurs with overall fragmentation by compulsion of 2D (17.67%) and 3D items (26.33%). Events described with s*pezzare* are mainly annotated with overall fragmentation of non-flexible 1D items with force unspecified or tension (respectively, 17.27% and 11.51%). All *snap* occurrences are found in overall fragmentation contexts, with 1D items, 20% unspecified for force and non-flexible, 20% non-flexible items with tension, 22.5% flexible items with tension. Table 7 shows the total percentage of the first three most recurrent combinations of parameters. As shown by the description above and by the table, our sample verbs do not have clear-cut profiles and clear boundaries in the parameters they select.

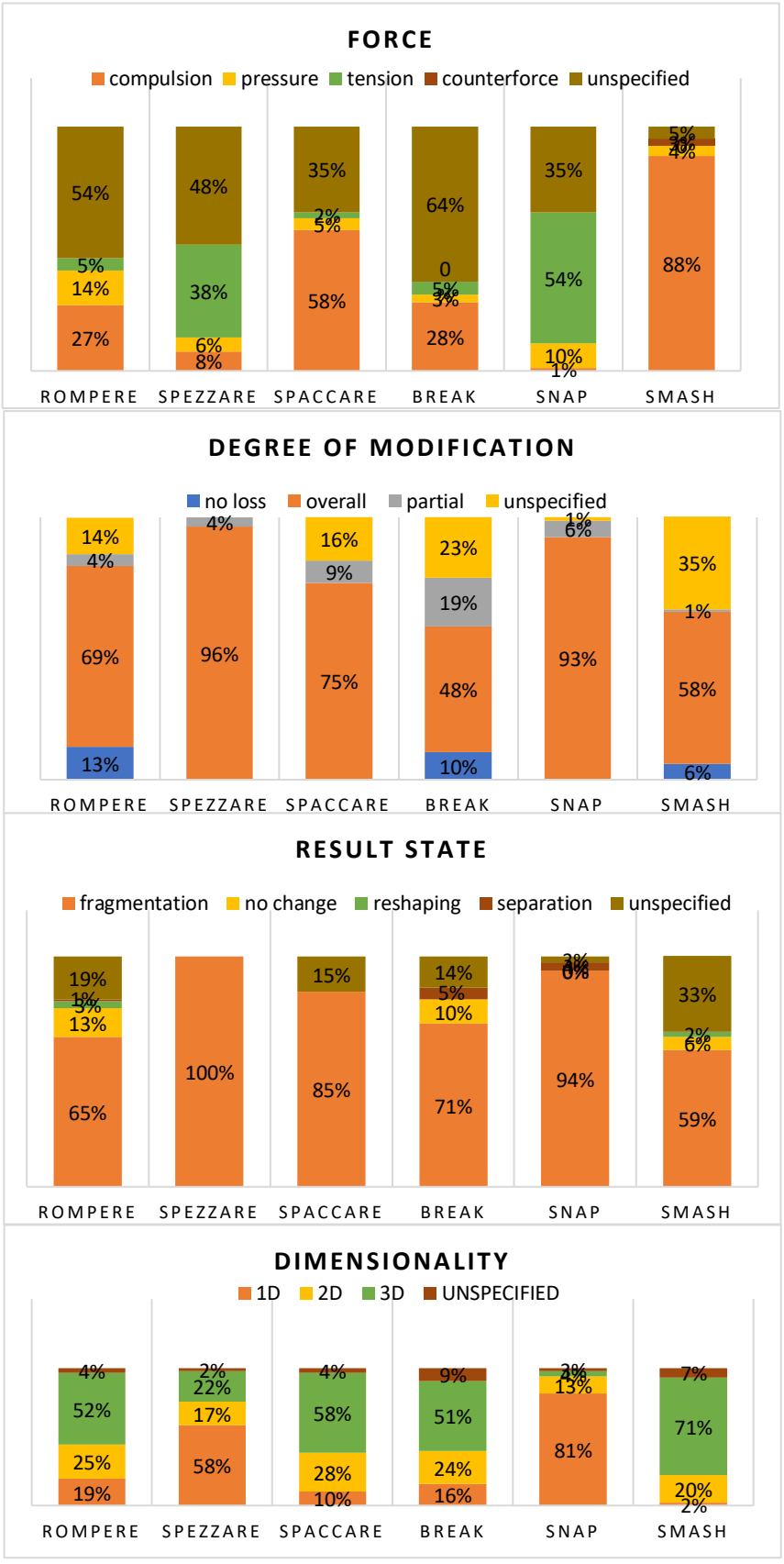
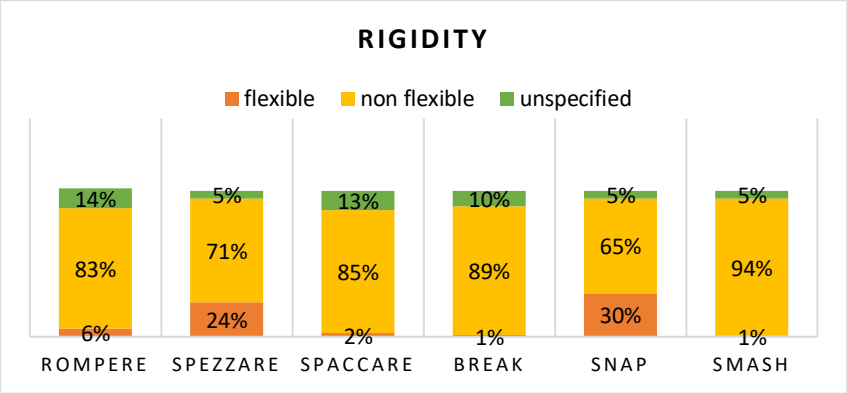


Figure 11. Stacked back graphs of each parameters' distribution across the verb



Continuing Figure 11.

MOST RECURRENT COMBINATIONS OF PARAMETERS					
BREAK					
FORCE	**DEGREE**	**RESULT**	**DIM.**	**RIGIDITY**	**%**
Unspecified	Overall	Fragmentation	3D	Non-flexible	10.4%
Comp.	Overall	Fragmentation	2D	Non-flexible	9.4%
Unspecified	No Loss	No Change	3D	Non-flexible	9.4%
		Tot.			29.2%
SMASH					
FORCE	**DEGREE**	**RESULT**	**DIM.**	**RIGIDITY**	**%**
Compulsion	Overall	Fragmentation	3D	Non-flexible	26.3%
Compulsion	Unspecified	Unspecified	3D	Non-flexible	20.3%
Compulsion	Overall	Fragmentation	2D	Non-flexible	17.7%
		Tot.			64.3%
SNAP					
FORCE	**DEGREE**	**RESULT**	**DIM.**	**RIGIDITY**	**%**
Tension	Overall	Fragmentation	1D	Flexible	22.5%
Unspecified	Overall	Fragmentation	1D	Non-flexible	20%
Unspecified	Overall	Fragmentation	1D	Non-flexible	20%
		Tot.			62.5%

Table 7. Aggregation of data for the first three most recurrent parameter combinations for the English verbs. (Cacioli & Vernillo 2023).

We used multiple correspondence analysis (MCA) as a visualisation tool (Levshina, 2015). This visualisation technique is suitable for this type of data as it displays the relationship between the values of multiple categorical variables – in our case, 5 parameters – and distributes the points on a space according to how strongly they are associated. MCA provides a typology of the individual observations (which, in our case, are the occurrences sampled from corpora), linking them to categories (Italian and English; the verbs) and describing them using the variables they are

annotated with (Husson & Josse, 2014) – that is, our parameters. The model is created in the RStudio environment using the package FactoMineR (Lê, Josse & Husson, 2008). We plotted the annotated data, adding the category *verb* as a supplementary variable (Figure 8) so that it could be mapped onto the space constituted of data points of the annotated parameters. To account for different aspects of semantic boundaries, we show three plots – each showing a different aspect.

MOST RECURRENT COMBINATIONS OF PARAMETERS					
ROMPERE					
FORCE	**DEGREE**	**RESULT**	**DIMENSION.**	**RIGIDITY**	**%**
Unspecified	Overall	Fragmentation	3D	Non-flexible	11.4%
Unspecified	No Loss	No Change	3D	Non-flexible	9.5%
Compulsion	Overall	Fragmentation	3D	Non-flexible	8.3%
		Tot.			29.2%
SPACCARE					
FORCE	**DEGREE**	**RESULT**	**DIMENSION.**	**RIGIDITY**	**%**
Compulsion	Overall	Fragmentation	3D	Non-flexible	25.6%
Compulsion	Overall	Fragmentation	2D	Non-flexible	14%
Unspecified	Overall	Fragmentation	3D	Non-flexible	9.8%
		Tot.			49.4%
SPEZZARE					
FORCE	**DEGREE**	**RESULT**	**DIMENSION.**	**RIGIDITY**	**%**
Tension	Overall	Fragmentation	1D	Flexible	11.5%
Unspecified	Overall	Fragmentation	1D	Non-flexible	17.3%
Tension	Overall	Fragmentation	3D	Non-flexible	13.7%
		Tot.			42.5%

Table 8. Aggregation of data for the first three most recurrent parameter combinations for the Italian verbs. (Cacioli & Vernillo 2023).

We are interested in looking at how parameters annotated are distributed in our sample of six verbs. Figure 12 shows the distribution of most characteristic parameters' values: Dimension 2 on the vertical axis separates occurrences denoted by a physical disruption from loss of functionality. The data annotation confirmed what emerged from the literature review about the physical affectedness of an object and its loss of functionality as the two main frames characterising the lexicalisations of BREAK. The model captures this component, showing that loss of functionality is visibly far from the other parameters, meaning that the two general verbs have these two frames that can occur separately where the loss of functionality frame can occur without an external, visible modification of the object (e.g., a computer, a washing machine, etc.).

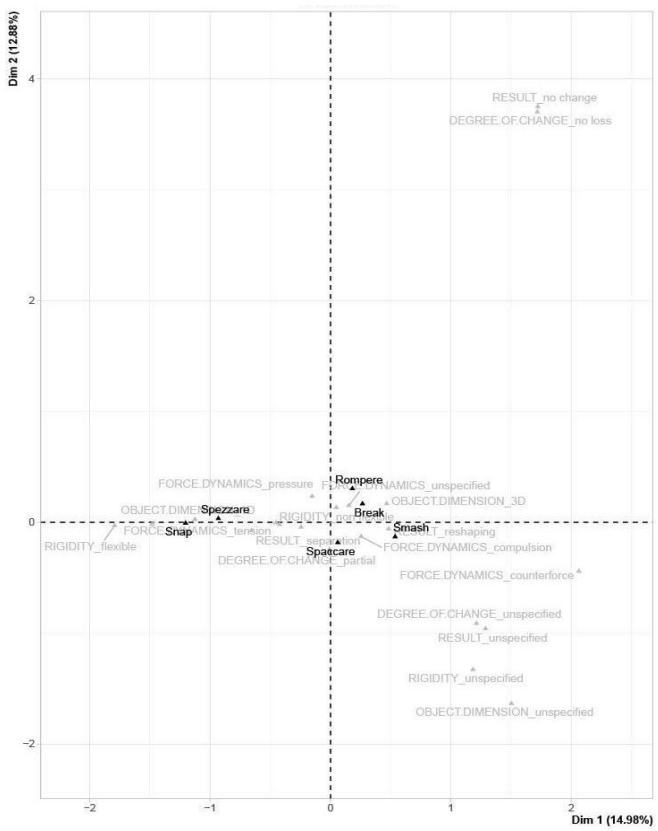


Figure 12. MCA map of the annotation on corpora occurrences

Only *break* and *rompere* occur in configurations of parameters with unspecified forces, with no loss of integrity and no change. Dimension 1 separates *snap* and *smash* on the values of flexibility of the object, object dimensionality (predominant 1D for snap and 2D-3D for smash) and force (predominantly tension for snap and compulsion for smash). *Spezzare* and *spaccare* are closer to each other (that is, share more parameters' values) than their English counterparts. For example: *spezzare* is mainly annotated with 1D objects but also occurs with a wide variety of 2D themes, as *spaccare*. The boundaries between them seem fuzzier than between *smash* and *snap*.

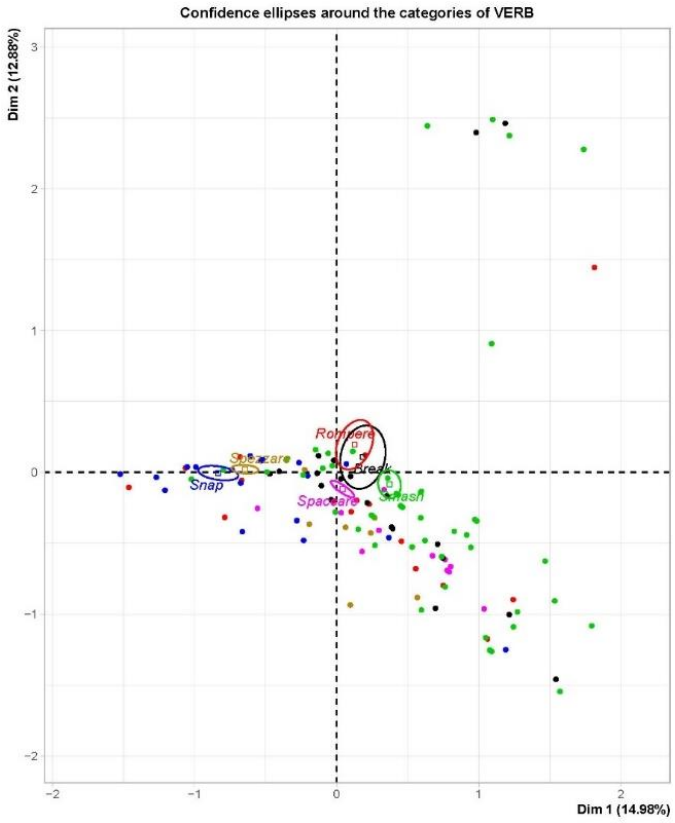


Figure 13: MCA map with confidence ellipses around the centroids of each verb

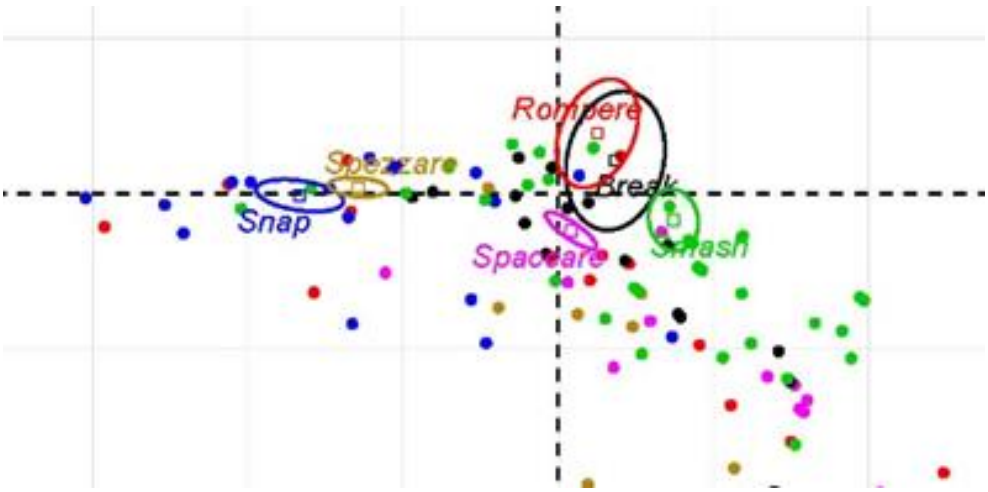


Figure 14: Zoom of the relevant areas of Figure 13.

Figure 13 shows the mapping of corpora occurrences in the semantic space, with centroid confidence ellipses drawn around clusters that the individuals form for each verb: these are the prototypical configurations of the annotated parameters (following Levshina, 2015). From a contrastive point of view, *rompere* and *break* are partially overlapped, meaning that their individuals are similar for the annotated parameters: this is read as the two verbs being close translation equivalents. The other four of our sample verbs do not overlap; their prototypical configurations are different. The smash-spaccare pair, has a significant difference in the force values: force is often unspecified for smash, because its meaning potential goes beyond breaking events (e.g., hitting events) and such these cases are not easy to judge and are potentially ambiguous in context. About *snap* and *spezzare*, *snap* has more vagueness concerning force than spezzare. In the annotation, examples such as 'the tornado snapped the tree' do not give indications or possibilities of inferring the force that could have brought about the event (pressure, compulsion, tension or a mix of force types). Although this can also happen with spezzare (*I capelli si spezzano con facilità*, "hair breaks easily"), it is less common in our sample. We take this to mean that other parameters are more central in the meaning of snap – such as the 1-D dimensionality of the object and the result state.

Confidence ellipses (Figure 15) around all exemplars for each verb show the extended area for each verb (beyond the prototypical configurations). Here, we can see overlapping areas for *snap, smash, spaccare* and *spezzare*. *Snap* and *spezzare* are considerably overlapped, while *spaccare* is entirely included in the extension of smash – showing that smash has a greater meaning potential. However, this shows the reasons and extent of their translation equivalence. Concerning boundaries within each language, the general verbs overlap almost entirely with the specific verbs for both languages. Moreover, it stands out that *spezzare* and *spaccare* overlap more than *snap* and *smash*.

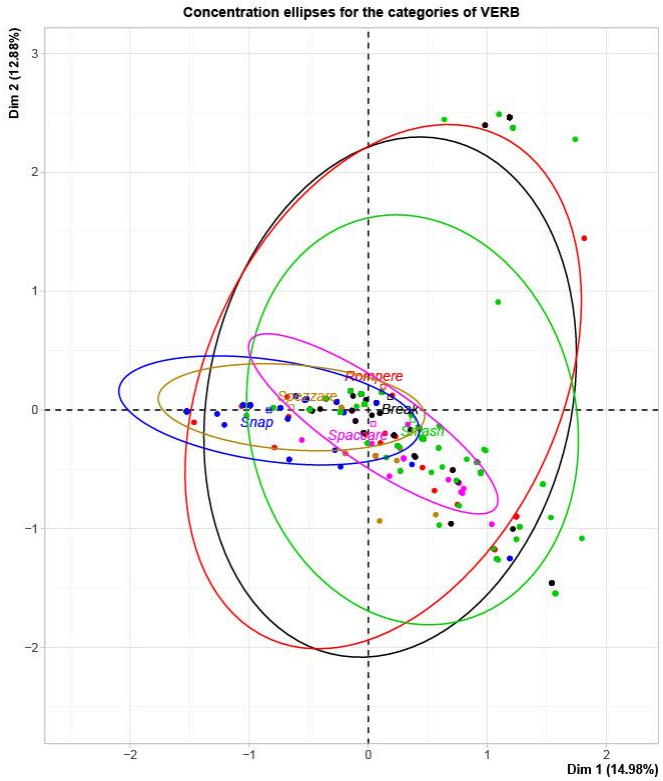


Figure 15: MCA map of confidence ellipses around the exemplars of each verb

5.5 Discussion

Following the procedure described above, Imagact facilitated the investigation of a set of breaking events using a world-to-word approach to mapping events and verbs. The corpora played a pivotal role in furnishing a usage sample of a set of verbs, thereby enabling contrastive research focusing on semantic boundaries and translation equivalents. The various occurrences from different texts and contexts provided complementary information to the Imagact resource. In this sense, the corpus part of the study took the opposite perspective - a word-to-world perspective, where the linguistic context was the key to infer the action schemas underlying the events described. We based the investigation on the premise that through an in-depth study of the context, we could get information about the semantics of the verb – what it is that in an event has to be described with a particular verb. Among the parameters, force and result state were those where we most often rely on inferences. For example, in sentence 4, breakage occurs through pressure, as hinted by the verb *premere* ('press, push'), while in example 6, through compulsion, as hinted by 'after a fall':

(3) Premere i bordi dell'impasto per rompere le bolle (kittyskitchen.it)

"Press the dough on the edges to break the bubbles".

(4) A gadget designed to keep the eggs from breaking after a fall (olemiss.edu)

This phenomenon aligns with the notion that verbs tend to be underspecified or general and acquire specific meanings in context and usage (Pustejovsky, 2017; Croft and Cruse, 2004; Viberg, 2020; Moneglia, 2014).

The annotated corpus occurrences revealed the most frequent parameter combinations in the samples, and MCA provided visual representations of the breaking domain. Results indicated that Italian and English use general verbs to describe breakage involving a modification in material integrity and/or loss of functionality, demonstrating no thematic restrictions and occurring with all the considered contact forces. The MCA visualisations highlighted typical configurations for each verb but also showcased occurrences drifting away from prototypical ones, resulting in overlapping areas. Corpus annotation further confirmed their nature as general verbs due to the high frequency of occurrences marked as unspecified, particularly for Force. Being annotated with a variety of the values of the parameters, they cover the other verb's semantic extensions. MCA visualisations show that they overlapped, showing a significant degree of translation equivalence.

More distinct parameters characterised the other four verbs in our selection. For instance, *spaccare* was associated with Force (compulsion) and non-flexible, 3D objects, while *spezzare* was linked to 1D and flexible objects, with tension as its predominant force. The English verb *smash* is associated with compulsion. Object dimensionality emerged as a crucial parameter for *spezzare* and *snap*. Semantic boundaries were found to be fuzzy, as exemplified by the peripheral use of *spezzare* with 3D objects. Concerning dimensionality, sentence (5) shows a peripheral use of *spezzare*: a 3D object with pressure (only occurring for 6% of *spezzare*):

(5) Un becco senza denti ed un palato con due ossa appuntite simile a quelle di cui sono dotati alcuni serpenti per spezzare le uova (unfuturoasud.it)
"A toothless beak and a palate with two sharp bones like those that some snakes use to break eggs".

The language pair's comparison indicated the degree of semantic correspondence between *snap-spezzare* and *smash-spaccare*, despite slightly different prototypical configurations. Additionally, a distinction emerged in the semantic relations of verbs between the two languages: *rompere-spaccare-spezzare* demonstrated closer and more cohesive internal relationships than break-smash-snap. This observation aligned with the semantic categorisation depicted in Imagact scenes, where spezzare and spaccare were enclosed in *rompere*, while *smash* and *snap* covered

actions were not included in *break*, suggesting less interdependence in the internal relations of the English set of verbs.

6 Conclusion

In this chapter, we introduced a video linguistic resource – Imagact – and the type of semantic investigation it allows, within languages and across languages. We explained how the resource can be used in Cognitive Linguistics research: it provides visual support to investigate actions and events. We compared the breaking events in Imagact as they appear with the English annotation of the verb *break* and the Italian annotation of the verb *rompere*. We provided a semantic analysis of these verbs and their local equivalences – that is, the other verbs that can be used to describe a set of events, irrespective of their status as verbs describing a *break* or not. This resource captures an event's possible varieties of construals: events are rich in information and have many aspects that can be chosen to describe.

We used these pieces of information to assess the extent of correspondences between verbs in a common domain. We selected a sample of three English verbs (*break, snap, smash*) and three Italian verbs (*rompere, spezzare, spaccare*), which resulted in pairs of translation equivalents. With this selection, we moved on to investigate the breaking domain and *break* verbs in corpora. We performed an extensive manual annotation of corpora occurrences of 5 parameters. The parameters emerged from the visual inspection of the sample of breaking scenes in Imagact and the aspects discussed in Chapter 2 as relevant in categorising breaking events. We visualised this data with a Multiple Correspondence Analysis (MCA), which showed the extent of semantic overlap between general verbs and specific verbs within a language and correspondence pairs between languages. We found that the general verbs extend similarly in English and Italian. The specific verbs in each language have their own semantic area and do not overlap if we consider their prototypical configurations. Italian verbs are closer to each other than English verbs (meaning that they share more parameters), and across languages we find that *snap* and *spezzare* are closer than *smash* and *spaccare*. Their correspondence is still visible in their extended configurations (rather than the prototypical ones). This methodology and analysis, by providing a qualitative evaluation and a visual display of a variety of annotated parameters, shed light on the extent of correspondence between verb pairs in two languages and allowed to find the categorial intersection between the semantically close verbs using both a world-to-word and word-to-world perspectives. We take a high degree of semantic overlap within verbs of one language as indicative of fuzziness of categorisation: the categories of spezzare and spaccare share

more members than snap and smash. However, as far as typicality is concerned, their typical members are similar.

Chapter 4. Event Categorisation and Construals

1. Introduction

Following up on Study 1, we set out to investigate event categories and their relationship to linguistic form-meaning mappings. The study centres explicitly on the linguistic categorisation of breaking events through predicate forms. The study emphasises linguistic construals as dynamic components of linguistic expression within a cognitive framework of semantic representation. As discussed in Chapter 1, construals, viewed as negotiable meanings in context (Paradis, 2015), have a pivotal role in accounting for the flexibility of language use by individuals and categorisation. To follow this premise, we look at instances of everyday actions relevant to our experience of the world and interpret the construals of each event as the speakers' conceptualisations. Contrary to previous approaches on event categorisation (Rifkin 1985; Morris and Murphy 1990), which refer to activities in Vendler's (1957) and Croft's (2012) sense (*housework activities; school activities*) and prompted nominal responses (e.g., cleaning, sweeping, washing-up), we focus on categorisations achievements as dynamically expressed by verbs. Compared to Study 1, the scope of this study is extended to a sample of three languages – English, Swedish, and Italian – provides a more refined set of scenes and introduce elicited written descriptions of the events for construal examination.

We use a video-description task with native speakers of English, Swedish and Italian under the assumption that cross-linguistic diversity accentuates the properties of meaning relevant to form-meaning mappings (König 2012). The research operationalises the study of meaning and categories through the analysis of construals, codability scores, and semantic maps.

Semantic maps provide a visualisation of the network established between construals (across participants) and events. They are used to empirically investigate meaning and show the relations between a grid of concepts (or senses, or meanings - see Cysouw 2010) and lexical forms. Moreover, they allow us to visualise linguistic diversity in investigating meaning. If categories operate on similarities (in perceptual characteristics, functions, and goals…), semantic maps offer a solution to operationalise similarity and linguistic categorisation. As discussed in Chapter 1-Section 2 and 4.1, categories operate on similarities (in perceptual characteristics, functions, goals…). We operationalised the study of semantic categorisation by analysing the collection of the construals used to describe event by our sample of population. Codability, as discussed in Chapter 2, is taken as the measure of how a group of speakers of the same language coherently conceptualises an event. Codability can be quantified using Simpson's Diversity Index, which aims

to show how a scene can be interpreted and described with various verbs. One aspect stressed in Chapter 2 was related to how usual or unusual the event is as a parameter that affects codability.

Compared to previous studies, linguistic categorisation is here investigated using stimuli from a video repository that were not created for the purpose of the study, aiming to eliminate a possible stimuli-creation bias.

The primary inquiry of this study is the cognitive status of semantic categorisation of events, specifically focusing on the levels of abstraction represented in individual speakers' construals. First, we determine the relationship between conceptual and linguistic categories, addressing how conceptual categorisation and construals interact by looking at lexical diversity and the variety of non-*break* verbs speakers produce to refer to the breaking domain. Furthermore, the research investigates the extent to which linguistically derived categories align with cognitive theories of categorisation, particularly concerning the basic-level theory: this is studied considering the level of abstraction of each verb and, in operational terms, the generality and specificity of the verbs.

2. Procedure and Data

The investigation unfolds in three phases. First, the conceptual domain is delimited by selecting events conforming to an operational definition of breaking events from a multilingual video resource (Imagact – www.imagact.it). This step ensures comparability across languages and minimises the researcher bias in stimuli selection. Second, we inspect how these verbs are used in the resource to describe these events, comparing lexical items across the three languages. In the following section, we describe this procedure. In the last phase we proceed to set up the video description task and collect video descriptions.

2.1 Step 1: Visual inspection

The first step aims to find an abstract category of breaking events at the conceptual level, irrespective of the verbs used to describe such events, to form a set of stimuli that is not task-specific and can be used for cross-linguistic investigation. To achieve this aim, we used Imagact, an online video resource (www.imagact.it). Upon visual inspection, 35 scenes conform to our definition of breaking events described In Chapter 2-section 6.1 (see Appendix A), the selection comprises a variety of actions featuring different action schemas, goals and objects.

2.2 Step 2: Linguistic inspection and selection of data

We checked which verbs describe the set of events isolated in Step 1. We take the predicate as the unit of analysis, according to the definition used in Van Staden and Narasimhan (2012): "a complex of one or more relational entities in the clause with a predicative function (including simple verbs, verb + particle constructions, verb + light verb constructions, etc. but excluding modifying elements such as adverbials)."

The general verbs discussed above for the three languages (*break, rompere, ha sönder*) are found in 16 scenes for English, 15 scenes for Italian and 9 scenes for Swedish. Having considered the literature about the breaking category in Swedish, we further consider two aspects relevant to the selection of the scenes:

1. The literature review revealed that the particle *sönder* is especially relevant. Considering the discussion on particles and the need to account for their role in event descriptions (Viberg 2020), we consider *sönder* a *general* expression of the breaking domain.
2. As described in Viberg (1985, 2007, 2020), *bryta* expresses a great variety of breaking events without overlapping with the most general *ha sönder*. We also consider *bryta* when selecting scenes.

Thus, we extend the selection of scenes to those predicates that are combined with the particle *sönder* and to *bryta*. After checking the linguistic mapping on the selection of breaking events, we extend the search on Imagact to each *break* verb. The resulting scenes relevant to the general *break* verbs in our sample of languages, as represented in Imagact, are 23. The mapping of verbs and scenes do not overlap across languages (see Appendix B and C) in the selected events:

- Of the initial 35 scenes, not all of them are described by a *break* verb in Imagact
- There are scenes annotated with *break* verbs that were not captured by the initial selection

There is a substantial degree of overlap between the conceptual domain as selected by our working definition and the linguistic mapping with *break* verbs in the three languages under consideration, but nonetheless, it is a partial overlap. The events left out from the selection, which conformed to the definition of *breaking event* but are not described with a lexical item that focuses

on the breakage, are described with verbs that pick up on different aspects of the event – such as *crushing, cutting, grinding,* and *dividing*. Thus, at this stage, we can interpret the two selections as a demonstration that a (ill)defined[13] conceptual domain is not overlapped to the linguistic categories.

In this step, we also notice that verbs not belonging to BREAKING can describe breaking events. An example is the English verb *open*: forcefully breaking a locked safe or a door can also be described as *opening a safe/a door*. In these events, it is impossible to reestablish the previous closing system, so it is defined as a breaking event because of the loss of functionality. At the same time, the agent gets access to the content, thus configuring an opening event. However, there is cross-linguistic variation in how events may be construed according to a perspectivization: the door scene in Italian is not described with *rompere,* and the safe scene is not described with *ha sönder* in Swedish.

Steps 1 and 2 served to create a category for investigation and to limit the scope of our study, checking the linguistic mapping of the general *break* verbs in the three languages under investigation. These languages' variation in categorising breaking against a common (however limited) universe of events is only partial. The next step focuses on the events described by at least one general *break* verb in at least one of the languages investigated (see Appendix B for the list). This means that the final selection contrasts three languages on a set of events described by a *break* verb among *break, rompere* and *ha sönder*, but not all the verbs describe all the scenes. Following this operation, our selection comprises 25 videos.

2.3 Step 3: Elicited descriptions

We asked native speakers of Italian, English, and Swedish to describe a set of breaking events to investigate the different possible construals of these events and the variation between speakers at the preferred level of abstraction. The level of abstraction to discuss hierarchical categorisation is operationalised in terms of semantic specificity of the verbs.

[13] As explained in Chapter 2, the domain has been defined differently on previous literature. However, as stressed previously and although according to our conceptual framework we do not subscribe to a view of lexical or conceptual meaning based on definitions and conditions, it seemed a necessary step to restrict the area of investigation with a working definition.

2.3.1 Participants

We collected video descriptions from 30 speakers per language, recruited through Prolific for English and Italian and at the University of Lund for Swedish[14]. In both cases, the task was performed online. Participants were students at the undergraduate or master's level, aged 18-28. They declared themselves to be native speakers of the language and mainly monolingual (defined as i. language acquired via natural process and ii. being raised with only one language). Participants were asked to describe short video clips (N=40), of which 25 were the target actions for the study (Table 3 and Appendix 2). The target actions are all agentive and intentional and combine different action qualities: forces, manners/instruments, and materials (dimensionality, softness/rigidity). Participants were asked only to write one sentence, conforming to a specific structure to ensure comparability among the languages. They were asked to start with a subject, in the form of a person's name of their choice, continue with a verb and add other elements relevant to describe the event (see instructions in Appendix F). They were not given instructions or clues on how to interpret the video. For example, they could describe the same event as "opening the door" or "entering the room".

2.3.2 Materials

The 25 target scenes are described in Table 3 below with the features that characterise them and have been discussed in the literature as relevant in breaking lexicalisations across-languages (Majid et al. 2007; Stathi 2023): effector (whether it is performed with body parts or tools), dimensionality (object which mainly extends on one dimension, such as a pencil, two dimensions, such as a tile, or a rounded, three-dimensional one) and performance of action (whether the event features a single action, such as a single hit toward an object, or repeated hitting actions). In the case of Majid et al. 2007 and Stathi 2023, the stimuli were created for the purpose of the study (including a variety of objects of different dimensionality and flexibility) and instruments combined in different action performances in usual or unusual events (Bohnemeyer, Bowerman & Brown 2001). In this study, the stimuli are not manipulated to have a balanced representation of these features. As mentioned in section 3, the stimuli are not created for this task but for general purposes (much like general-purpose corpora), and the actions are performed in real-life background contexts. This aims at giving ecological validity.

[14] The different types of data collection are due to practical reasons, as the target population for Swedish was available to the authors at the time of data collection *in situ* at Lund University.

SCENES	Effector	Dimension	Part-whole relations	Performance of action
Smash bottle on a wall	Hands	3D	Whole	Single action
Wrestling-breaking arm	Hands	2D	Whole	Single action
Breaking tile with tool	Screwdriver	2D	Whole	Single action
Breaking a wheel off a toy car	Hands	3D	Part of a whole	Single action
Make a fan non-functional	Screwdriver	3D	Whole	Continuous
Crack a mirror with tool	Screwdriver	2D	Whole	Single action
Force with a tool to open a safe	Crowbar	3D	Part of a whole	Single action
Computer display not working	Hands	3D	Whole	Repeated
Divide twig in parts	Hands	2D	Whole	Repeated
Divide pencil in two	Hands	1D	Whole	Single action
Detach a branch from a plant	Hands	1D	Part of a whole	Single action
Smash egg with a hand	Hands	3D	Whole	Single action
Smash wooden box with hammer	Hammer	3D	Whole	Repeated
Create a hole in a wooden box	Hammer	3D	Whole	Single action
Break chair separating legs from the back	Hands	3D	Part of a whole	Single action
Smash furniture in a room	various tools	3D	Whole	Repeated
Peel peanuts	Hands	3D	Whole	Continuous
Crack walnuts	Nutcracker	3D	Whole	Single action
Smash car on a wall	-	3D	Whole	Single action
Crack an egg	Hands	3D	Whole	Single action
Tear paper in two	Hands	2D	Whole	Single action
Tear paper in four	Hands	2D	Whole	Repeated
Break open door	Shoulder	2D	Whole	Single action
Smash tomatoes with machine	Mechanical machine	3D	Whole	Repeated
Divide rope in two by pulling the extremes	Hands	3D	Whole	Single action

Table 1 List of events included in the event description task with their visual features.

2.3.3 Coding

The responses were collected, and the predicate form was coded in a matrix with the events in the first column and the predicates per participant in the rows. Incomplete descriptions, video fails, and misinterpretations were discarded. The predicates were taken as a single lemma or as verb + particle. For the predicate to include a particle, the construction had to conform to the following criteria:

- the verb + particle construct is conventional (e.g., found in lexical resources)
- The particle contributes meaning that relates to the manner or result of the action – e.g., in a resultative construction such as 'break the door open'.

When the same verb is found with different particles, we evaluated whether they contributed differences in meaning that are relevant from an actional point of view. We checked whether the other particles were used as alternatives in the same event only or in different events. In the second case, we coded one verb-particle construct only. The only case in the English data is reported in example 1, and examples 2-3 report the cases for Swedish.

(1) a. barge into a door
    b. barge a door in
(2) a. plocka av
    b. plocka bort
(3) a. tackla upp
    b. tackla i

3. Results

The analysis considers the lexical diversity that emerges from the participants' construals assessing the quality and quantity of a variety of verbs per language to investigate categorisation through network-based semantic maps.

3.1 Lexical diversity

Starting from lexical diversity, Table 2 reports the number of predicates per language and the median verbs per event with the median of predicates per video. Total verbs with their counts are reported in Appendix D. Swedish and English participants used a greater variety of predicates than Italian. This is consistent with the typological distinction between satellite-framed and verb-framed languages, which predicts a higher type of number for satellite-framed languages, testifying higher *granularity* (Stathi 2023) for English and Swedish. Accordingly, the median of

predicates per event is lower for Italian. However, the quantitative difference is cancelled out if we count verb roots rather than predicates.

	English	*Italian*	*Swedish*
Total responses (690 possible answers per language)	670	653	665
Number of Predicates	101	67	100
Of which verb roots	69	67	63
Median of Predicates per event	8	6	8

Table 2 Total responses with the number of predicates produced per language and the median value per event.

	English		*Italian*		*Swedish*	
	Breaking	Other	Breaking	Other	Breaking	Other
Type	36	65	22	45	46	55
Tokens	463	207	443	210	501	164
TTR	0.15		0.10		0.14	
Median	0.77		0.51		0.91	
SDI	0.73		0.64		0.72	

Table 3 Types, Tokens, Type-Token Ratio (TTR) and mean of Simpson's Diversity Index (SDI).

The resulting construals differed in the aspects of the event under focus. It emerged that construals chosen by participants involve verbs belonging to the break category (= the meaning of the verb includes a material disintegration and/or loss of functionality) and other verbs evoking different domains. For all languages, the verbs evoking the break category are used in most cases: 69% for English, 68% for Italian, and 75% for English. The general verbs – *break*, *rompere*, *ha sönder* – are used, respectively, in 22%, 26% and 4% of the cases. This result shows an unexpected divergence of English from the satellite-framed typology: it is expected for satellite-framed languages to have a lower token frequency of general verbs compared to verb-framed languages (Stathi 2023). The incidence of *ha sönder* in the Swedish data is considerably low if compared to other lexical items: *knäcka* 11%, *slå sönder* 10.5%, and *bryta* 7%. If we take together the tokens for the predicates composed of *bryta* and a particle, this is the highest value in our data: 16.5% of the construals. Considering the counts of the particle *sönder* combined with different verbs, these account for 21% of the construals.

To assess the diversity of construals in the domains they evoke, the verbs are investigated for their semantic contribution to the description of the event and classified into verbs from the break category and from other domains. An example is provided in Table 4: the event can be interpreted as the material disintegration of the object or the access to its content. Moreover, the focus can be on specific sub-events, such as the compromission of the outer layer of the object or

its removal. The complete list of verbs divided per domain is available in Appendix C. All quantifications – TTR and SDI – used to assess linguistic diversity seem to show that there are no relevant differences between the languages. SDI means are not significant in one-way ANOVA ($F = 1.25$, $p = 0.29$). SDI is further evaluated in section 4.2.

Event*: A person crushes a peanut and removes the skin with hands*		
Perspective	**Construal**	**Verb**
Sub-events	Removal of the outer layer	*Take off, remove, shell, de-shell, peel, peel off*
	Outer layer compromission	*Crack, crush*
Whole event	Access to content	*Open*
	Material disintegration	*Break, break apart*

Table 4 Construals of the event: a person crushes a peanut and removes the skin with hands.

	Semantic descriptors	**English**	**Italian**	**Swedish**
Functionality	*Compromise functionality without external modification*	Damage	Manomettere	
Results	*Disruption in pieces*	shatter, rip, shred, tear, tear open, crack, snap	Fare a pezzi, fare a pezzetti, ridurre in pezzi, stracciare, strappare	riva (+ sönder, isär, itu, upp, av), splittra, spräcka, slita (+ loss, av, isär), dra sönder (+loss, isär), knäcka, rycka (+ isär, sönder bort), ta itu, ta isär, knäcka
	Rough disruption	Burst, demolish, destroy, smash,	Distruggere, spaccare, sfasciare, scassare	Demolera, förstöra, slå sönder, krossa
	Rough disruption with a tool			Hamra sönder, banka sönder
	Divide into pieces in a controlled way with some effector (hands or tool)	Chop	Recidere, sminuzzare	
	Creating an opening	Break open	sfondare	ha hål, slå hål
	Separation of pieces from a whole	Chip		
Implication of breaking	*Create an opening or a hole*	force open, make a hole, put a hole	Bucare	slå genom, bända upp
	Apply pressure or create tension that result into loss of integrity	Crush	Schiacciare, spiaccicare, spremere, spezzare, spezzettare	trasa sönder, trycka sönder, skruva sönder

Table 5 Cross-linguistic classification of verbs according to semantic contribution

Within *break* verbs, *break, rompere, and ha sönder* are at the most abstract level. Verbs from the breaking domain are reported in Table 5 and classified according to the aspects of their semantics to capture similarities between the languages and better account for patterns of specificity. The verbs classified as *break* verbs evoke mainly the domains of compromission of functionality (even without the material loss of integrity), type of result in pieces, or they pick on aspects of the event that evoke a material change of state and *may* implicate a *breaking* through our world knowledge or causal links. As an example of this last class, the English *force open:* using force on an object until the object is opened is, by common knowledge, an event in which a material disintegration may occur, with loss of functionality. The construals that evoke different domains (e.g., *separate*, *open, divide*) do not appear to have a superordinate category that is inclusive of the break category. Regarding the levels of abstraction, the three languages all have different levels of generality and specificity in domains shared across the languages or language-specific domains.

3.2 Codability

To better account for the role of frequency of occurrence of each verb in each scene and the levels of abstraction chosen in the verbalisations, we used the Simpson Diversity Index (SDI). Drawing from Majid et al. (2008) and Majid et al. (2021), the SDI is taken as a measure of diversity in the responses and interpreted as the *codability* of a scene. The index is calculated with the *Vegan* package in Rstudio with the 'diversity' function, which returns a value between 0 (no diversity) and 1 (full diversity). No diversity means a scene is referred to with only one verb, while full diversity means that all participants used a different verb for a single scene. In our sample, only Swedish scenes are coded with 0, which means there is no diversity (*break an arm* and *crack an egg*). As mentioned above, the differences among the languages are not significant. Table 6 shows that each language has low standard deviation values, showing that the data is clustered around the mean (see Figure 1). We take this to mean that there is no linguistic-internal pattern in the SDI. However, we investigate which scene each value corresponds to across the languages (Figure 2).

Language	**Mean**	**Median**	**SD**
English	0.69	0.73	0.18
Italian	0.6	0.64	0.18
Swedish	0.65	0.72	0.26

Table 6 Mean, Median and SD values for Simpson's Diversity Index (SDI)

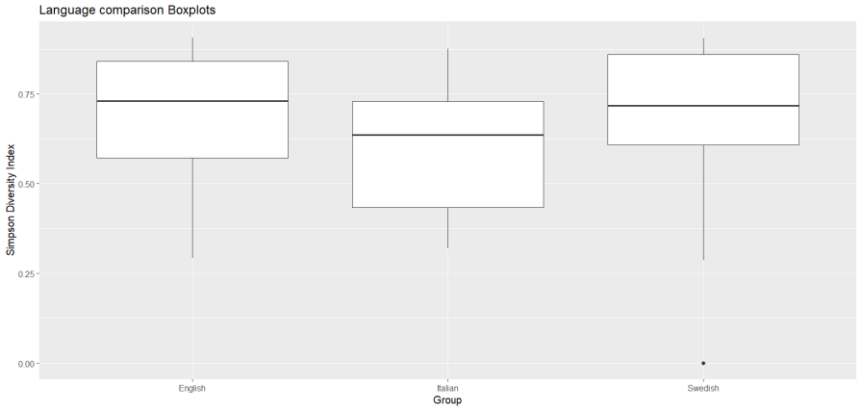


Figure 1 Boxplot of the Simpson's Diversity Index (SDI) means in English, Italian and Swedish

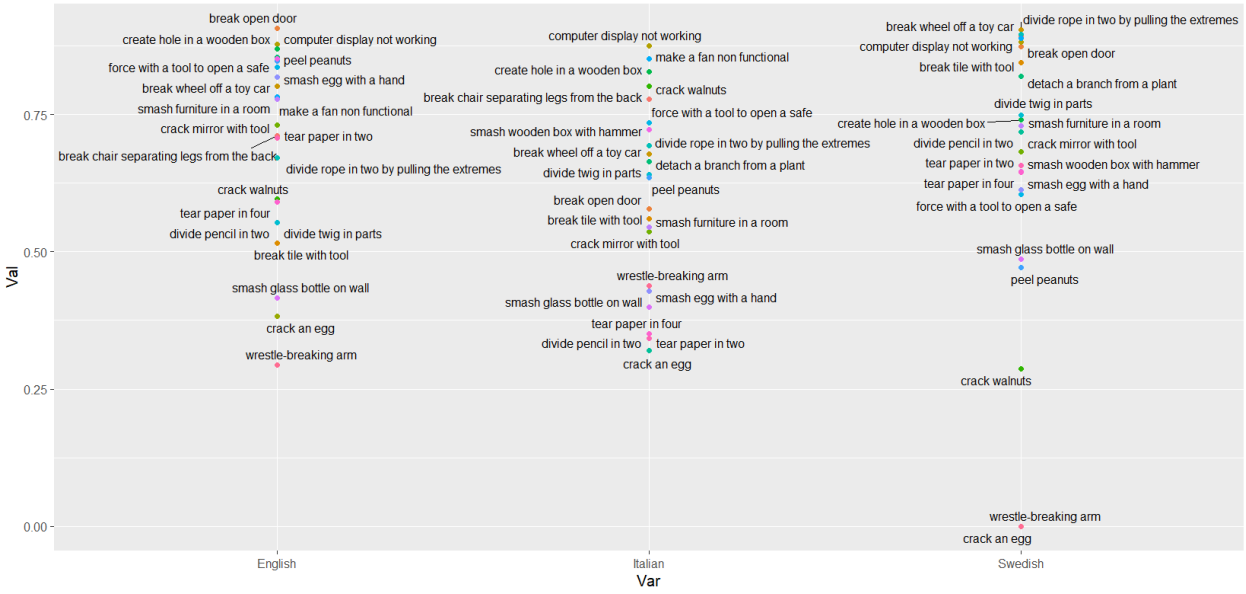


Figure 2. Simpson's Diversity Index (SDI) of individual events per language.

Figure 2 shows the SDI of each scene in the three languages. Swedish's three most codable events are 'crack an egg', 'wrestle-breaking arm' and 'crack walnuts', where the first two events have an SDI of 0, that is, no diversity. The qualitative result is similar for English, whose most codable events are 'wrestle-breaking arm', 'crack an egg' and 'smash a glass bottle on the wall'. Italian's three most codable events are 'crack an egg', 'divide pencil in two', and 'tear paper in two'. Their distribution is qualitatively different but similar if we look at the numbers only (Figure 1).

3.3 Categorisation through semantic maps

The relations described above can be visualised with semantic maps. We use a connectivity map (Croft 2022; van der Auwera 2008; Georgakopoulos & Polis 2018; Georgakopoulos & Härtl 2022, Georgakopoulos & Polis 2022) to visualise the relationship between construals and the event stimuli. The data set is visualised in a graph with nodes and edges that show the frequency of association of each verb to each event across participants and within language. This type of map was chosen over other semantic map approaches. For breaking events, similar maps have been used: multidimensional scaling in Stathi (2023) and correspondence analysis in Majid et al. (2007). The advantage of the connectivity map for our data lies in the representation of the relationship between events and verbs, together with a measure of how strong the relationship is in our sample, which is determined by how many times speakers chose a specific construal. The graphs are based on counts of occurrences in the sample (that is, how many times participants used a verb for each event) rather than on a yes/no matrix, which would represent whether an event can, in principle, be described with a specific verb.

Table 7 shows an extract of the data: each scene is repeated in a column as many times as verbs that describe it. The database is elaborated in R to fit the *igraph* package (Csardi & Nepusz 2006) requirements to plot a network and is rendered dynamic with the visNetwork package[15] (copyright holder: Benoit Thieurmel).

Scene	**Predicate**	**Counts**
break tile with tool	break	20
break tile with tool	crack	5
break tile with tool	push into	1
break tile with tool	shatter	1
break tile with tool	smash	3

Table 7 Extract of the coded file for English used for plotting the graph.

The count column in Table 7 is fundamental to our research question. The literature often uses semantic maps to show relations between possible meanings and lexical items. Still, they do not account for the relevance of lexical items in speakers' verbalisations.

The graphs in Figures 3-5 have thicker edges when the relationship between a predicate and an event is stronger. The arrows go from an event to its verbs (meaning that events are distinguished as the dots with arrows going out). The colours represent the community detection:

[15] https://datastorm-open.github.io/visNetwork/,

community detection algorithms are a tool for clustering and apt for contrasting different languages. The algorithm individuates a set of shared properties in the relations between verbs and clusters the nodes that are densely connected (Louvain Method - Blondel et al. 2008). The maps are dynamic and can be interacted with at the following links:

- http://lablita.it/app/visnet/visnetSWE.html
- http://lablita.it/app/visnet/visnetITA.html
- http://lablita.it/app/visnet/visnetENG.html

Each resulting categorisation is visually different in terms of the thickness of edges and directions of relations. In Swedish, the events and predicates are spread out in the semantic space, which means that each event shares fewer verbs. This is the result of the higher type count for Swedish. However, the English type count is similar, but the relation is visualised differently. In fact, it is dependent on the three-way relation of types, tokens and scenes. The difference in the Swedish map can be explained by pointing out the use of the general verbs in Swedish: *ha sönder* is used in 8 events out of 25 (while *bryta* occurs in 9), against the English *break*, which is used in 17 events and the Italian *rompere*, used in 20 events. This is visible in the English and Italian maps, where a core of events sharing verbs is clustered together with more, thicker edges. Reading this in terms of communities, Italian and English establish a central network of convergence of arrows with their general verbs. The effect of a populated core does not result in the Swedish categorisation. The verbs with the most connections in Swedish (beyond the general verb) are *knäcka* and *slå sönder*.

It is interesting to notice that, in all languages, each scene is connected with each other. The events are, in fact, selected according to the mapping of verbs onto breaking scenes, taking together all mapping for the three languages. This means that not all scenes were described by a general *break* verb in each language. To this, we add that construals were evoking domains different from the breaking one. Nonetheless, the scenes are all interconnected.

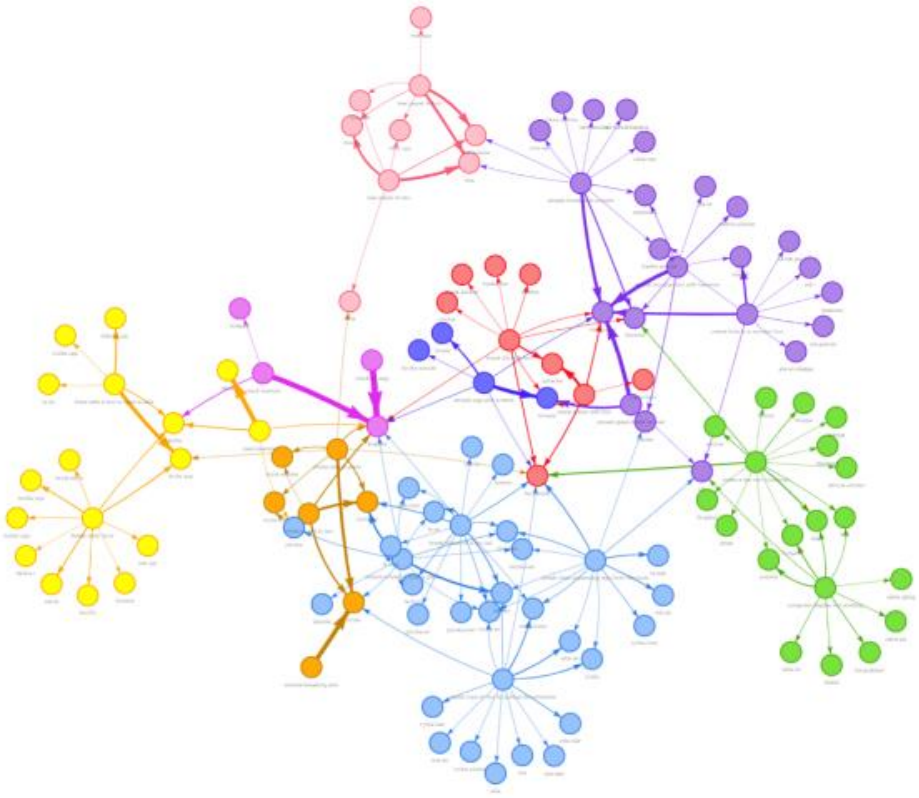


Figure 3 Swedish semantic map

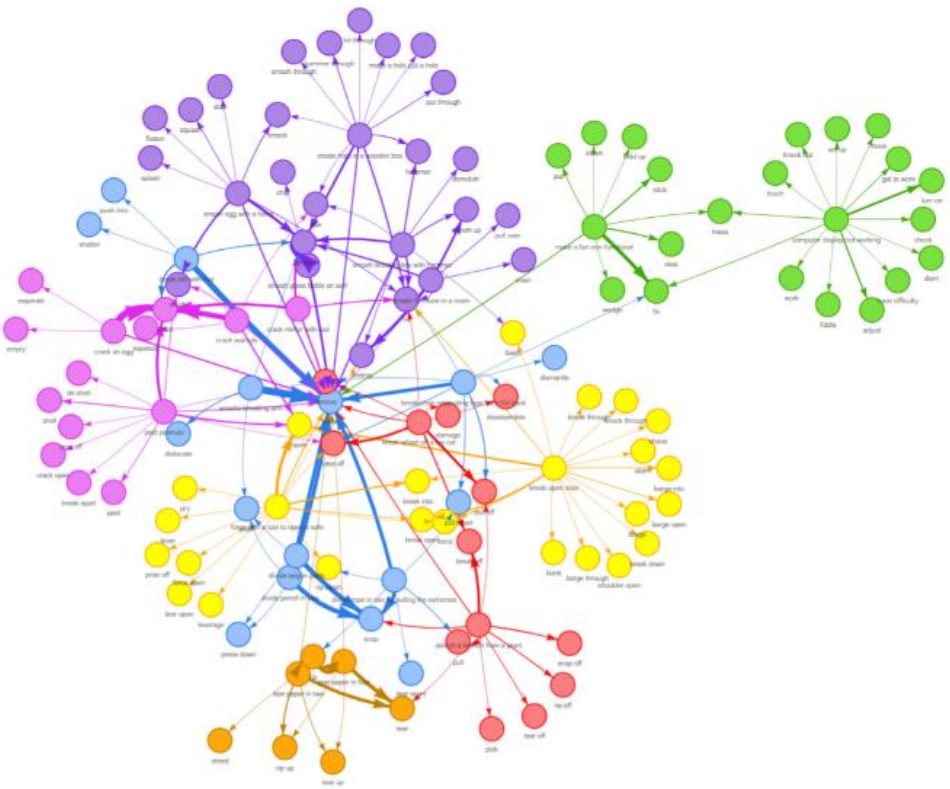


Figure 4 English semantic map

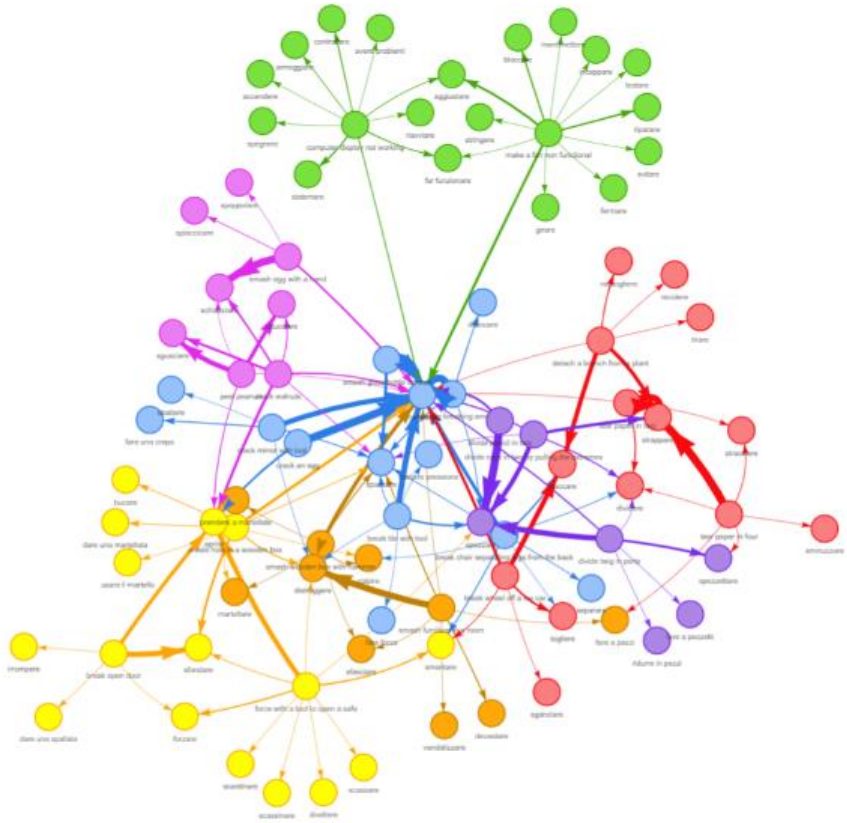


Figure 5 Italian semantic map

The detection of communities reveals the network established between verbs within this set of events: six communities are identified for English and Italian, while 9 for Swedish. Only one community is shared between the three languages: the two events with electronic/motor (break a computer monitor and break an electric fan) devices linked to the concept of functionality. Swedish and English share one community: a tearing cluster, 'tear paper in two' and 'tear paper in four' (in the map: orange for English; pink for Swedish). This cluster in Italian also includes 'break a wheel off a toy' and 'detach a branch from plan' (in red in the map). This community can be an example of how community detection works. In English and Swedish, these events are described with verbs that do not describe other scenes in a recurrent manner. In Italian, 'tear paper in two' and 'tear paper in four' are mainly described with the verb *strappare* ('tear', 'rip'), which is also strongly linked to 'detach a branch from plan'. In turn, 'detach a branch from plan' is strongly linked to *staccare* ('detach'), which is strongly associated to 'break a wheel off a toy'.

The one-dimensional properties of objects are responsible for establishing a community in Italian (purple community, defined by *spezzare*) and Swedish (orange community, represented by *bryta)* but not in English. Swedish is sensitive to the parameters of 2-dimensional, brittle objects acted upon with an instrument ('break tile with tool'; 'crack mirror'; see the red

community). Swedish is strongly associated with the verb *spricka*, which describes a type of result (a split in the integrity) achievable with a brittle, thin material.

4. Discussion

The results and analysis of the construals used to describe the set of breaking events show that the three languages' predicates for the break category map differently to a set of instantiations of breaking events.

Overall break category

Despite the differences in mapping *break* verbs to events and other domain verbs, the maps show a shared cross-linguistic categorisation of breaking, with verbs that interact in a network of intertwined categories. Construals show that actions are also described with verbs belonging to different categories (e.g., *open*, *separate, divide*), showing that an event can be conceptualised from different perspectives. What emerged from the construal-based graphs is that despite the linguistic mapping not overlapping for these three languages, they still share how other predicates pick up on aspects of the events.

Comparison of categorisation to previous literature

The graphs show a categorisation that is coherent with the results from Majid et al. 2008 as far as Swedish and English are concerned: Swedish has more categories than English, which is visually shown by the clusters of events spread out in the map. However, the emerging verbs are different - also due to the different stimuli set. For example, the verb *hugga* ('chop') is the main *break* verb in Majid et al. (2008) but is not attested in our data. Unlike previous literature, more attention is paid to verb particle units, which brings other predicates into discussion. Moreover, a tearing cluster appears separate from the *break* cluster in English (contrary to Majid et al. 2008 and Stathi 2023), as well as in Swedish.

The codability scores do not reveal a language-internal pattern: diversity is not a property of the languages according to the results from this data set. The events tested here show that languages may have conventionalisations (e.g., Swedish construals of the 'crack an egg' event), especially for everyday events, that override bigger-picture typological patterns (e.g., lack of a general verb for 'crack an egg') and could not be captured in a fine-grained investigation. Related to conventionalisation and mapping imposed by languages, the codability scores show that

physical or conceptual continuities of some events do not lead to a coherent linguistic categorisation. For example, a group of scenes in the semantic domain of accessing *the content of an edible object* ('crack an egg', 'peel a peanut', 'crack walnuts') are linked in the networks because they have verbs in common but have different codability scores in the three languages and different lexical variation.

Generality-specificity

In terms of patterns of granularity, we found that the count of types is consistent with the typological distinction between satellited-framed and verb-framed languages, which predicts a higher type number for satellite-framed languages, testifying higher *granularity* (as found by Stathi 2023). This difference depends on the coding system used for the data: this is true if we look at predicates (verb root). If we only consider verb roots, the satellite-framed specificity does not emerge in the types. In terms of general verbs frequency, Swedish has a low token frequency of the general verb *ha sönder* (expected for satellite-framed languages) but this is not true for the English general verb *break*.

In Swedish, the partile *sönder* emerged as representative of the break category, accounting for 21% of the construals, while its most frequent verb-particle predicate accounts for 16.5% of the construals. This data indicates that satellite-framed languages are, in fact, characterised by satellites, and a comparison that does not carefully take them into account could be misleading (e.g., only analysing verb roots as in Majid et al. 2007). In the domain of breaking events, *sönder* seems the relevant item for categorisation purposes: by encoding the results and the non-functionality (remember that the stative construction is rendered as *vara sönder* [Eng. 'to be in pieces'] for both material disruptions and non-functionality) it encodes the two basic frames of the breaking category.

Levels of abstraction

Our results indicate that there are different levels of abstraction based on the generality and specificity of verbs that are relevant to speakers according to language-specific lexicalisations and within-language conventions. There is a lack of a higher-level term superordinate to the general *break* that is linguistically relevant, and subordinate items vary across languages.

Regarding categorisation theory, we can go back to the initial question on the hierarchy of event categories. Concerning the subordinate level, in Italian, *rompere, spezzare* and *spaccare* have a close relationship. Each event in our dataset that is described with *spezzare* or *spaccare* is also described with *rompere* at least once. This would be the configuration of a *hyponymic relation.* The same relation is found, for example, between *break* and *snap,* but not in the Swedish verbs. According to this data set, specific break-type verbs such as *bryta, knäcka,* or *krossa* are not in a

subordination relation. The specific verbs can be posited to be in a troponymic relation, but not as strongly as with activity verbs. These verbs generally specify the change-of-state result and less so the manner. Thus, it seems more challenging to find troponyms for an abstract category, such as the break category, compared to EAT or RUN.

The variety of construals per action belonging to the same domain changes in languages and across languages: a wide variety of construals per action shows that there is less of a basic level preference. In contrast, a dominant verb per action indicates that there is a linguistically preferred level of abstraction. A case at the extremity is Swedish construals of 'crack an egg', where no diversity is recorded and only the verb *knäcka* (lit. 'crack') is used. In Swedish, there is a conventional description with no possibilities for perspectivization. Italian and English extend the categorisation of this event to other abstraction levels, using the general *break* and *rompere* verbs, as well as *open* and *aprire* ('open'). This suggests that in Swedish, the preferred level of abstraction is more specific than in Italian and English. Moreover, in Swedish, the particle *sönder,* encoding a primary component of breaking (the result in pieces), being efficiently combined with other roots, provides dynamicity to the Swedish breaking categorisation. Concerning this last point, the results of this data propose that the basic term for the break category in Swedish is an adjectival particle (*sönder,* 'in pieces') rather than a verb.

5. Conclusion

As categorisation theories continue to posit questions and challenges, this study aimed at discussing cross-linguistic categories and levels of abstractions in event categories with a case study on breaking events. Taking together the theoretical (Chapter 2) and empirical contributions on the topic, we showed that there is still ground to discuss the status of actions in the investigation of categorisation through form-meaning mappings. Two key points have guided the research: the cross-linguistic perspective, which helps to uncover differences in categorisation, and the focus on *general* verbs, which capture a category at a high level of abstraction, which has not been the focus of the research in event categories yet. The linguistically relevant verbs for the description of this domain are general verbs for English and Italian (*break* and *rompere*). In contrast, in Swedish, the verb at the highest level of abstraction, *ha sönder*, seems to be less extended to our set of events. These verbs describe a visually varied set of events brought together by their conceptualisation for functional similarity. The study of categorisation is operationalised through the quantitative and qualitative analysis of lexical diversity in the construals of a set of breaking events given by native speakers, which are analysed by quantifications, measures of

codability, and semantic maps. The cross-linguistic analysis showed the across-language differences in lexicalisations on a generality-specificity cline that establish different levels of abstraction in category formation at the basic and subordinate level. At the superordinate level, they appear to be missing relevant lexicalisations of an (even more) general level. Our findings are that the categorisation of events cannot be equated to that of discrete entities, as the levels of abstraction are not neatly divided into superordinate-basic-subordinate relations. This seems related to the nature of events as being relational, abstract, and encompassing more aspects of world knowledge than discrete entities or activities. The interconnection of categories at the linguistic level seems more relevant in event categorisation: events can be described with various construals pointing at different conceptualisations and evoking different categories. This variation appears conceptually driven but is limited by each language's conventions and lexicalisations. Each language has a preferred level of abstraction (or a conventionalisation) which picks on different aspects of the event – visual feature, function, or goals. The verbs used are similar cross-linguistically concerning the semantic dimensions they realise: the construals conceptualise the events in terms of result, manner of execution, tool of execution, force used, and space. The discussion above shows that categorisation mediates the interplay of conceptual structures and the operating level of construal. Categories seem flexible, and their members change based on their needs: an item can be an exemplar of one category in one context and another category in another context, as well as being a member of different categories for different speakers.

Chapter 5. Language Planning and Production

1. Introduction

Cross-linguistic variation has great impact in the speakers' conceptualisation of crucial aspects of events. It was discussed in Chapter 1-Section 3 the role of construals in directing the attention to specific aspects of events (selection of semantic roles, fore-grounding and back-grounding, generality, and specificity of actions) and it was shown in Study 2 that construals vary in terms of lexical variety cross-linguistically. We now investigate how speakers of different languages visually direct their attention to the unfolding of an event while producing descriptions in the form of a spoken sentence. We select Italian and Swedish (leaving out English) as they are the two more genetically different out of our three-languages sample and have shown more differences in their lexical variation and predicates.

The research questions in this field of research stem from the interest in linguistics, psychology, and cognitive sciences in finding out how we distribute the panning and producing phases of spoken language, that is, how we manage to move from a conceptualisation phase to the delivery of a message. The literature reviewed in Chapter 1 supported that language production rapidly affects how people allocate visual attention to components of a scene, in particular in the order in which the elements are attended. If people need to talk about what they see, they are likely to shift their focus of attention toward the aspects of the scene that are relevant for linguistic purposes. Language-driven effects on attention are found mainly in the domain of motion events, characterised by established cross-linguistic differences in verb typology (concerning the path of motion and the manner of motion). There has been evidence for a tight link between event conceptualisation and language-specific motion event descriptions in eye-tracking studies, which suggests an interconnectedness of gaze behaviour and language during production tasks. However, there is still a need for experimentation in other semantic domains. Chapter 1-Section 4.2.3 briefly presented models of message planning (lexical vs structural; Linear Incrementality vs Hierarchical Incrementality). From the literature on speech planning, we take the perspective adopted by Konopka and Meyer (2014) and Konopka (2019), which focuses on the planning of relational and non-relational information. The first is the information concerning individual participants (e.g., the discrete entities involved in the event), while the second is the information

concerning the action. We are concerned with relational information in this work, which is our focus of analysis.

We highlighted in the previous chapters the linguistic differences in the verbs and structures Italian and Swedish use to describe breaking events. Italian uses a general verb that encodes results and more specific verbs that convey different meanings about the type of results. Occasionally, manner verbs may be used to focus on the manner of action performance. Swedish expresses breaking in a predicate composed of a light verb and an adjectival particle (*sönder*, 'in pieces'), where the adjectival particle contributes the crucial semantics of the breaking category – the result in pieces. Other specific result verbs or manner verbs can substitute the light verb. In this last case, the predicate efficiently combines manner and result. We hypothesise that this difference is reflected in different sentence planning. To test this hypothesis, we turn to eye tracking to investigate the allocation of attention in two types of events: events showing a manner of execution without a change of the state of the object (Manner condition) and events showing the same action that leads to a breaking of the object (Result condition). The events are further divided according to the effector of the object and the type of patient in a coherent manner: in one condition, one-dimensional items are bent by the agent using their hands (Hand condition), and in another condition, the agent uses a tool on a three-dimensional object (Tool condition). We hypothesise that Swedish speakers are facilitated in the encoding of relational information, having the possibility to condense information about manner and result in a predicate.

2. Finding the action encoding window

As discussed in Chapter 1-Section 4.2.3, speech planning and production models investigate the formulation of a sentence in terms of how lexical and syntactic realisation takes place, as well as the type of incrementality in building the message (linear or hierarchical). These models predict how the information is structured into a sentence, whether precedence is given to lexical retrieval or syntactic structure, and linked to this aspect, whether the message is built by small units corresponding to the linear order of the sentence using a higher-level plan first (corresponding to the gist of a scene) and then building the increments. A challenging issue when using eye tracking to investigate Event Cognition, in particular actions, is how to isolate the encoding of the relational information. We discussed the approach taken by Konopka and Meyer (2014), which operated a distinction between relational and non-relational information as two aspects that are relevant in describing an event. This view on planning and production concerns this research as it predicts and focuses on the encoding of relational information, which is linguistically realised by

predicates in the languages and events here investigated. At what point during message and sentence formulation do speakers encode actions and verbs?

To do this, experiments can be designed to uncover the relational encoding processing. However, this is a challenging task, especially in eye-tracking. The encoding of non-relational information, as discrete entities in the stimuli, can be easily identified and analysed by looking at fixations to the areas of the event participants. Relational encoding is more complex to capture. It does not occur through fixations in a specific part of the stimuli but requires a conceptual interpretation of all the characters or objects in the stimuli. Additionally, considering parafoveal preview, an area does not need to be fixated to be processed if the relation is not ambiguous (Konopka 2019). The example proposed by Konopka (2019) is that of a kicking event: in 'John kicks the ball', the verb *kick* implies the use of a foot. Therefore, we could expect a kicking action to be performed with a foot and a speaker fixating the relevant area in a task of event description. However, other cues (body posture, object) may be sufficient to encode the action with a parafoveal view. Konopka and Meyer (2014) tested whether and when speakers prioritise encoding of different types of information (relational or non-relational) depending on the task. They compared the production of sentences of transitive events when the participant is primed with knowledge about the agent (e.g., asking 'what did the tiger do?' and instruct to answer starting with "it"). In this way, when the message planning begins, information about the agent has already been partially processed, and what is left to encode is relational information. Taking this research as a reference point, we use a strategy to prime participants prioritising relational encoding: participants see events performed by one and only agent. They are asked to describe them with an active sentence, always starting with the same subject name.

3. Research questions, design, and hypothesis

We ask whether cross-linguistic differences influence language production during event description and how they guide eye movements. The motivation lies in the lexicalisation and granularity differences between Swedish and Italian expressions of the same events. In particular, having shown how Swedish and Italian differ in terms of encoding of manner and result, we ask whether manner-focused and result-focused events are investigated differently depending on the language during verbalisation.

3.1 Selection of target events

As shown in Study 2 (Chapter 4), linguistic construals of the same event generally pick out different aspects of the events shown and evoke different categories. Thus, while expecting lexical variation in event descriptions, we chose the events described with verbs evoking the break category only (table 1). The events conform to the conditions of our interest for the investigation of breaking events:

- they are carried out with different action schemas (hitting with a tool vs. bending with hands)
- they can be manipulated for manner and result components.

These events highlight the cross-linguistic differences and similarities between the target languages: Swedish and Italian differ most in the integration of manner and result in actions described respectively by *ha sönder* and *rompere* ('break box with a hammer') but not in those described by *spezzare* and *snap* ('snap a pencil'), as highlighted by the categorisation map in Study 2 (Chapter 4-Section 3.3).

	Event	Italian	Swedish
1	**Break box with hammer**	Distruggere, rompere, spaccare, colpire, sfasciare, prendere a martellate, martellare	hamra sönder, slå hål, demolera, slå sönder, banka, banka sönder, slå ur, förstöra
2	**Snap pencil**	Spezzare, rompere	knäcka, bryta, bryta av, bryta sönder

Table 1. Construals of events in Swedish and Italian as resulted in Study 2

As shown in Table 1, event one is expressed in Italian with verbs encoding a result or the manner of the action, leaving the result unspecified. Swedish combines the two components in a predicate, occasionally choosing between the two (*banka* 'bang, pound' vs *förstöra* 'destroy'). This lexical diversity can be categorised as follows:

- Break verb root: *distruggere* ('destroy'), *rompere* (Eng. 'break'), *spaccare* ('smash), *sfasciare* ('break, smash');
- Hit verb root: *colpire* ('hit'), *prendere a martellate* ('hit with hammer'), *martellare* ('hammer');

- Hit + Break predicates: *hamra sönder* ('hammer in pieces'), *slå hål* (Eng. 'hit producing a hole'), *banka sönder* ('bang in pieces')

Event 2 is expressed in Italian by result verbs at various levels of specificity (*rompere* > *spezzare*), and so is Swedish (*Bryta* > *knäcka*). In both cases, all verbs' meanings belong to the breaking category.

3.2 Experiment Design

To explore these research questions, we use a within-subject design with 2-level variables and four conditions: Manner with tool, Result with tool, Manner with hands, and Result with hands. Figure 1 shows the design with screenshots of the video stimuli.

Figure 1. Conditions in the experiment design

3.3 Hypotheses

Our hypotheses concern lexical variety, onset times and eye movements.

From the literature review, we suggested that manner and result are two components of breaking events that are represented in construals. With the languages under investigation here, it is the case that verbs encode them differently. For events represented through videos, where an agent and objects move on the screen, we hypothesise that event semantics (whether the same performance of action ends with a result or not) influences the visual inspection. We have also shown with literature review and throughout Study 2 that actions performed with hands on one-dimensional objects are closely connected to two verbs in our sample languages: *spezzare* for

Italian and *bryta* for Swedish. We have seen that tools were predominant in Majid et al. (2007) and that verbs can lexicalise instruments. This offers construal possibilities to participants in terms of lexical variety and possibilities of generality-specificity. Speakers can choose whether to describe the action with a general result verb or use a lexicalisation of the instrument to define a manner of execution of the action. In both languages, this either-or situation is available. In Swedish, speakers can also choose to efficiently use a predicate with the lexicalised tool as the verb and the adjectival particle *sönder* encoding the result (e.g., *hamra sönder*, 'hammer in pieces'). Thus, we expect a cross-linguistic difference concerning the performance of the action condition.

3.3.1 Hypothesis on Hand condition

Combining these two conditions and obtaining the four conditions mentioned above, we expect the Manner-hand and Result-hand conditions to be similar in Italian and Swedish, as the linguistic encoding does not have significant differences concerning the verb. We expect hands not to attract attention for two reasons. First, they are not lexicalised in the verbs. Second, the event is not ambiguous, and the affordance is straightforward: a bending action of a thin, one-dimensional object that results in breakage is a frequent affordance (attested by corpus research in Study 1). Thus, we expect hands not to be frequently mentioned in prepositional phrases. Even if acknowledging them is necessary, given the non-ambiguity of the event, we hypothesise parafoveal encoding to be enough (as explained in Chapter 1-Section 4.2.3).

3.3.2 Hypothesis on Tool condition

We expect the main differences to be in the Tool conditions. We hypothesise a *facilitation of encoding* for Swedish. The *facilitation of encoding* hypothesis predicts that Swedish speakers can afford to plan the message more linearly. Encoding the verb should be more linear because it can be planned early, irrespective of the ending of the event. In the tool conditions (e.g., agent hitting a box with a hammer), the speaker can plan the verb *hit* or *hammer* to then encode the result with the particle *sönder* if the action performed leads to a result. We hypothesise this to be visible in earlier verb onsets, which would testify that lexical planning and retrieval are more immediate for Swedish speakers.

3.3.3 Speech planning and fixations

Predictions on the allocations of fixations according to speech planning theories can only be tentative in this experiment. In this regard, the study should be treated as exploratory: as the body of literature discussing speech planning and incrementality is based on picture viewing, our data only partially matches the data incrementality models usually explore.

A fundamental difference between picture-based and video-based experiments for event descriptions is that the action does not have a climatic point in picture-based studies. Speakers can start planning as soon as they see the picture. With video stimuli, speakers inspect before the video reaches a climax of informativeness: it is the moment in which the semantics of the event becomes clear, and speakers can make a final decision on the construal. Thus, the planning of sentences during video inspection deviates from the predictions of incrementality hypotheses. When viewing the event, speakers go through the pre-message and sentence-level planning with incremental updating during the unfolding of the event. These phases also stretch for longer before reaching production compared to picture viewing: event apprehension studies (Chapter 1-Section 4.1.2) show that apprehending the gist of an event for description is a rapid process, much more rapid than the natural unfolding of an action. In the specific case of the Tool conditions used in this work, both Italian and Swedish speakers, upon seeing a hitting event, are planning manner verbs and, possibly, inferring a result of the actions. For Italian speakers, this means planning a verb such as *colpire* ('hit') and a verb such as *rompere* (Eng. 'break'), which, in terms of structural planning, would result in a set of possibilities:

- Maria colpisce/martella la scatola
  Maria hits/hammers the box
- Maria rompe la scatola
  Maria breaks the box
- Maria colpisce/martella la scatola, rompendola
  Maria hits/hammers the box and breaks it
- Maria rompe la scatola colpendola/con un martello
  Maria breaks the box by hitting it / with a hammer

In Swedish, the planning of the Tool condition has different possible outcomes:

- Maria slår/hamrar burken
  Maria hits/hammers the vase
- Maria ha sönder burken
  Maria breaks the vase
- Maria slår/hamrar sönder burken
  Maria hits/hammers the vase broken

Italian speakers need to update both lexical choices and structure depending on the outcome of the event (the box breaks or does not break). Swedish speakers can potentially prepare a structure

and update only the lexical items. Given the Swedish possibility of manner and result encoding in a predicate, Swedish speakers can decide on lexical choices more efficiently at the event's climax. Italian participants might stretch the planning longer after the action climax.

The areas relevant for relational encoding are the tool and the object. The motion that is associated with the manner of action is in the tool, while the result is in the object. All areas are distinct but occasionally close to each other: the moment the hammer hits the object, the two areas merge. Thus, we need to keep in mind the parafoveal processing that is occurring. Participants can fixate on the object or the tool; either way, they are also aware of the other. Therefore, we take that if one of the areas is fixated more than the other, it means that the component is particularly relevant. Tentatively, we can hypothesise that with more linear planning of Swedish and less updating required for structural planning, Swedish speakers allocate more sustained looks to the object and process parafoveally the other relational areas.

The hypotheses outlined above are summarized below:

1. we expect hands not to be frequently mentioned and not a salient area drawing attention;
2. we expect a *facilitation of encoding* for Swedish speakers, who can afford to plan the message more linearly (earlier in the trial) in the tool condition, irrespective of the ending of the event, thus producing earlier onsets;
3. we expect Swedish speakers to allocate more sustained looks to the object and process parafoveally the other relational areas.

4. Building the experiment

4.1 Stimuli

The stimuli consist of short videos (5 to 12 seconds) with audio created for the purpose of the experiment. Creating the videos was judged as necessary in this type of study to guarantee that the target items were comparable and to control all variables except for the variable of study. Thus, the videos feature the same agent, background, lightning and zooming. The background colour is black, and the agent is wearing an outfit with neutral colours to avoid possible distractors. The scenes are shot in the recording studio[16] at the Humanities Lab of Lund's

[16] LARM Studio, Swedish acronym for Ljud, Animering, Rörlig bild, Musik (English: Sound, Animation, Video, Music). Website: https://www.humlab.lu.se/facilities/larm-studio/

University with the help of the lab technician[17]. The stimuli are thus recorded in the real world with a human agent and authentic props. The total number of stimuli is 44, of which 2 are practice videos and 8 are target videos. The target video features the agent in the same position (seated at a table). The fillers are balanced between full-length shots (agent only), half-length shots (agent only) and half-length shots with the agent seated at a table.

4.1. 1 Why natural scenes

Eye-tracking allows for both explorative and experimental studies, depending on the research's aim to favour ecological validity or conclusive results on a specific hypothesis. If explorative studies provide more ecological validity, tracking the interaction in a context where it usually occurs, they do not give definite conclusions about the relation between causes and effects of the viewing behaviour (Holmqvist et al. 2011). Hypothesis testing provides more conclusive results, testing theory-driven hypotheses with material designed for the purpose, controlling for one aspect at a time (Holsanova et al. 2009). An alternative is to create a balance between ecological validity and experimental control (Holsanova et al. 2009). In this experiment, we chose this last option. The close-to-natural scenes described above were selected to balance ecological validity and experimental control. Being in a recording studio, it was possible to create coherent stimuli but at the expense of action videos included in real-life situations (e.g., a home, a kitchen, a park…). In principle, these choices allow to have controlled results that can be generalised to natural situations of use (Holsanova 2014). Moreover, there is evidence of differences in eye-tracking research between naturalistic and animation stimuli (de Almeida et al. 2019 for a discussion). Our choice of naturalistic stimuli is to aim at ecological validity.

4.1.2 Characteristics of the target events

Table 2 describes the target items concerning the parameters of breaking events discussed in the literature review in Chapter 2 and Studies 1 and 2. It features the expected verbal descriptions regarding manner and result. This prediction is based on the semantics of the item and the linguistic characteristics of the languages.

[17] We thank Peter Roslund, research engineer at the LARM Studio, for his help and patience in shooting the videos.

Condition	Action	Effector	Obj. dim.	Type result	Expected linguistic encoding: Italian	Expected linguistic encoding: Swedish
Result tool	Repeated blows	Hammer	3D	In pieces	Result, periphrasis	Result or Manner.result
Result hands	Single bending	Hands	1D	In two parts	Result	Result
Manner tool	Repeated blows	Hammer	3D	No change	Manner or Attempted result	Manner or Attempted result
Manner hands	Repeated bending	Hands	1D	No change	Manner or Attempted result	Manner or Attempted result

Table 2. Description of the stimuli

4.1.3 Event timings

The videos were shot in a natural way, with actions performed by an agent. As a result, each item has a slightly different timing concerning the event's unfolding (Table 3). As noted in section 6.1, this is a shortcoming in the design, especially concerning the time of first bending or hitting actions. This shortcoming is considered in the analysis. The end time difference does not have a negative impact, as the analysis only considers earlier windows of event unfolding.

Table 4 describes the timings of the stimuli. The scenes are trimmed at the video start[18] so that the hands are in a resting position in the first 400ms of the video, the event apprehension window. After that, the first action begins: grabbing the objects (in the Hand condition) or the tool (in the Manner condition). The video's crucial point is the first bending action (in the Hand condition) or the first hitting action (in the Manner condition). At this point, the break will be immediately visible in the Result condition. In the Manner condition, it takes a few bending actions or hitting actions to show that the object is not affected, and its original state is not disrupted.

[18] Adobe Premiere Pro is used for all video editing, which consisting in segmenting the videos in single items and adjust the starting and ending of the video.

item	condition	grab tool	grab of object	first bending or hitting action	end action
break box	Result tool	500	600	1567	8500
smash plate	Result tool	450	430	3268	6300
snap branch	Result hands	-	1050	2480	5500
snap pencil	Result hands	-	1034	2767	5000
try break box	Manner tool	850	800	2934	8700
try smash plate	Manner tool	1500	550	2650	12400
try snap branch	Manner hands	-	1300	2300	11000
try snap pencil	Manner hands	-	918	2900	11000

Table 3. Event timing per item

Timing	Description
0-400	the actor in the video is still. The object and effector are visible as separate areas. Participants make sense of the scene through the event participants' affordances (e.g., the hammer close to a box affords a hitting motion).
First grab	The agent starts moving their hand and grasping the object (in the tool condition, also grabs the hammer).
First hitting or bending action.	The agent first hits or bends the object. After the first hit or bend, the object is visibly broken in the Result condition. In the Manner condition, the video continues with a few attempts at hitting or bending. At this point in the unfolding of the action, we hypothesise that the participant can plan the verb at the lexical level—choosing the preferred item.
End of action	End of action encoding window. The agent rests their hand on the table, separately from the object and the tool. The video stops, and the last frame remains visible to participants.

Table 4. Description and meaning of event timings.

4.1.4 Areas of Interest (AoI)

Areas of interest are managed in Data Viewer, the proprietary software provided with Eye Link trackers. They are drawn with circles of equal size in all items. They are dynamic—that is, they move in the video with the areas they encircle (shown in Figure 2).

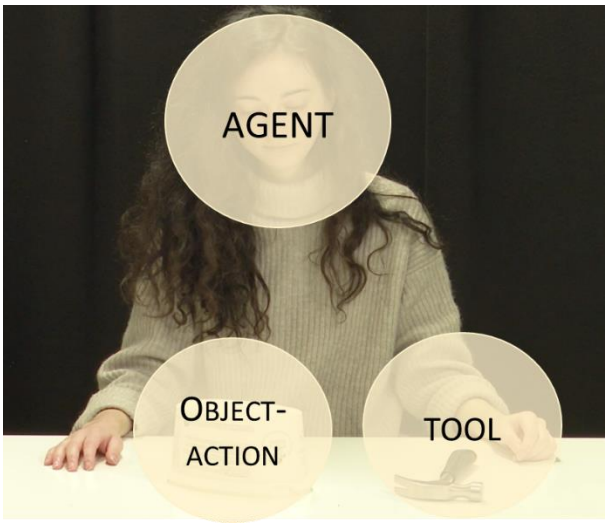


Figure 2. Example of Areas of Interest (AoI).

4.2 Data collection

Data collection took place at the University of Florence and the Humanities Lab of Lund University. The eye trackers used are the Eye-Link Portable Duo in Florence and the EyeLink 1000 in Lund. The machines feature the same configuration and use the same algorithms to process data (as stated by the manufacturer). The only difference is that one is a desktop version (has a desktop computer as the operating computer for the eye tracking software), and the other has a portable PC to operate the machine.

In Majid et al. (2007) consultants were shown the video clips, and they described the event after having seen each clip, with further linguistic probing for alternative descriptions carried out at the discretion of the researcher. Thus, each participant contributed, possibly, with more than one verb. This same procedure was used for the Imagact resource. Here, as in Study 2, we only asked participants to produce one description without further elicitation and without intervening in the production process. Differently from Study 2, we asked for spoken descriptions to prevent re-elaboration and recorded them.

4.3.1 Participants

Forty-four Italian native speakers and forty-three Swedish native speakers participated in the experiment, of which 37 per language could be used for the analysis (see section 5.2). Italian participants were recruited as volunteers, and Swedish participants received lunch vouchers. Participants signed an informed consent and answered a three-question language questionnaire to assess their native speaker status. Participants declared (i) having acquired the target language

from infancy in a natural way and not via explicit instruction, (ii) having spoken only the target language in domestic contexts, and (iii) at present, using the target language as the main language of communication and study.

4.3.2 Instructions

The instructions for both languages are reported in Appendix F. These instructions result from a series of pilots with Swedish and Italian participants, carried out before the final data collection. One of the challenges was to find appropriate wording for instructions. Our aim was for the design to have natural responses, and it was also important to have descriptions while the video unfolded. Early verb onset would be the most informative part of planning processes, and eye movements would be cleaner. When piloting, we found that participants were uncomfortable when asked to describe the sentence as quickly as possible and produced inaccurate responses (e.g., in the manner scenes, participants would start describing too soon with a resultative sentence, inferring the end of the event). At the same time, not indicating when to begin describing resulted in participants taking longer to produce the sentence. The final instructions asked participants to describe "as soon as you understand what is happening in the video".

4.3 Protocol

The protocol of the experiment followed the standard steps and recommendation of eye tracking research (Conklin et al. 2018; Attardo and Pickering 2023):

1. Welcome participants and read the informed consent.
2. General instructions.
   Participants were told what to expect (e.g., 'You are going to describe some short videos') and explained how the eye-tracking machine worked.
3. Calibration: the participant sits on the dedicated position, and calibration starts
4. Start of the experiment
    a. Instructions on screen
    b. Practice trials: participants familiarise themselves with the task, stimuli, and agent in the video. When participants first see experimental items, they have inspected the agent and setting. At this stage, participants can ask questions about their

descriptions, and the experimenter can assess whether the task is straightforward for them.

c. Recording of the experiment: participants start describing the items on the screen in a randomised sequence. There is no interaction between the participant and the experimenter.

5. Debriefing
   Participants are allowed to ask questions about the aim of the experiment and about their performance.

4.4 Measures and windows of analysis

We are interested in the following data to investigate our research question:

- Speech onsets
- Construals
- Eye movements
  - Fixations to Areas of Interest (AoI)
  - Time course of fixations

The trials unfolding through time are segmented into windows for analysis. Each action starts with the agent in a seated position with hands resting on the table. After 400ms, the hands start moving in the direction of the object and the tool. The first action performed by the agent in the video is the grasping the object in the Hand condition and the grasping of the hammer in the Tool condition. At this point, participants have had enough time to look at the agent, objects and tool to make predictions on the relationship between these items based on their affordances. The crucial moment for action planning is the first action performed on the object (first hitting or bending action). In the Result condition, the first hit of the hammer on the object or the first bend of the object shows the result in pieces. After that, the event ends. In the Manner condition, at first hit or bend, it appears that the action does not change the physical shape of the object. The videos show a series of repeated hits or bending motions, after which the event ends. The first hitting or bending action time through the trials varies within the items (Figure 3). Speakers can start speaking at their will. We expect verb onsets to start after the first hitting or bending action.

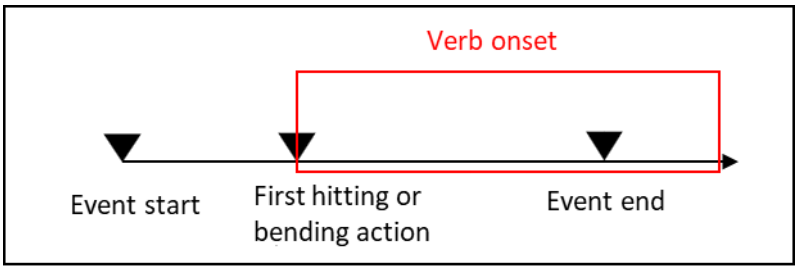


Figure 3. Unfolding of the events with salient actions

The time-windows of interest are three:

- Early window: from the event start to the climax of the event. The climax corresponds to the first hitting or bending action. This window includes the initial message planning, which is crucial for the encoding of relational information. The linguistic differences should be visible here as explained in section 3.3.3: Swedish speakers can plan the structure (and possibly a verb) for the hitting or bending action and adjust their construal with a result particle if the event ends in a result or continue with the initial planning. Having each item different timings of first hitting or bending action, the windows for each item differ in length.
- Late window: From the first hitting or bending action to verb onset. This window captures the revealing moment of the semantics of the event and the production phase. Verb onsets can occur at different times in the window, and thus, trials in this window have different lengths.
- Middle window: 1000ms before the first hitting or bending action and 1000ms after the first hitting or bending action. This window captures the fixations before and after the event's crucial point. Compared to the other two windows, it captures the same length for each trial.

5. Results

5.1 Data coding

Audio files were transcribed and segmented in Praat (https://www.fon.hum.uva.nl/praat/). Each word's beginning and ending were segmented. The Italian transcription was done manually, and the Swedish transcription was processed with an automatic speech recognition software, Whisper[19], and checked manually. An automatic aligner for word segmentation was used and it was then checked manually. The eye-tracking files were managed in Data Viewer to position the Areas of Interest, check the resulting files for errors and download fixation reports. Further data coding (merging the files, processing, and analysis) was done in RStudio.

5.2 Data cleaning

The total number of participants for the Italian dataset was 44 and for Swedish, 43. Each participant's data files were evaluated to remove material errors – such as occasional technical issues and calibration errors for eye-tracking, or failed audio recordings. Participants who had to be discarded were 7 for Italian and 6 for Swedish, with the following motivations:

- Italian:
  - Eye tracking: participants 2, 9 and 37 were discarded for missing eye-tracking files; participants 4 and 10 were removed for calibration errors;
  - Audio files: participant 38 was discarded for missing audio file.

- Swedish:
  - Eye-tracking: participants 1 and 2 were discarded for missing eye-tracking file; participant 11 left the session before finishing stimuli; participant 41 was removed for calibration errors;
  - Audio files: participants 12 and 34 were discarded for missing audio files.

[19] With the help of Lund's Humanities Lab in the person of Johan Frid, which edited the Pratt script to include the automatic Whisper annotation.

Furthermore, a few individual trials had to be removed: 4 for Italian (items: 1 break box, 1 snap branch, 1 try smash plate, 1 snap pencil) and 3 for Swedish (items: 1 smash plate, 2 break box) because of a single missing file audio or missing eye-tracking data within a participant data.

Each trial was evaluated for accuracy and conformance to the instructions (whether the sentences followed the required syntactic structure and interpreted the video according to the target semantic meaning). As participants did not have instructions on how to interpret the scene, accuracy was only evaluated in the case of manner scenes. If the participants used a result verb in a manner video, it was discarded – e.g., 'John breaks the box' for the manner-tool video, where the agent do not break the box. Sentences construed on an attempted result (e.g., 'John tried to break the box') were kept. The recordings report was annotated with a scoring for sentences to keep (marked as 'YES') and to discard ('NO'). The inaccurate sentences were 6 in the Italian dataset and 11 in the Swedish dataset. Other discarded sentences were 10 for Italian and 1 for Swedish, motivated below:

- Italian:
  - participant 1 construed a Manner-hand scene with a negation (*Maria non spezza il ramo,* 'Maria does not snap the branch')
  - participants 11, 14, and 17 used the verb *usare* ('use')
  - participants 14 and 42 (x2) started to construe the sentence as *Maria vuole* ('Maria wants to')
  - participant 30 used the verb *vedere* ('see') in a result-tool scene and the verb *prendere* ('grab*')* in a Manner-hand scene.

- Swedish
  - participant 10, Manner-tool condition, verb *nudda*, 'touch'*).*

After data cleaning, 37 participants per language were used for the analysis and a total of 276 trials for Italian and 281 for Swedish, out of 296 possible trials.

5.3 Linguistic production and verb onsets

The first step in analysing the results considers the types of construals, the lexical diversity, and the time of onset.

6.3.1 Construals and lexical choices

As per instructions, speakers produced active sentences in the SVO structure. The types of construals concerning Manner-result videos were five:

1. Result of action with a result video: congruent description (e.g., 'Mary broke the vase')
2. Result of action with a manner video: non-congruent descriptions (e.g., 'Mary hit the vase'). These construals were excluded, as mentioned in the section above.
3. Manner of the action with a manner video: congruent description (e.g., 'Mary hit the vase')
4. Manner of the action with a result video: underspecified for result but valid response. Unlike the construal 2., it cannot be denied that a manner of performance occurred. This can be expressed with a contact or manner verb (e.g., *hit, hammer*)
5. Attempted result: in manner videos, a possible construal could be the interpretation of the action, instead of describing the performance of the action, participants described what they thought was the intention of the action (e.g., 'Maria tries to break the box'). This construal recurred equally in both Italian and Swedish.

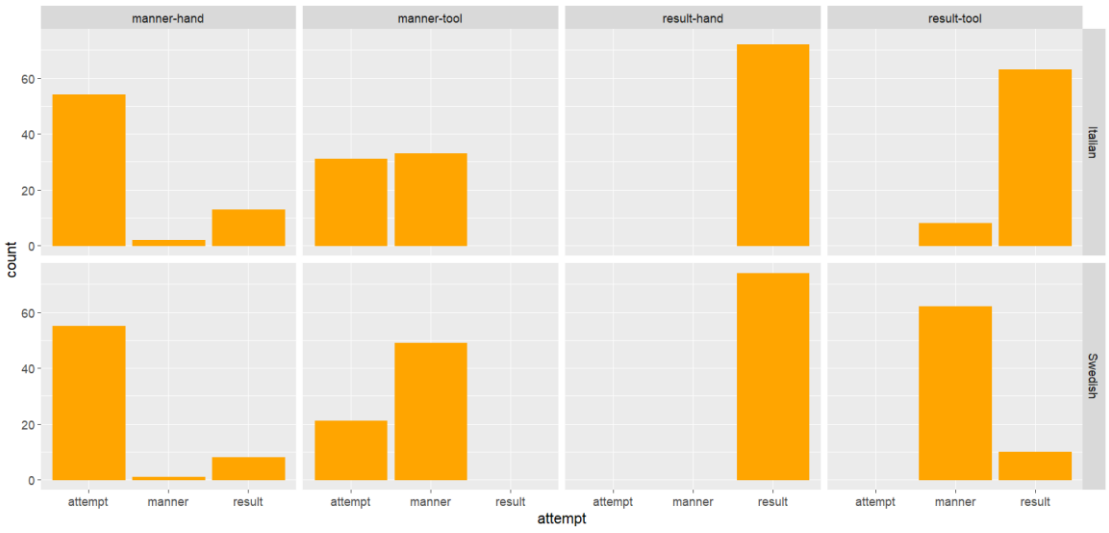


Figure 4. Distribution of construal types per language and condition

Figure 4 shows that in the Hand conditions, Italian and Swedish speakers used the types of construals similarly. In the Manner-Hand condition, the 'attempted' result construal is the most frequent, while participants only used result construals in the Result-hand condition. In the Tool condition, differences emerged. In the Manner-tool condition, Swedish speakers preferred the manner construal, while Italian speakers used equally the manner construal and the attempted result. In the Result-tool condition, it is evident that Italian and Swedish speakers chose opposite

construals. Swedish speakers preferred to include manner, while Italian speakers focused on the result. Overall, Swedish speakers chose to use manner verbs more often than Italian speakers. Figure 5 shows the detail of verbs used in the construal, according to a three-way distinction: verbs encoding a manner of the action (e.g., *hammer*), a result (e.g., *break*) and both components (which we coded *manner.result* in our files). This last possibility is only possible in Swedish, and it includes predicates such as *hamra sönder* ('hammer in pieces').

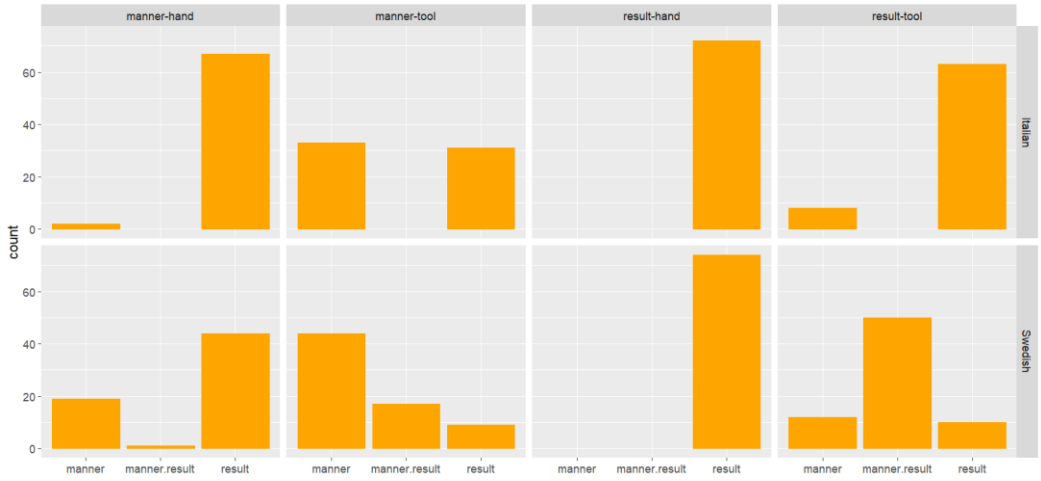


Figure 5. Distribution of verbs per language and condition

In the Manner-hand condition, Swedish has more manner verbs. There are no differences in the Result-hand condition. In the tool conditions, Swedish has manner.result as the most typical configuration for the result videos and more manner verbs in the manner video. In the tool condition Italian has result verbs and manner verbs equally distributed in the manner videos and predominantly result verbs in the result video.

The chi-square tests of the two variables are independent. Setting our null hypothesis as 'language and produced verb type are independent', and our hypothesis as 'language and produced verb type are dependent', we see that the two variables are dependent (performed without Yates' continuity correction: X-squared = 77,28, df = 1, p-value < 0). Chi-square tests were performed on individual conditions (except for Result-hand condition, which has the same outcome for Swedish and Italian) and showed significant differences. Overall, Italian speakers prefer result verbs, and Swedish speakers prefer manner verbs, having the possibility to include the result in the particle *sönder.* There is a difference between manner-hand and manner-tool scenes in that hand scenes were described more consistently with result verbs (e.g., in the attempted result construal) than tool scenes.

Focusing on the performance of the action, *hand* and *tool* distinguished actions performed manually and actions performed with a hammer used for repeated hitting movements on the objects. Hands were mentioned only 3 times by Italian speakers and never by Swedish speakers. The tool was explicitly mentioned 34 times by Swedish participants and 47 by Italian participants with a prepositional phrase (Swe. *med hammare*; Ita. *con un martello* [both 'with a hammer']). For the Tool condition, the information about the action being performed with a hammer could also be introduced in the verb, with the lexicalisations *hamra* for Swedish and *martellare* for Italian (or the expression *dare una martellata,* 'give a hammer blow'). This occurred in 18 construals for Italian and 34 for Swedish. In total, the tool was mentioned 65 times by Italian participants and 68 times by Swedish participants (Table 5).

	Verb 'hammer'		**Phrase 'with a hammer'**			
	Raw	%	Raw	%	Total raw	Total %
Italian	18	6.5	47	17	65	23.5
Swedish	34	12	34	12	68	24

Table 5. Counts of occurrence of tool encoding in the verb or a prepositional phrase.

5.3.2 Verb onsets

Verb onsets are investigated across languages and conditions. The trend we can assess visually is that Swedish participants have, on average, earlier onsets than Italian participants (Figure 6). The trend is also confirmed by comparing manner and result, and language (Figure 7).

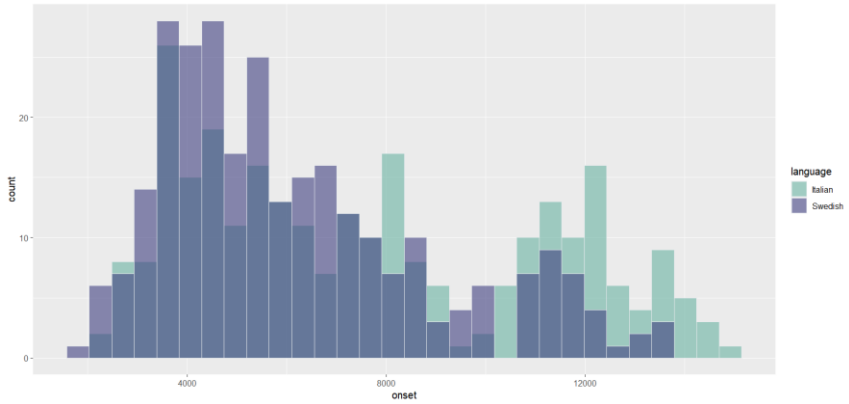


Figure 6. Verb onset times in Swedish participants and Italian participants

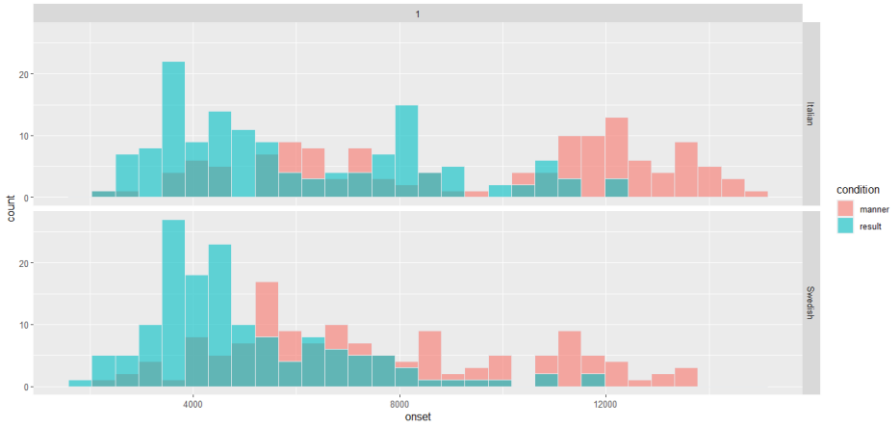


Figure 7. Verb onset times in Manner and Result conditions

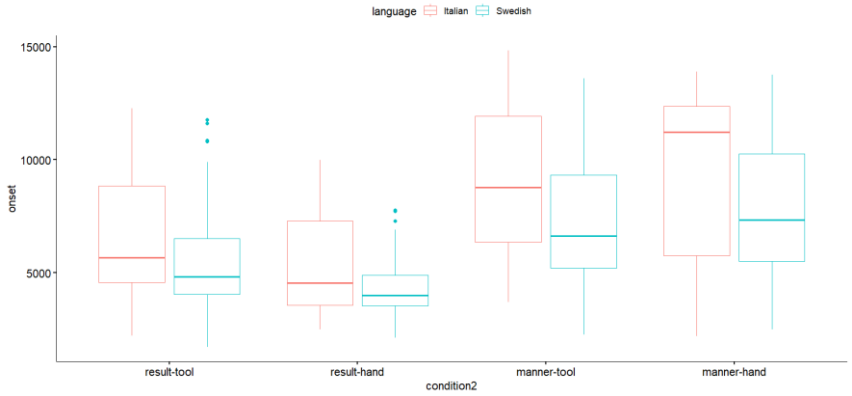


Figure 8. Boxplot of onsets for Swedish and Italian participants compared in each condition.

A one-way ANOVA was performed to compare the effect of language on verb onsets. Results show a significant difference (Mean Sq. 290680449; F-value 290680449; P-value < 0). The ANOVA test was repeated to check the effect of the four conditions on the onsets, and the results confirmed a significant difference in the groups (Mean Sq. 1326909470, F-value: 56.28; P-value < 0). The post-hoc test (Tukey Multiple comparisons of means) shows a difference in the means for all pairwise comparisons except for Manner-tool and Manner-hand conditions (Figure 8, Table 6). This analysis compares the conditions without distinguishing between languages. Thus, the next step is to check the interaction between the language and the conditions. The significant differences in the onsets and conditions, per language, are in the manner videos.

	Difference	Upper	Lower	P-value adj
Swedish:result-tool-Italian:result-tool	-1318.9110	-2699.26007	61.43808	0.0730
Swedish:result-hand-Italian:result-hand	-809.2661	-2175.45472	556.9224	0.618
Swedish:manner-tool-Italian:manner-tool	-1983.0095	-3416.26252	-549.7565	0 ***
Swedish:manner-hand-Italian:manner-hand	-1705.3843	-3137.66327	-273.1053	0.007 **

Table 6. Report of Anova tests on the differences between languages and conditions

In the Wilcoxon rank sum test (without continuity correction), the difference between Italian and Swedish with respect to the time of onset is significant (Table 7). Performing the test in each individual condition, the group difference emerges in all conditions except in the Result-hand condition. Moreover, time onset and produced verb type (manner verb or result verb) are not dependent.

Wilcoxon rank sum test (without continuity correction)		
	p-value	
Language	0	***
Tool vs hand	0.015	**
Manner vs result	0	***
Produced verb type	0.165	
Result-tool	0.006	**
Result-hand	0.073	
Manner-tool	0	***
Manner-hand	0.001	**

Table 7. Wilcoxon rank sum test (without continuity correction) on pairwise variables

We now test onsets and verbs with a mixed model. As it has been found that time of onsets and language are dependent, we will now proceed to fit a mixed model.

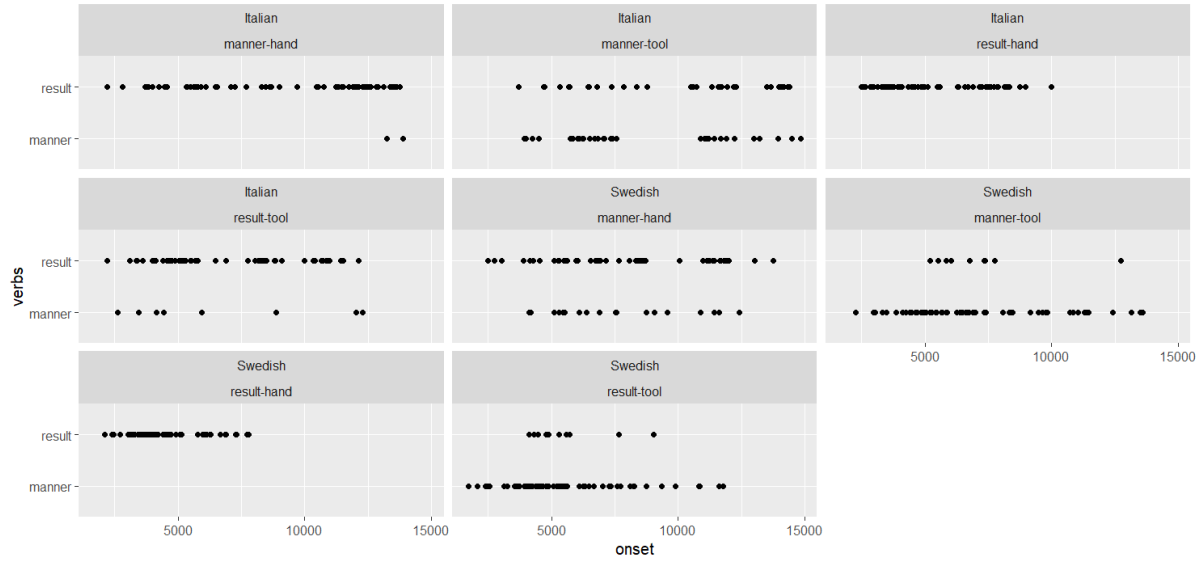


Figure 9. Distribution of time onsets per verb produced, conditions, and language.

We fitted a linear mixed model (estimated using REML and BOBYQA optimiser) to predict onset with language, the Manner-result condition and the Tool-hand condition, with participants and items as random effects. Results indicate that onset times are significantly lower for Swedish compared to Italian (beta = -1295.20, 95% CI [-1693.79, -896.62], t(547) = -6.38, p < .001; Std. beta = -0.41, 95% CI [-0.53, -0.28]). Comparing conditions across languages, we find that result scenes have shorter onsets (result-hand: beta = -3597.31, 95% CI [-4514.52, -2680.10], t(547) = -7.70, p < .001; Std. beta = -1.12, 95% CI [-1.41, -0.84]; result-tool: beta = -2387.55, 95% CI [-3306.37, -1468.74], t(547) = -5.10, p < .001; Std. beta = -0.75, 95% CI [-1.03, -0.46]).

Taken together, we found an overall language effect that was persistent within conditions. Overall, the Result condition has shorter times for onsets, in both Italian and Swedish. We did not find an effect for verb onsets and verb type.

5.4 Fixations and Multilevel Mixed Models (MLM) Analysis

To visualise our data, we plot the proportions of fixations to each interest area. A crucial part of the analysis is the time-windows analysis described in section 4.5.1. We visualise the proportions of fixations in boxplots and fit multilevel mixed models to test main effects and interactions in three windows of analysis: the early time window (from event start to first hitting or bending action, adjusted with the individual items' event timings), middle time window (a total of 2ms around the first hitting or bending action) and a late window (from first hitting or bending action to verb onset, adjusted to each participants' onset time).

Mixed models are statistical models combining fixed effects and random effects (Bates, 2010). The factors that are part of the design and have been considered so far are listed below:

- Onsets (numerical)
- Fixations (numerical)
- Produced verb type (2 levels, Manner and Result)

These variables are analysed according to the following groups:

- Language (2 levels, Swedish and Italian)
- Condition 1 (2 Levels: Result video and Manner video)
- Condition 2 (2 Levels: Tool and Hand)

We consider these factors as fixed effects to fit a mixed model: we are interested in the effects of language and conditions on the fixations. Participants and items are the random effects, representing participants' or items' deviations from the fixed-effect parameters (Bates, 2010). In the models, we look at each interest area separately, excluding hand fixations because we do not have enough data points. We start fitting the models in a bottom-up fashion – starting from a simple model, we proceed to adding factors. We begin with one-predictor models to check overall cross-linguistic differences and the effect of conditions. Then we check interactions between language and the two conditions. Next, we fit multilevel mixed models to see the main effects and interactions. We report the result for each time window.

5.4.1 Early time window

Here, participants look at the scenes but have yet to produce a verb. However, they have started speaking, producing the agent of the sentence (*Maria*).

The plot in Figure 10 shows that the most fixated area is that of the object, followed by the agent. The effectors received less fixations. In the case of the hand effectors, they appear irrelevant. Overall, the plots show a difference between languages in all conditions and a difference in manner and results within each language. Swedish participants seem to direct more looks to the object. In the Tool condition, Swedish speakers seem to look more to the tool in the Manner condition (0.14) than the Result condition (0.08). For Italian speakers, the difference is less prominent (manner: 0.12; result in 0.11).

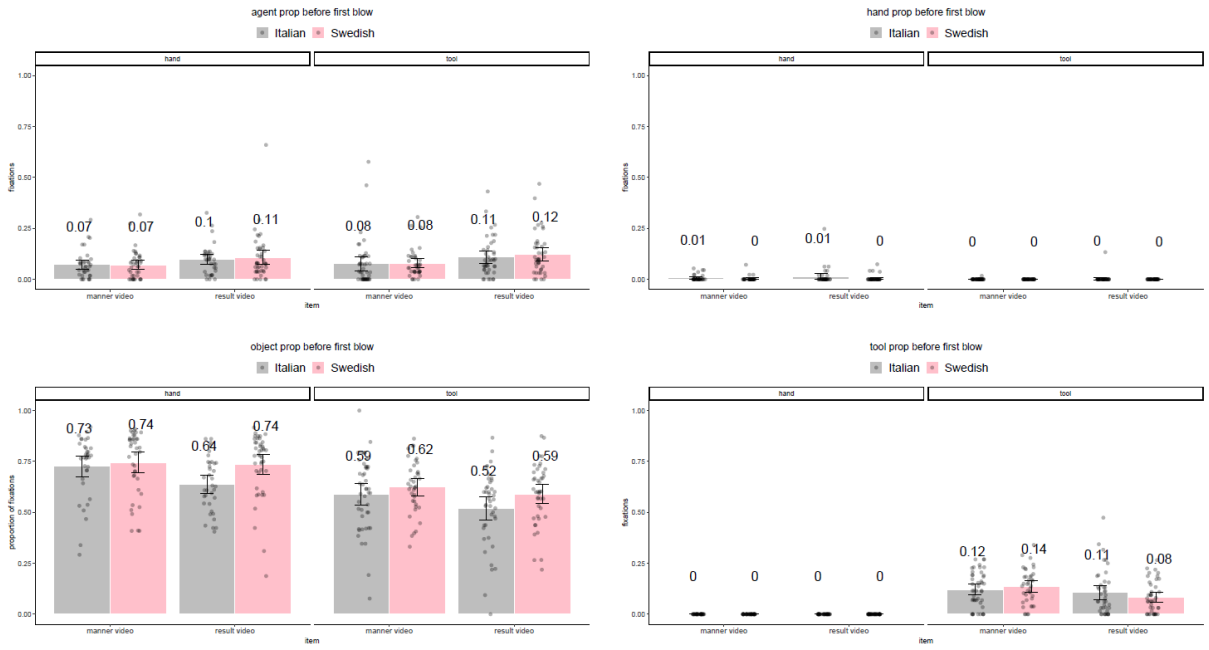


Figure 10. Boxplot of proportions of fixations in each AoI per language and condition in the early time window.

Main effects for Language - Fixed effects						
Object fixation	Estimate	Std. Error	df t	value	Pr(>|t|)	
(Intercept)	0.64439	0.03819	818.257	16.874 0	.00000012	***
language1	0.05743	0.02505	6.982.613	2.292	0.0249	*
Tool fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.058473	0.026411	7.093.242	2.214	0.0619	.
language1	-0.003256	0.008022	70.544.442	-0.406	0.6860	
Agent fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.092233	0.011528	12.782.424	8.001	0.00000249	***
language1	0.009423	0.014494	70.816.663	0.650	0.518	
Language with conditions - fixed effects						
Object fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.64443	0.03192	627.117	20.191	0.000000607	***
language1	0.05750	0.02507	6.980.863	2.294	0.0248	*
tool.hand1	-0.13139	0.06025	498.666	-2.181	0.0812	.
Manner.result1	-0.04797	0.06026	498.706	-0.796	0.4622	
Tool fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.058488	0.016482	5.167.657	3.549	0.0155	*
language1	-0.003286	0.008016	70.568.857	-0.410	0.6831	
tool.hand1	0.116937	0.032677	4.992.193	3.579	0.0159	*
Manner.result1	-0.019460	0.032677	4.992.258	-0.596	0.5775	
Agent fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.092105	0.010599	10.393.089	8.690	0.00000436	***
language1	0.009428	0.014493	70.804.145	0.651	0.517	
tool.hand1	0.006597	0.017494	4.976.987	0.377	0.722	
Manner.result1	0.033755	0.017496	4.979.923	1.929	0.112	

Table 8. Result of the multilevel mixed models for main effects in the early time window.

To interpret this data, we recall that both Swedish and Italian participants mention tools with equal frequency, but Swedish speakers encode them more often in the verb. This means that looks to the tool should be located in an earlier window compared to Italian speakers, as the verb is produced earlier in the sentence structure than a prepositional phrase ('Maria hammers a box' vs. 'Maria breaks a box with a hammer').

In Table 8 and Table 9, we report the results of MLM testing performed in each AoI. We find a main effect of language in object fixations: Swedish speakers dedicated more fixations to the object. We see an interaction between language and the Manner-result condition: Swedish speakers' fixations are directed more to the object compared to Italian speakers' fixations when it is a result video. The main effect for Tool-hand condition is almost significant (P 0.08). It becomes significant if we drop the items and random effects: this would tell that the proportions of fixations to the object are less when it is a tool video. There is an effect for tool fixations and Tool-hand condition. This result should not be considered, as in the Tool-hand condition, one level has the tool as the interest area, and the other level does not.

Interactions - 3 factors (Language, condition 1, condition 2)						
Object fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.6444916	0.0352641	48.058.519	18.276	0.0000126	***
language1	0.0561849	0.0250345	697.912.755	2.244	0.0280	*
tool.hand1	-0.1306506	0.0673187	39.919.417	-1.941	0.1244	
Manner.result1	-0.0482333	0.0673200	39.922.605	-0.716	0.5134	
language1:tool.hand1	-0.0008514	0.0271548	4.630.985.366	-0.031	0.9750	
language1: Manner.result1	0.0547420	0.0271681	4.636.001.845	2.015	0.0445	*
tool.hand1: Manner.result1	0.0086055	0.1346366	39.918.609	0.064	0.9521	
language1:tool.hand1: Manner.result1	-0.0099832	0.0543032	4.630.008.800	-0.184	0.8542	
Tool fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.058343	0.017770	4.121.273	3.283	0.0291	*
language1	-0.002889	0.008043	70.616.358	-0.359	0.7205	
tool.hand1	0.116675	0.035266	3.996.267	3.308	0.0297	*
Manner.result1	-0.019229	0.035266	3.996.368	-0.545	0.6146	
language1:tool.hand1	-0.006036	0.013451	467.403.825	-0.449	0.6538	
language1: Manner.result1	-0.022251	0.013453	468.212.079	-1.654	0.0988	.
tool.hand1: Manner.result1	-0.038203	0.070531	3.996.239	-0.542	0.6168	
language1:tool.hand1: Manner.result1	-0.044456	0.026900	467.201.144	-1.653	0.0991	.
Agent fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.092068	0.011414	7.445.951	8.066	0.0000616	***
language1	0.009250	0.014507	70.857.696	0.638	0.526	
tool.hand1	0.006645	0.019441	3.982.980	0.342	0.750	
Manner.result1	0.033768	0.019443	3.984.317	1.737	0.158	
language1:tool.hand1	0.001649	0.016405	464.309.509	0.101	0.920	
language1:Manner.result1	0.008295	0.016412	464.832.537	0.505	0.613	
tool.hand1: Manner.result1	0.009226	0.038881	3.982.646	0.237	0.824	
language1:tool.hand1: Manner.result1	-0.011498	0.032806	464.205.105	-0.350	0.726	

Table 9. Result of the multilevel mixed models for interactions in the early time window.

5.4.2 Late time window

In this time window, speakers know whether the event ends with a break or not. Again, the object is the area receiving more fixations, followed by the agent (Figure 11). Compared to the early window, the agent receives more looks, and the interest in the effector seems to drop (for both hands and tools). We find more looks to the tool for Italian speakers, in both manner and result tool videos compared to Swedish. Focusing on the conditions where language differences are more relevant, the plots show fewer looks at the objects in the Manner-tool condition than in the Result-tool condition, both for Italian and Swedish speakers, with Italian showing overall fewer fixations. Compared to the early time windows, there are more looks to the object in the late window, in the Tool conditions. The interest to the object in the Hand condition, instead, is stable from early to late window, while it increases in the Tool condition after the first blow. The properties of the videos themselves explain this: the action is more static in the Hand conditions. In contrast, in the Tool conditions there is dynamicity in the action going from hands to tool to the

object. It is relevant to see that in the dynamicity of the event and the areas to fixate, the interest is more stable to the object after the first action. We perform the same MLM in this time window.

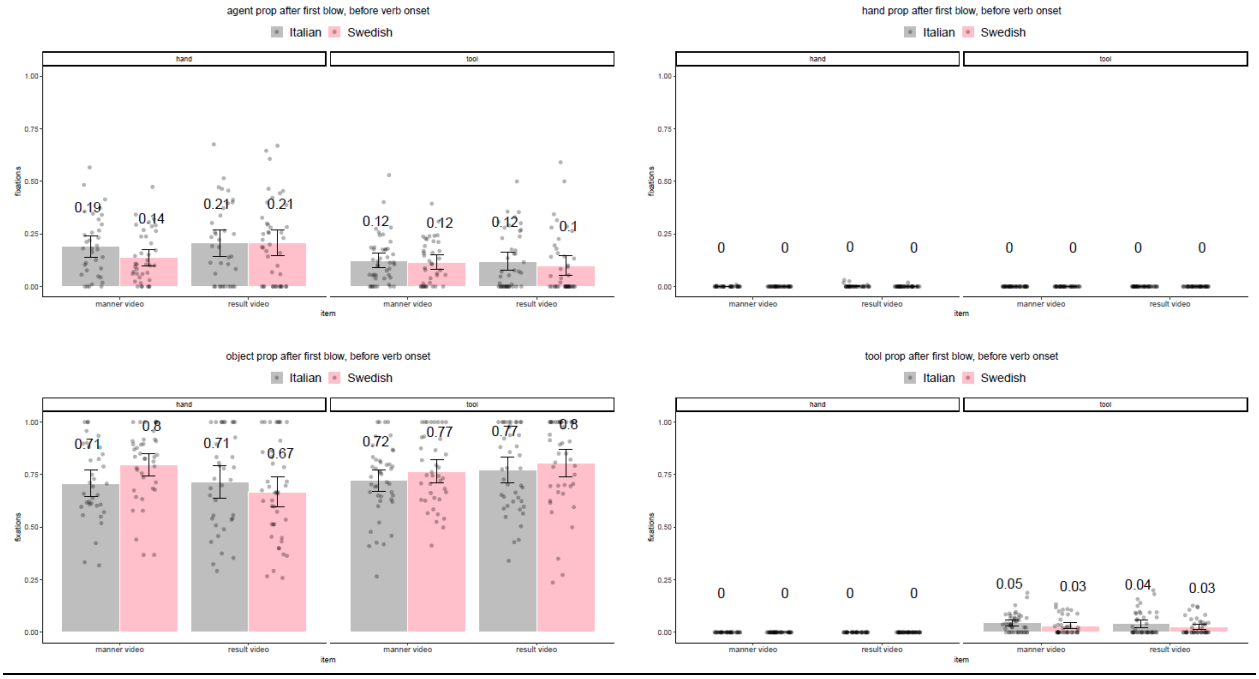


Figure 11. Boxplot of proportions of fixations in each AoI per language and condition in the late time window.

Main effect for language - Fixed effects						
Object fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.74201	0.03335	1046914	22.249	0.000	***
language1	0.02744	0.03319	7.045.802	0.827	0.411	
Tool fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.018049	0.007276	7.241.076	2.481	0.0411	*
language1	-0.005905	0.003979	69.639.445	-1.484	0.1423	
Agent fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.15141	0.02567	1.086.380	5.898	0.000109	***
language1	-0.01369	0.02599	7.082.139	-0.527	0.600047	
Main effects with Condition 1 and 2 - Fixed effects						
Object fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.740437	0.016427	70.901.995	45.074	<0.000	***
Manner.result1	-0.005608	0.018052	462.152.921	-0.311	0.7562	
tool.hand1	0.042497	0.018042	462.856.617	2.355	0.0189	*
Tool fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.018072	0.002762	6.509.389	6.542	0.000436	***
tool.hand1	0.036221	0.005176	5.157.740	6.997	0.000808	***
Manner.result1	-0.002683	0.005177	5.164.244	-0.518	0.625683	
Agent fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.15149	0.02454	808.132	6.173	0.000256	***
tool.hand1	-0.07056	0.04345	501.253	-1.624	0.165137	
Manner.result1	0.01844	0.04344	501.186	0.424	0.688874	

Table 10. Result of the multilevel mixed models for main effects in the late time window.

Interactions between language, condition 1 and condition 2 - Fixed effects						
Object fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.742255	0.038385	5.419.339	19.337	0.000	***
language1	0.029040	0.033230	70.576.423	0.874	0.3851	
Manner.result1	-0.006034	0.071151	4.009.723	-0.085	0.9365	
tool.hand1	0.039244	0.071152	4.009.997	0.552	0.6105	
language1: Manner.result1	-0.067154	0.033054	454.047.522	-2.032	0.0428	*
language1:tool.hand1	0.019103	0.033064	454.725.024	0.578	0.5637	
Manner.result1:tool.hand1	0.121043	0.142297	4.009.144	0.851	0.4428	
language1:manner.result1:tool.hand	0.140295	0.066063	453.550.533	2.124	0.0342	*
Tool fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.018150	0.002944	5.024.921	6.165	0.00161	**
language1	-0.005982	0.003967	69.813.551	-1.508	0.13614	
tool.hand1	0.036417	0.005588	4.140.769	6.517	0.00253	**
Manner.result1	-0.002893	0.005588	4.140.765	-0.518	0.63112	
language1:tool.hand1	-0.011349	0.007011	460.094.693	-1.619	0.10621	
language1: Manner.result1	0.002172	0.007011	459.960.912	0.310	0.75682	
tool.hand1: Manner.result1	-0.006004	0.011174	4.139.665	-0.537	0.61863	
language1:tool.hand1: Manner.result1	0.003978	0.014020	458.862.543	0.284	0.77673	
Agent fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	***
(Intercept)	0.151037	0.025753	6.238.920	5.865	0.000945	
language1	-0.013939	0.026042	70.837.443	-0.535	0.594155	
tool.hand1	-0.070307	0.046059	4.015.984	-1.526	0.201327	
Manner.result1	0.019008	0.046058	4.015.579	0.413	0.700914	
Language1:tool.hand1	0.003318	0.024232	454.381.371	0.137	0.891146	
language1: Manner.result1	0.013361	0.024223	453.695.597	0.552	0.581512	
tool.hand1: Manner.result1	-0.063244	0.092112	4.014.795	-0.687	0.529938	*
language1: tool.hand1: Manner.result1	-0.099809	0.048412	453.253.181	-2.062	0.039811	

Table 11. Result of the multilevel mixed models for interactions in the late time window.

In this window, we do not find a main effect of language. We find a significant and positive effect for Tool-hand condition in the agent fixations: the agent received more fixations in the Hand conditions. We find one interaction of language and condition and one of the two significant conditions: there are fewer fixations to the object in Swedish participants to the result videos and fewer fixations to the agent in result-tool videos than in result-hand videos. There is also a three-way interaction for Swedish, Result condition and Tool condition that is significant and positive, meaning that Swedish speakers allocate more fixations to the object in the Result-tool condition.

5.4.3 Middle time window

This time window captures fixations around the climax of the action (first hitting or bending action). We did not find main effects or interactions in the conditions of interest. However, the plot (Figure 12) shows a difference in the fixations to the agents, which are more in the Result-hand condition compared to all other conditions. The Result-hand condition is the easiest to plan

in both languages for the immediateness of the semantic content. For this reason, it might free the participant's attention from the relation-encoding areas.

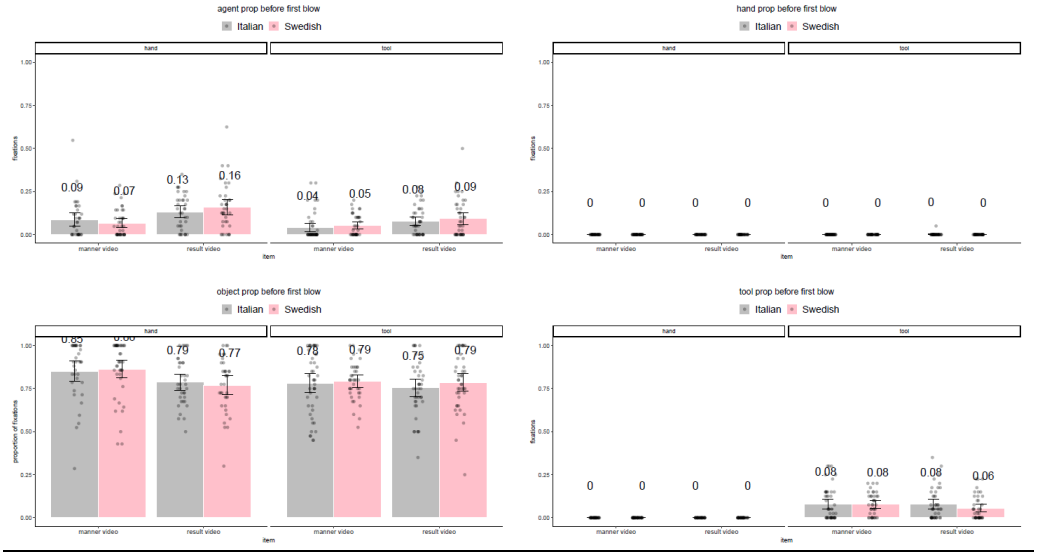


Figure 12. Boxplot of proportions of fixations in each AoI per language and condition in middle time window.

Main effect of language						
Object fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.79601	0.02771	859.348	28.729	0.000	***
language1	0.01239	0.02301	6.791.733	0.538	0.592	
Tool fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.038545	0.017849	7.121.468	2.159	0.067	.
language1	-0.006302	0.007232	69.847839	-0.871	0.387	
Agent fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.087418	0.019021	8.492.630	4.596	0.00151	**
language1	0.008238	0.014779	70.086139	0.557	0.57903	
Main effect with conditions						
Object fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.79609	0.02945	597.414	27.033	0	***
Manner.result1	-0.04353	0.05633	500.916	-0.773	0.474	
tool.hand1	-0.04000	0.05633	500.866	-0.710	0.509	
Tool fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.038545	0.012009	5.182.244	3.210	0.0226	*
Manner.result1	-0.006267	0.023784	4.985.759	-0.263	0.8027	
tool.hand1	0.077178	0.023784	4.985.483	3.245	0.0229	*
Agent fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.08737	0.01539	674.374	5.679	0.000857	***
Manner.result1	0.05052	0.02856	502.950	1.769	0.136756	
tool.hand1	-0.05063	0.02856	502.861	-1.773	0.136047	

Table 12. Result of the multilevel mixed models for main effects in the middle time window.

Interactions						
Object fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.79616212	0.03183544	467.221.484	25.009	0.000	***
language1	0.01258337	0.02302082	6.787.306.567	0.547	0.586	
condition1	-0.04330857	0.06126484	400.986.203	-0.707	0.519	
tool.hand1	-0.04032950	0.06126372	400.956.810	-0.658	0.546	
language1: Manner.result1	0.00008158	0.03029170	46.216.546.074	0.003	0.998	
language1:tool.hand1	0.03692129	0.03028215	46.159.169.207	1.219	0.223	
condition1:tool.hand1	0.05858498	0.12252685	400.949.162	0.478	0.657	
language1: Manner.result1:tool.hand1	0.06130919	0.06055980	46.147.894.150	1.012	0.312	
Tool fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.038552	0.013318	4.120.244	2.895	0.0428	*
language1	-0.006181	0.007249	69.843.804	-0.853	0.3967	
Manner.result1	-0.006318	0.026423	3.991.859	-0.239	0.8228	
tool.hand1	0.077208	0.026423	3.991.768	2.922	0.0433	*
language1: Manner.result1	-0.009432	0.012851	467.467.745	-0.734	0.4633	
language1:tool.hand1	-0.012994	0.012850	466.751.542	-1.011	0.3124	
condition1:tool.hand1	-0.012277	0.052846	3.991.742	-0.232	0.8277	
language.x1: Manner.result1:tool.hand1	-0.018734	0.025699	466.572.033	-0.729	0.4664	
Agent fixations	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	0.087166	0.016712	5.186.491	5.216	0.00307	**
language1	0.007826	0.014816	70.098.715	0.528	0.59900	
Manner.result1	0.050696	0.031345	4.024.635	1.617	0.18068	
tool.hand1	-0.050430	0.031345	4.024.173	-1.609	0.18250	
language1: Manner.result1	0.017176	0.018421	463.727.932	0.932	0.35160	
language1:tool.hand1	0.001255	0.018415	463.204.445	0.068	0.94569	
condition1:tool.hand1	-0.024475	0.062689	4.024.055	-0.390	0.71601	
language1: Manner.result1:tool.hand1	-0.058291	0.036826	463.105.319	-1.583	0.11413	

Table 13. Result of the multilevel mixed models for interactions in the middle time window.

5.5 Timecourse visualisation and Divergence analysis

In a design that features video stimuli, we are interested in the eye movements during the unfolding of the video. To account for the time variable, we can investigate the time course of trials, incorporating the proportion of eye gaze samples in the AoI. The time-course analysis starts extracting the time-course data from the fixation report. The experiment was conducted by tracking the sampling frequency of 500 Hz (1 sample every 2ms). To extract a time course from the eye tracker's samples, the fixations were binned into 100 ms intervals, and the length was cut to a maximum of 15 seconds. In the following sections, we visualise the data with time course plots and perform a divergence analysis.

5.5.1 Plotting the time-course

Before looking at patterns, we plot each item in each language (Figure 13) according to the length of the video displaying the event.

We found each item to be quite different from each other, even those which are in the same conditions (e.g., the two items for the Result-tool condition)[20]. The most similarities are in the Manner-hand and Result-hand conditions. However, we find similar patterns across the languages in each item. This shows that in the bigger picture, the eye gaze is driven by the scenes' semantics – thus, there are differences driven by the manner-result semantics and tool-hand – and the two languages show similar time courses per item. Table 13 shows the time course in the four conditions to visualise patterns across items.

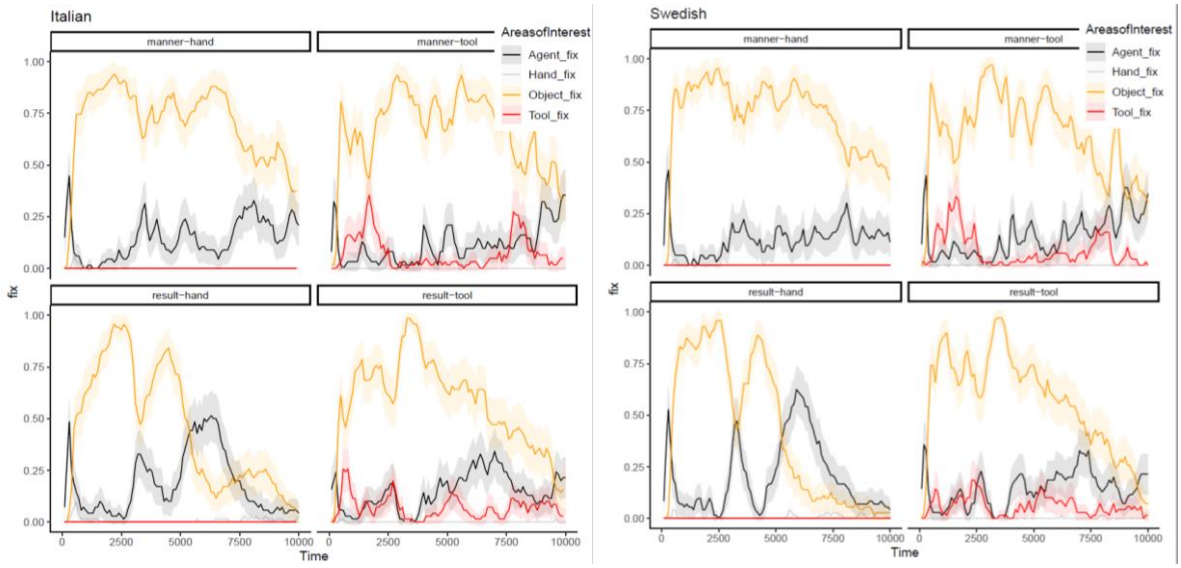


Figure 13.. Time-course per condition and language

[20] This is one of the limits of the study and a challenge in ecological stimuli. We need to keep in mind this diversity in drawing conclusions from this study. These issues are discussed in section 6.2

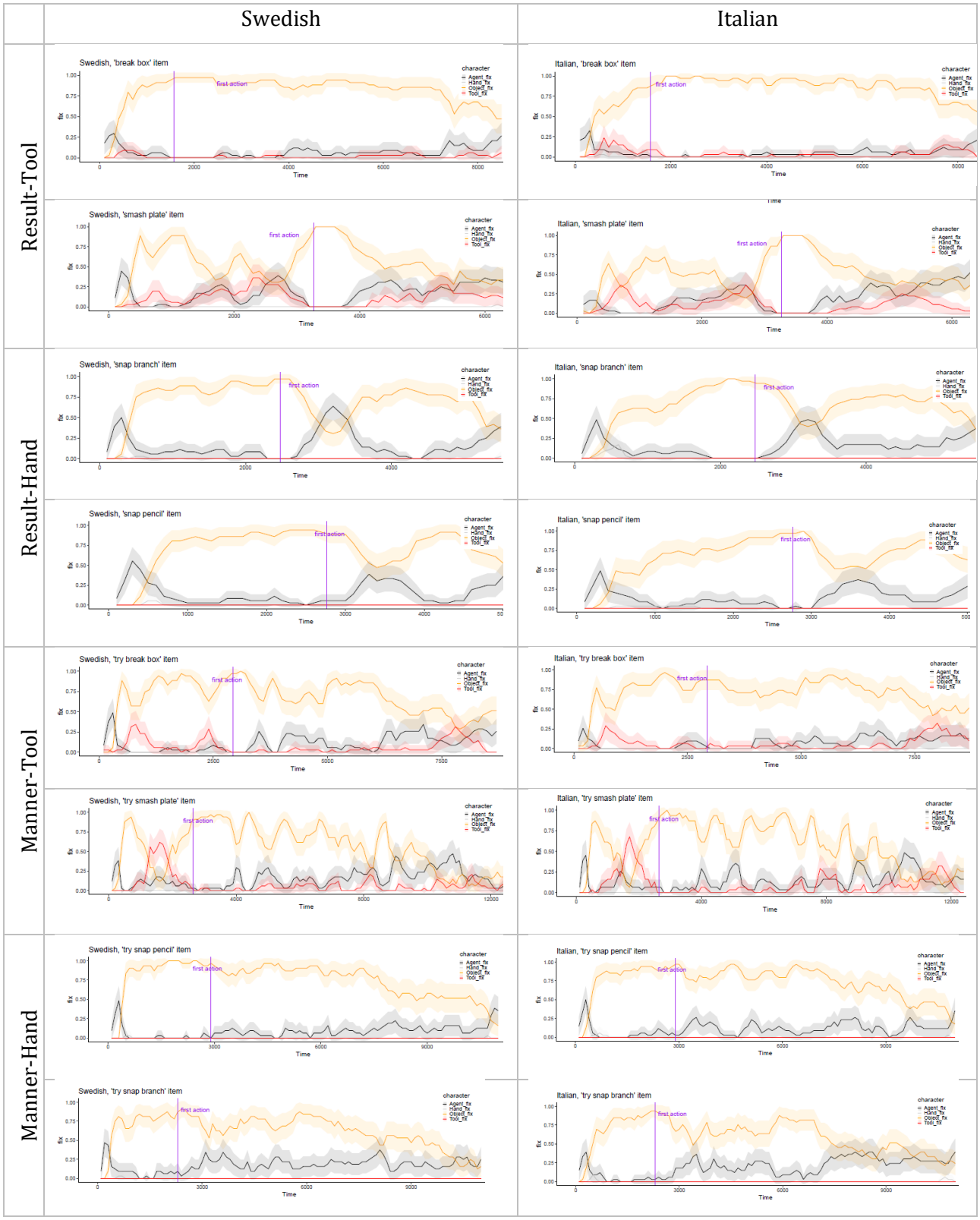


Table 13. Display of time-course per item, condition and language.

5.5.2 Divergence analysis

To assess the statistical effects of the eye-gaze data, we used the divergence analysis in the EyetrackingR package (Dink &Ferguson, 2015). This analysis divides the participant's averaged eye-movement data into time bins. It allows us to choose the preferred length of the time bins (in our case, 50 ms). Divergence analysis comprises 4 different analyses to test for divergences in two groups. The first analyses are a series of t-tests performed for each bin (t-test, Wilcox. test, linear models). It tests the difference in the proportion of looks towards the areas of interest. The second analysis is a bootstrapped smoothed divergence analysis. The third analysis is Bootstrapped cluster-based permutation analysis. The analysis can be performed on two conditions. We will subset our data and test conditions separately to compare results from analyses two and three. We are interested in language's influence on fixations. Thus, we compare Italian and Swedish fixations on each condition: manner-tool, manner-hand, result-tool, and result-hand.

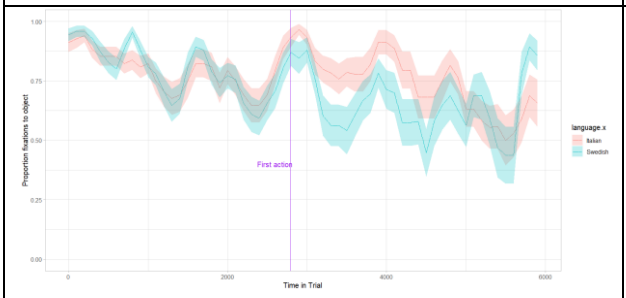
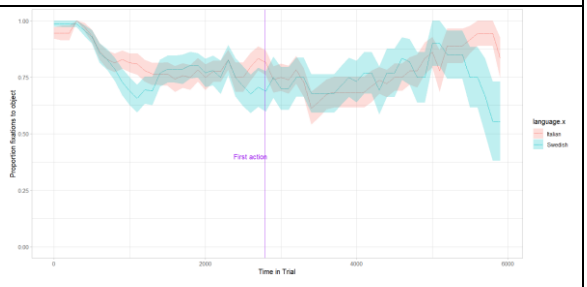
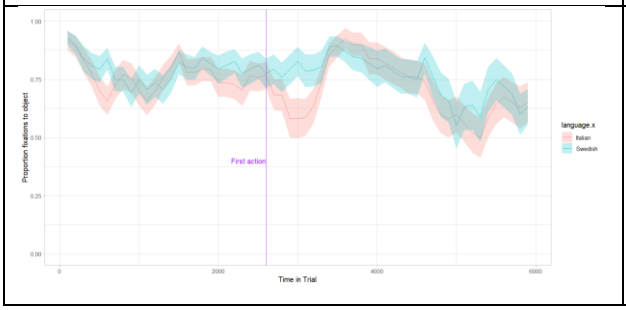
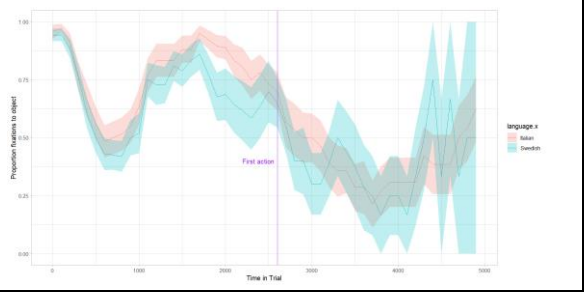
Tool manner	Tool Result
	
Hand manner	**Hand Result**
	

Table 14. Time-course of divergences per language and condition.

<u>Bootstrapped smoothed divergence analysis.</u>

This test operates on a smoothed version of the data. It resamples the data, replacing it with the smooth curve (e.g., randomly resampling subjects from each condition from our data) with small deviations in noisy data that could be a significant period of divergence. The process is reiterated thousands of times to obtain the 95% confidence intervals of this distribution. Time-bins whose

confidence intervals do not overlap with zero can be considered statistically significant at alpha = .05. (as used in Ferguson, Graf & Waxman 2017). The smoothed plots show a considerable time window of divergence in the Manner-tool condition from 4100ms to 5200ms (Table 15). The other conditions do not show as solid clusters as the Manner-tool condition. Following the divergence analysis pipeline (Dink & Ferguson, 2015), we compare these results to bootstrapped cluster-based permutations.

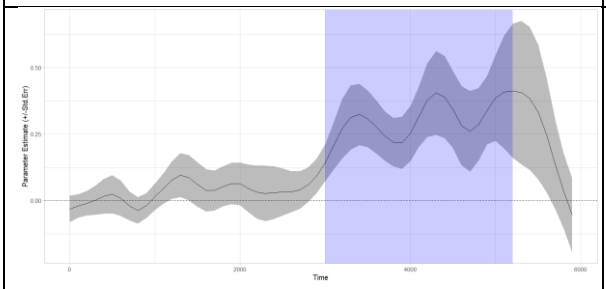
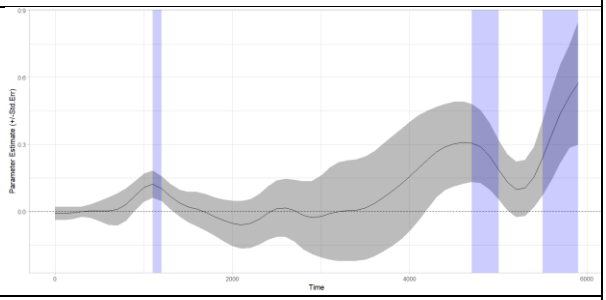
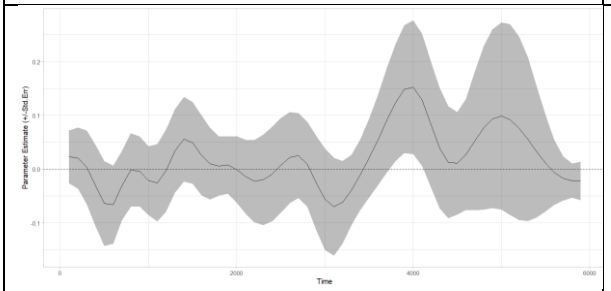
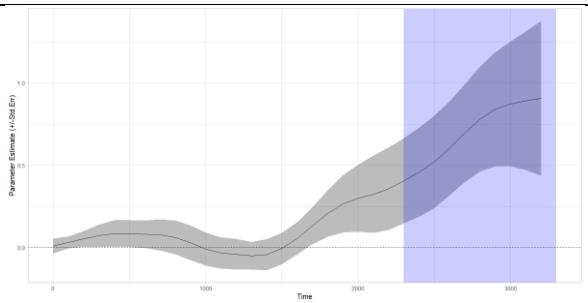
Manner Tool	Result Tool
	
Positive Run 1 ===== Time: 4100 - 5200	Runs of Significant Time Bins: Positive Run 1 Time: 1100 - 1200 Positive Run 2 Time: 4700 - 5000 Positive Run 3 Time: 5500 - 5900
Manner hand	Result hand
	
Runs of Significant Time Bins: -	Runs of Significant Time Bins: Positive Run 1 Time: 2300 - 3300

Table 15. Visualisation of the bootstrapped smoothed divergence analysis. The error bands represent 95% confidence intervals. Blue-coloured areas represent the time clusters during which fixations and conditions differed significantly.

Bootstrapped cluster-based permutations

Cluster-based permutations (Maris & Oostenveld, 2007) is a non-parametric test used to investigate the effect of a measure's condition over time. The analysis is performed by running a test on each time bin. The significant bins that are adjacent are kept and grouped together – forming a cluster. The data is then shuffled within subjects, and the procedure is reiterated (1000 times or more) to create a large amount of shuffled data. The shuffled data tells us what clusters we should expect if there were no effect – in a null distribution hypothesis – to compare with the original data (Ito & Knoeferle 2023). Ito & Knoeferle (2023) describe the advantages of this analysis as the possibility of detecting an effect of condition while controlling for autocorrelation of eye-movement data. Moreover, it maintains statistical power while controlling for multiple comparisons across many time bins – reducing the chances of false positives and false negatives. Ito & Knoeferle (2023) report as disadvantages the difficulties in the interpretation of the results. These data do not reveal the onset, offset or duration of an effect (as explained by Maris & Oostenveld 2007) because the *p*-value is derived from cluster-level statistics, so there is no statistical certainty or confidence about a specific time bin in the cluster. The authors' examples are that if a cluster individuated in the data goes from 200ms to 500 ms, this should not be interpreted as the start and end points of the effects or that the effect lasts 300ms. Rather, the effect should be interpreted as a significant difference between the two conditions somewhere in the time course.

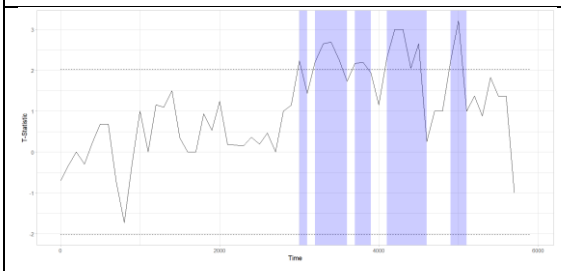
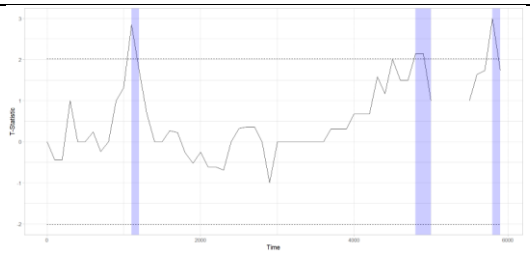
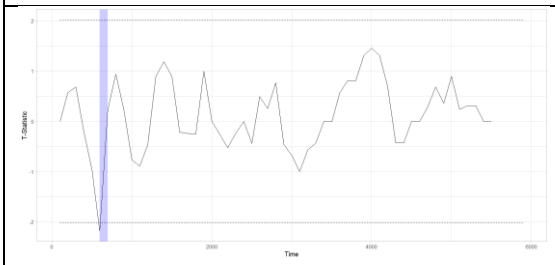
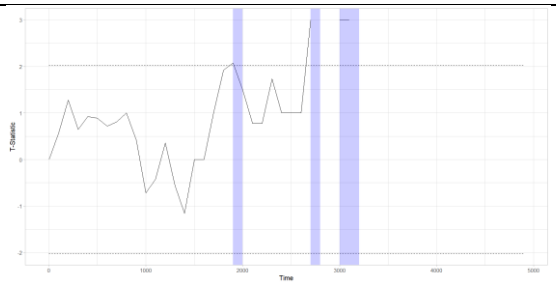
<table>
<tr><td>Tool manner</td><td>Tool Result</td></tr>
<tr><td>  </td><td>  </td></tr>
<tr><td>Cluster Direction SumStatistic StartTime EndTime Probability
1 Positive 2.2360 3000 3100 0.900
2 Positive 9.7803 3200 3600 0.210
3 Positive 4.3638 3700 3900 0.639
4 Positive 13.004 4100 4600 0.044*
5 Positive 5.4432 4900 5100 0.544</td><td>Cluster Direction SumStatistic StartTime EndTime Probability
1 Positive 2.845 1100 1200 0.56
2 Positive 4.276 4800 5000 0.46
3 Positive 3.000 5800 5900 0.56</td></tr>
<tr><td>Hand manner</td><td>Hand Result</td></tr>
<tr><td>  </td><td>  </td></tr>
<tr><td>Cluster Direction SumStatistic StartTime EndTime Probability
1 Negative -2.178193 600 700 0.753</td><td>Cluster Direction SumStatistic StartTime EndTime Probability
1 Positive 2.076 1900 2000 0.813
2 Positive 3.000 2700 2800 0.613
3 Positive 6.000 3000 3200 0.480</td></tr>
</table>

Table 16. Visualisation of the bootstrapped cluster-based permutations and summary statistics of the cluster analysis. The error bands represent 95% confidence intervals. Blue-coloured areas represent the time clusters during which fixations and conditions differed significantly.

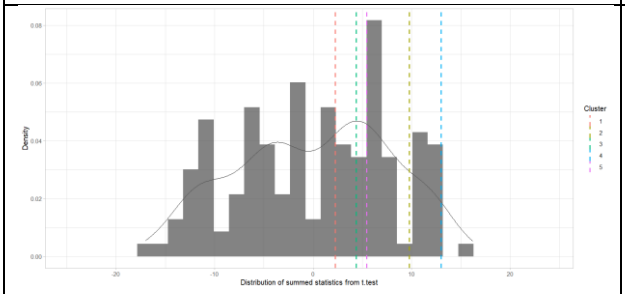
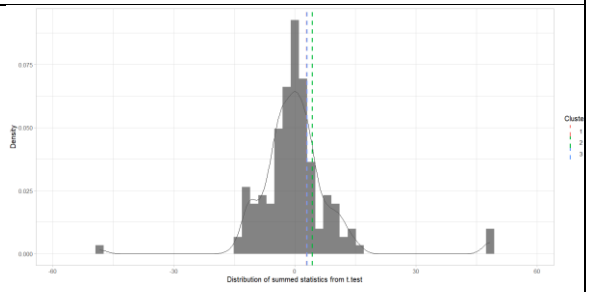
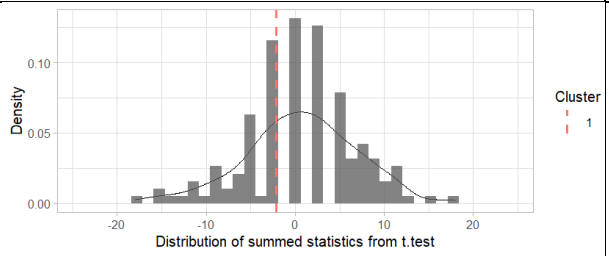
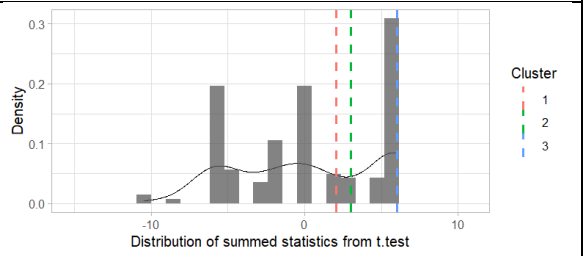
Tool manner	Tool result
	
Hand manner	**Hand result**
	

Table 17. Null distribution plot showing the summed statistics of a t-test performed on the clusters reported in Table 16.

In the Manner-tool condition, the significant cluster covers an area close to the one that was significant in the smoothed analysis. As we can see in the plots (Table 16 and 17), the time-bin clusters are close but separated, probably due to noisy data. The window of divergence is close to what we call the late time window.

6. Discussion

This study investigated cross-linguistic differences in the encoding of events, which differed in two characteristics: resultative or not resultative (thus, only displaying a manner of performance of the action) and hands or tools effectors of the action. These characteristics combined formed four conditions: Manner-tool, Manner-hand, Result-tool, and Result-hand. The conditions are motivated by linguistic differences in the verbal encoding of the action in Italian and Swedish. We tested these conditions in a within-subject design, asking Italian and Swedish native speakers to describe these events, tracking their eye movements and speech onsets.

Results suggest cross-linguistic differences in verb choices and sentence planning, which are assessed through verb onsets and eye movement. Verb choices reflect differences in event construal: Swedish and Italian participants expressed the events differently, with Swedish participants using manner more often with manner-result predicates. Construals provided different perspectives on event interpretations, among which the expression of an attempted result. We found earlier verb onsets for Swedish speakers. Proportions of fixations on interest areas analysed in three time-windows suggest that the differences between languages are more evident in the early time windows, with a main effect of language on object fixations and interactions of language and manner-result videos. Finally, a divergence analysis of the time course of video inspection suggested that the cross-linguistic differences are primarily evident in the Manner-tool condition, coherently with the linguistic differences, with Italian speakers allocating more attention to the objects after the first hitting action.

These results are consistent with our hypothesis, which predicted earlier verb onsets for Swedish due to the efficiency of message planning for Swedish participants for its verbal characteristics. Considering linguistic features, we hypothesised Manner-tool condition as the most revealing of differences: if Swedish speakers and Italian speakers use similar result verbs (as discussed in Study 2) for breaking actions performed with hands, Swedish speakers may encode manner and result separately for other categories of breaking events (which here we individuated as tool actions on three-dimensional objects). Taking these results together, we found an overall language effect that was persistent within conditions. Overall, the results condition has lower onsets, in both Italian and Swedish. We did not find an effect for verb onsets and verb type. To interpret this last result, we need to relate it to results above: as the count of verbs is skewed in favour of result verbs, we cannot exclude that the type of verb has an effect on onset times.

Manner-result and hand-tools

As visible in the time course plots, there are differences overall (non-language dependent) according to the manner or the result of the event.

Concerning the effectors, Result conditions indicate that hands in our events did not capture looks – crucially, they were not mentioned in the descriptions. Hammer or hands do not need to be mentioned. However, speakers can choose to mention them, and they did for the hammer – which also received more looks together with the verbal mention.

The role of the agent

Further considerations on message planning concern the role of the agent. Participants start the session with practice videos, where they have the chance to familiarise themselves with the background and the agent in the video, which is the same for all target and filler videos. They are asked to produce a sentence that always starts with a subject (*Maria*). Thus, participants do not strictly need to focus their attention on the agent. Fixations to the agent are still numerically relevant: the agent always catches attention. Coherently with the expectation and the literature review, we see from the time course plots that the agent has a spike in each item at the beginning. This testifies that it is fixed more consistently when mentioned, in a linear incrementality fashion, even if the participant knows who the agent is and does not need to plan the mention.

Message planning

The Manner-tool conditions can shed light on interpreting these results concerning message planning theory. First, earlier verb onsets for Swedish could be taken as indicative of a facilitation of encoding. Considering the discussion in section 3.3.3, this result is more significant than it would be in a picture description task. With video stimuli, participants of both languages must necessarily wait for the video's unfolding to produce the sentences, which is a period in which they are forced to continue planning and updating. Videos lead to longer times before onsets and induce a planning phase, which may be longer than needed (some participants wait until the end of the video before producing a sentence). For these reasons, it seems particularly relevant that onset times are significantly different between the two language groups.

Over the course of two different analyses (MLM and divergence analysis), we found that Swedish speakers allocate more fixations to objects than Italian speakers in the early time window and that Italian speakers allocate more fixations to the object later in the trial, approximately after the first hitting action. In the Swedish data, object fixations decrease after the first hitting action. The allocation of attention to the object in the earlier encoding window for Swedish speakers could be indicative of a more linear planning and the less effort needed to plan and update the structure and lexical choices of a sentence. This divergence is consistent with our tentative hypothesis but should nonetheless be taken with caution in consideration of some limitations in the design.

6.1 Limitations of the study and future directions

The experiment encountered several limitations that should be acknowledged and considered when interpreting results. Firstly, the production of the three types of construals (attempted result/manner/result) introduced variability in participant responses and diminished the data's homogeneity. This proved to be a limitation in testing the data according to the verb produced, as the result verbs produced were in the context of an attempt result construal of manner videos, thus confounding the analysis. An alternative data collection method could have imposed further constrained instructions on participant responses (but potentially sacrificing the ecology of the task). Second, the experimental design would have benefited from increased items per condition to enhance statistical power. Additionally, more coherence in event timings across items would have strengthened the results. In future studies, using a between-subjects design to test manner and result videos separately and avoid any priming effects seems promising. In terms of future directions, it could be interesting to better explore the role of effector of actions: for example, eliciting mentions of hands and tools and assessing their influence on eye movements when mentioned. Concerning manner and result in video and the attempted result construal, it could be used to explore event predictions of results and our inclination to over-interpret events and make sense of actions on a broader picture (thus, inferring a break from a hitting action).

7. Conclusion

This study moved on from previous work described in Chapters 4 and 5, which were focused on categories and construals cross-linguistically, to investigate the breaking category with an online task. The aim was to examine the speech planning of event descriptions, motivated by linguistic differences in Italian and Swedish in how breaking events are described. Causative breaking events are characterised by a result state occurring in the object after an action that produces a material separation and the loss of functionality of said objects. The result can be achieved with different action performances – that is, manners. In Italian, a verb root can indicate the result state (*rompere,* 'break') without conveying any information on the action performances. In Swedish, a predicate composed of a light verb and an adjectival particle (*sönder,* 'in pieces') is used to convey the result state of most breaking actions and a verb root (*bryta,* 'snap') is used primarily for breaking actions performed on one-dimensional items. To describe the manner of the action, both Italian and Swedish may use verbs denoting a hitting action, a bending action, or lexicalise an instrument. In addition, Swedish use a predicate composed of a manner verb and adjectival

particle *sönder* (e.g., *hamra sönder*, 'hammer in pieces'), which allows to pack together the manner and result pieces of the breaking events. Testing Italian and Swedish native-speaker participants, this experiment revealed linguistic differences in construals, as well as onsets and eye movements – taken to reflect message planning. Results show that Swedish speakers use manner verbs more frequently, especially in manner-result predicates construals. Swedish participants' early onsets and the differences in the allocation of attention to the object (Swedish participants allocating more attention to the object in the first window of analysis) are interpreted as facilitation of Swedish in the encoding of breaking events: the structure and lexical possibilities of Swedish in the expression of breaking events allows participants to plan more linearly. Finally, having highlighted how the models of language planning and their predictions cannot be compared accurately to our data, further studies could be directed at investigating models of speech production in video-based elicitations.

General Conclusions

Throughout this work, we explored the tight link between categorisation and construals in event cognition. Categorisation plays a crucial role in making sense of the world, allowing us to organise information efficiently. Under a Cognitive Linguistic framework, construals represent our conceptualisation of events and give access to the interpretations we make. Languages vary in how they represent categories in lexical items and structures, showing different patterns and construal possibilities. Lastly, events are dynamic processes that we perceive and describe through different linguistic perspectives.

This work was structured in 5 chapters. Chapter 1 introduced the theoretical framework that underlies this research – that is Cognitive Linguistics assumptions, which connect the mind and language to what humans perceive in the world and give meaning to. It emphasized the essential aspects of cognition and language that are particularly relevant for studying event descriptions. To understand how we conceptualise, process, and communicate about events, we discussed the cognitive mechanisms of categorisation and construal, to then present research in event cognition and discuss the stages of apprehension and planning. A brief review of research in scene descriptions, particularly with eye tracking, illustrates how language influences eye movements in verbal description tasks, impacting planning and cognitive processing. In Chapter 2 we reviewed the linguistics studies in actions and events, presenting different classifications that have been used to investigate verbs. We then presented the linguistics research on verbs that describe breaking events, highlighting the differences across languages. We operate a delimitation of the semantic domain of investigation by providing a definition of the breaking category to be used for the practical purposes of the research. Chapters 3 to 5 presented empirical and experimental research. In Chapter 3, we introduced an online video resource, Imagact, and explained how the resource can be used to investigate construals: it provides visual support for investigating actions and linguistic annotation of different verbs that can be used to describe the same event. As concepts are lexicalised into items, but lexicalisations vary across languages, cross-linguistic differences within a common domain are investigated in Study 1 with a fine-grained contrastive investigation. The study aimed at exploring category boundaries of a sample of verbs, to assess their fuzziness, and typicality of the category members. We used corpora occurrences and operationalised the study of category boundaries by extracting from literature and visual inspection of scenes the relevant parameters of *break* verbs. The parameters annotated corpora occurrences and contrasted the verbs in the two languages based on the annotations. Chapter 4 investigated the levels of abstraction of lexical items that describe a set of 25 breaking events. We asked English, Italian, and Swedish participants to describe the events in a written format and

analysed the result in terms of lexical variety, codability and according to verb-scene relations visualized in a graph-based connectivity map. In chapter 5 we examined the cognitive mechanisms underlying spoken production with breaking events. The interest was in linguistic differences between Italian and Swedish, which display different encoding possibilities in the domain of breaking events. The planning of event descriptions in these two languages was investigated by means of verb onset timings and eye fixations and analyses with statistical testing.

Event categorisation

Categories play a vital role in assigning meaning to events and structuring experiences in the world, while construals provide the linguistic interpretation and subjectification of events. In Cognitive Linguistics, categorisation theory abandons the classical view based on logic and necessary conditions to adopt a theory of prototypes based on typicality and fuzziness. This is linked to the assumption of the organisation of categories in a hierarchical model (superordinate – basic – subordinate) where basic level categories are the most relevant in cognition due to their relevance for our everyday purposes. Basic-level categories help organise concepts in efficient ways and are found to be represented at the lexical level. The hierarchical organisation corresponds to levels of abstraction that are linguistically realised on a generality-specificity cline and termed hypernyms and hyponyms. The relevance of basic level categories in cognition, especially for objects, is an established finding in the literature, but this is not necessarily reflected in event categories. We showed that the literature review reveals a focus on activities (e.g., eat or run) in the lexical category of verbs but leaves open the discussion about events described with a higher level of abstraction captured by *general verbs*. For activities which generally involve the description of some action (e.g., eat, walk), troponymy relations at the basic level are also realised – extending on a horizontal level, compared to hyponymy and hypernymy: it is a relation of inclusiveness of the higher level of abstraction, but also a basic way of carrying out any activity that can stand in opposition to its lower level term (e.g., eat, devour). We propose to direct the attention to general verbs, like *open* or *break*, which categorise events without common perceptual features (unlike activities) at a more abstract level.

Speech production and message planning

A point of interest for event description is to investigate how we extract information from an event during its unfolding, conceptualise the event and plan a sentence to describe it. Studies in production focus on the integration between sentences and world scenes. This integration depends on the rapid processing of dense information. Speech planning research is concerned

with lexical retrieval and structural building, as well as how these components are built within two phases: pre-message planning, that is, conceptualisation, and sentence-level planning, and the way they interact in the incremental planning that leads to sentence production (Linear Incrementality or Hierarchical Incrementality). Within these phases, the event components that need to be processed can be divided into relational encoding (event participants) and non-relational encoding (action information). Eye-tracking is a successful methodology in this research area to investigate which parts of the event are inspected and promote models of sentence planning. These studies are relevant for linguistic and Cognitive Linguistics purposes: they allow us to investigate the interplay of cognitive mechanisms and linguistic realisations as they occur.

To pull the threads together we sum up the main findings of the research.

Verb category boundaries are fuzzy, and category membership shows typicality effects, as emerged in Study 1. In a usage-based approach, data from corpora can be used to find the typical events referred to by a verb (operated through parameters that characterise events) and determine different degrees of fuzziness in the boundaries comparing verb pairs. Italian verbs *rompere-spezzare* and *rompere-spaccare* have fuzzier boundaries than English *break-snap* and *break-smash*. From a contrastive point of view, *break-rompere* are suitable translation equivalents, while snap-spezzare and smash-spaccare show different prototypical configurations while showing a degree of overlap at the category periphery.

Event categorisation is different in different in languages and, crucially, it is proven to be different from categorisation of discrete entities and activities. In this research, this is explained by acknowledging the nature of events as relational, abstract, and encompassing more aspects of world knowledge than discrete entities or activities. A relevant aspect of event categorisation appears to be that of the interconnectedness of categories: at the linguistic level, events can be described with various construals pointing at different conceptualisations and evoking different categories. The domain of breaking events, as captured by general verbs, shows a high degree of abstraction, allowing us to capture visually different events. There seems to lack a superordinate, more abstract level than that of general verbs that is relevant in lexicalisations, and the subordinate level does not rely on troponyms (as for activities).

Event categorisation in languages is affected by properties of the event itself (as demonstrated by the underlying shared categorization of the breaking domain to some extent which emerged in Study 1 and 2) but at the linguistic level, it is also limited by each language's conventions and lexicalisations. In fact, each language has a preferred level of abstraction or conventionalisations.

The breaking category as expressed by construals in an event description task show differences across languages based on the lexical variety of each language and the degree of generality-specificity of the verbs. Coherently with previous studies, Swedish has a higher specificity compared to English, and, as emerging in this study, compared to Italian. This is in accordance with the typological distinction of satellite-framed and verb-framed languages. In fact, Italian is more general than Swedish. However, our results show that English is closer to Italian than Swedish. What emerges clearly is that the Swedish categorisation of the breaking category revolves around an adjectival particle (*sönder*, 'in pieces') rather than a verb, like in English and Italian. This finding should be taken into consideration in further exploration of event categories, extending the focus of research to other elements than verbs.

Linguistic differences emerged in the planning of breaking events between Italian and Swedish, visible in verb onsets and locations of fixations throughout experimental trials. Swedish participants early onsets and their sustained allocation of attention to the object early in the unfolding of the event are interpreted as facilitation of Swedish in the planning of a description, resulting in a linear planning. This result highlights the influence of language in cognitive processing and conceptualisation of events.

Future studies should continue the investigation of event categories, turning to different domains and more languages. While we found language differences in speech planning, we found that this experimental design had some limitations. As mentioned in Chapter 5, the number of items was limited per condition was limited, and items should have more similar timings. Moreover, further manipulations in the conditions and improvements in the design could enhance the relational encoding window. For example, controlling for participants' time of video exposure. However, since videos provide more ecological validity in testing speech planning (it mimics how we look at events in real life), we believe it could be worth improving this design. Future research should be directed to using videos in event descriptions and refining models of speech planning.

Appendices

Appendix A

	Scene description	Italian verbs	English verbs	Swedish verbs
1	Crush flowers with foot	Schiacciare, Pestare, Calpestare	Trample, mash, flatten, crush, squash, tread, smash	Trampa, kliva, platta till
2	Machine cuts and gathers crops	Raccogliere, Mietere	Gather, pick, harvest, reap	Skörda
3	Blending fruit inside a blender	Frullare	Blend, liquidize	Blanda, Mixa
4	Passing tomatoes	Passare	Mash, crush, puree	Mosa, pressa, krossa
5	Pierce sheets of paper	Mettere, Applicare	Pierce	Göra hål
6	Cut a rope	Tagliare, Accorciare	Shorten	klippa
7	Cut a bun open	Tagliare	Cut open	kära
8	Grind flour	Macinare	grind	mala
9	Crush a piece of fruit with foot	Schiacciare, Pestare	Trample, mash, flatten, crush, squash, tread, smash	Trampa, mosa, kliva, trampa sönder, trampa itu
10	Crush a potato with a fork	Schiacciare	Mash, crush	Mosa
11	Crush a can with a pressing machine	Schiacciare, Pressare, Comprimere	Mash, flatten, crush, press, compress, smash	Trycka ihop, platta till, pressa, ihop, knyckla ihop
12	Crush basil to make pesto	Pestare	Crush	Mosa
13	Divide twig in parts	Spezzettare, Dividere	Split up, split, divide	Dela, bryta, knäcka
14	branch from plant	Staccare, Levare, Togliere, Rimuovere	Take, pull, snap, remove, detach	Bryta, knäcka av, bryta loss
15	Peel peanuts	Sgusciare, Aprire	Shell, open up, open	Knäcka, skala
16	Smash car on a wall	Schiantare, Picchiare, Sbattere	Hit, collide, crash, ram, smash, strike, slam	krocka
17	Tear paper in two	Strappare	Tear, rip	Riva, riva itu, riva sönder
18	Tear paper in four	Strappare, Stracciare	Tear, rip, tear up, rip up	Riva, riva itu, riva sönder
19	A brick wall collapsed	Cadere, Crollare, Cascare, Cedere	Collapse, fall	Rasa, störta, falla samman, kollapsa, falla sönder
20	A building explodes	Distruggere, Fare crollare	Destroy	Falla ihop, förstöra, rasa, rasera
21	Break chair separating legs from the back	Rompere, Staccare Danneggiare,	Break, bust	Slita av, slita loss, slita sönder, göra sönder
22	Smash different pieces of furniture in a room	Sfasciare, Distruggere	Destroy, smash	Göra sönder, Vandalisera, Slå sönder, Slå itu, Förstöra
23	Crack walnuts	Rompere, Aprire, Schiacciare	Crack, break, open, open up	Knäcka
24	divide rope in two by pulling the extremes	Rompere, Spezzare	Break, snap	Bryta av, Slita isär, Slita sönder
25	Hulls the peas	Sgranare	Hull, Shell, Shuck	Skala, Rensa
26	smash bottle on a wall	Spaccare, Rompere	Break, smash, bust	Krossa, göra sönder, slå sönder
27	Crack an egg	Rompere, Aprire	Crack, break, open, open up	Knäcka
28	Breaking tile with tool	Rompere, Spezzare, Spaccare	Break, bust	Bryta, spricka, göra sönder, bryta sönder, bräcka, bräcka sönder
29	Breaking a wheel off a toy car	Rompere, Staccare Danneggiare	Break off	Göra sönder, bryta av
30	crack a mirror with tool	Rompere, Spaccare	Crack	Slå sönder, Spräcka, Göra sönder
31	Force with a tool to open a safe	Rompere, Aprire, Forzare	Break, open, open up	Öppna, Bryta upp, Bräcka upp
32	divide pencil in two	Spezzare, Rompere	Break, snap,	Bryta av, Slita isär, Slita sönder
33	Smash egg with a hand	Schiacciare	Mash, flatten, crush, squash, smash	Krossa, Göra sönder, Mosa
34	Smash wooden box with hammer	Sfasciare, Rompere, Distruggere	Break, smash, destroy, bust	Krossa, Förstöra, Banka, Göra sönder
35	Create a hole in a wooden box with a single hammer blow	Sfondare, Rompere, Danneggiare	Break, bust	Göra hål, Hamra

Appendix B

	Scene	Italian Verbs	English Verbs	Swedish Verbs
1	smash glass bottle on wall	Spaccare, Rompere	Break, Smash, Bust	Krossa, Göra sönder, Slå sönder
2	wrestle-breaking arm	Rompere	Break, Bust	Bryta
3	break tile with tool	Rompere, Spezzare, Spaccare	Break, Bust	Bryta, Spricka, Göra sönder, Bryta sönder, Bräcka, Bräcka sönder
4	break wheel off a toy car	Rompere, Staccare Danneggiare	Break off	Göra sönder, Bryta av
5	make a fan non-functional	Rompere, Danneggiare	Break, Bust, Damage	Förstöra, Göra sönder
6	crack mirror with tool	Rompere, Spaccare	Crack	Slå sönder, Spräcka, Göra sönder
7	force with a tool to open a safe	Rompere, aprire, forzare	Break, Open, Open up	Öppna, Bryta upp, Bräcka upp
8	computer display not working	Danneggiare, Rovinare	Break, Damage	Förstöra, Göra sönder
9	divide twig in parts	Spezzettare, Dividere	Split, Split up, Divide	Bryta, Knäcka, Dela
10	divide pencil in two	Spezzare, Rompere	Break, Snap	Bryta av, Slita isär, Slita sönder
11	detach a branch from a plant	Levare, Togliere, Staccare, Rimuovere	Detach, Take, Pull, Snap, Remove	Knäcka av, Bryta loss, Bryta
12	smash egg with a hand	Schiacciare	Mash, Flatten, Crush, Squash, Smash	Krossa, Göra sönder, Mosa
13	smash wooden box with hammer	Sfasciare, Rompere, Distruggere	Break, Smash, Destroy, Bust	Krossa, Förstöra, Banka, Göra sönder
14	create hole in a wooden box	Sfondare, Rompere, Danneggiare	Break, Bust	Göra hål, Hamra
15	break chair separating legs from the back	Rompere, Staccare Danneggiare,	Break, Bust	Slita av, Slita loss, Slita sönder, Göra sönder
16	smash furniture in a room	Sfasciare, Distruggere	Destroy, Smash	Göra sönder, Vandalisera, Slå sönder, Slå itu, Förstöra
17	peel peanuts	Aprire, Sgusciare	Open, Shell, Open up, Break apart	Skala, Knäcka
18	crack walnuts	Rompere, Aprire, Schiacciare	Crack, Break, Open, Open up	Knäcka
19	smash car on a wall	Schiantare, Picchiare, Sbattere	Hit, Collide, Crash, Smash	Krocka
20	crack an egg	Rompere, Aprire	Crack, Break, Open, Open up	Knäcka
21	tear paper in two	Strappare	Tear, Rip	Riva itu, Riva sönder, Riva
22	tear paper in four	Strappare	Tear, Rip, Tear up, Rip up,	Riva itu, Riva sönder, Riva
23	break open door	Sfondare	Break, Break open	Slå in, Bräcka upp, Forcera
24	smash tomatoes with machine	Passare	Mash, Puree, Crash	Pressa, Krossa, Mosa
25	divide rope in two by pulling the extremes	Rompere, Spezzare	Break, Snap	Bryta av, Slita isär, Slita sönder

Appendix C

	Other domains	Breaking domain
Italian	Accendere, aggiustare, aprire, armeggiare, avere problemi, bloccare, colpire, controllare, dare una martellata, dare una spallata, devastare, dislocare, dividere, fare funzionare, fare forza, fermare, forzare, girare, inceppare, irrompere, martellare, mettere pressione, prendere a martellate, raccogliere, riavviare, riparare, sbattere, sbucciare, scardinare, scassinare, separare, sganciare, sgusciare, sistemare, smontare, spegnere, staccare, stringere, svitare, testare, tirare, togliere, usare, vandalizzare	Bucare, distruggere, divellere, fare a pezzi, fare a pezzetti, fare una crepa, manomettere, recidere, ridurre in pezzi, rompere scassare, schiacciare, sfasciare, sfondare, sminuzzare, spaccare, spappolare, spezzare, spezzettare, spiaccicare, stracciare, strappare
English	Adjust, barge into, bash, check, de-shell, disassemble, dislocate, dismantle, empty, fiddle, fix, flatten, force, get to work, hammer, hammer through, have difficulty, hit, hold up, insert, knock out, lever, leverage, make, mess, move, open, peel, peel off, press, press down, prise off, pry, push, push into, put, put through, remove, separate, set up, shell, shove, slam, slap, smack, splash, squash, squeeze, stick stomp, storm, take off, touch, trash, turn on, use, wedge, work	Barge through, Barge open, Break, Break into, Break open, Burst, chip, chop, crack, crush, damage, demolish, destroy, force open, knock through, hit through, make a whole, rip (+ apart, off, up), shatter, shred, smash (+ up, through), snap, snap up, tear (+ open, off, apart), shoulder open, put a whole
Swedish	Banka, dela, dela upp, dra (+ ur, av), drämma, dyrka upp, få igång, fixa, forcera, fungera, gå bärsärkagång, göra till, greja, ha problem, justera, knäppa, knuffa, laga, lossna, mixtra, mosa, öppna, passera, plocka av/bort, plocka, rätta till, reparera, rycka (+ av, loss), sätta på, sätta igång, skala, skörda, skruva, slå, slå ur, slå på, slå in, slå upp, stänga, starta, stoppa, strimla, ta av, ta loss, ta bort, tackla upp, tackla i, trycka, trycka ner, trycka upp, välta ner, vandalisera	Bända upp, banka sönder, bryta, bryta sönder (av, itu, loss, upp) bryta sig in, demolera, dra sönder (+loss, isär), förstöra, ha hål, ha sönder, hamra sönder, knäcka, krossa, riva (+ sönder, ner, isär, itu, upp, av), rycka isär, rycka sönder, rycka bort, skruva sönder, slå sönder, slå hål, slå genom, slita (+ loss, av, isär), splittra, spräcka, ta itu, ta isär, trasa sönder, trycka sönder

Appendix D

English	Counts	Italian	Counts	Swedish	Counts
break	145	rompere	173	knäcka	74
crack	64	strappare	64	slå sönder	70
smash	60	spezzare	61	bryta	45
snap	47	aprire	42	krossa	31
open	26	distruggere	34	bryta av	29
rip	26	staccare	31	ha sönder	27
tear	25	spaccare	29	riva	27
hit	22	schiacciare	28	bryta upp	23
fix	15	sfondare	24	skala	21
destroy	13	sgusciare	21	spräcka	19
pull off	12	sbucciare	14	öppna	17
break off	12	dividere	10	riva sönder	17
take off	11	smontare	9	dra av	16
crush	10	aggiustare	8	förstöra	13
smash up	10	spezzettare	8	riva itu	12
turn on	7	togliere	6	slå hål	11
hammer	7	forzare	5	dra sönder	8
break into	6	riparare	5	laga	8
break open	6	colpire	4	dra loss	7
pull	5	far funzionare	4	slå	7
rip up	5	sistemare	4	bända upp	7
remove	4	bloccare	4	mosa	7
force	4	controllare	3	slita av	6
adjust	4	fare una crepa	3	ta av	6
mess	4	raccogliere	3	justera	6
smack	4	devastare	3	rycka	5
snap off	4	stracciare	3	rycka av	5
stick	4	dislocare	3	slå in	5
stop	4	separare	2	tackla upp	5
pull apart	3	armeggiare	2	fungera	5
use	3	bucare	2	bryta itu	5
hammer through	3	dare una martellata	2	banka	4
rip off	3	martellare	2	slå på	4
push	3	prendere a martellate	2	ta isär	4
shell	3	spiaccicare	2	fixa	4
tear up	3	fare a pezzi	2	bryta sönder	4
dislocate	3	sfasciare	2	demolera	4
disassemble	2	dare una spallata	1	knuffa	3
barge into	2	irrompere	1	ta bort	3
barge open	2	fare forza	1	ha problem	3
bash	2	mettere pressione	1	banka sönder	3
check	2	sganciare	1	plocka	3
get to work	2	accendere	1	slita isär	3
put through	2	avere problemi	1	skruva	3
tear off	2	riavviare	1	stoppa	3
press down	2	spegnere	1	hamra sönder	3
rip apart	2	sbattere	1	bryta loss	2
put	2	usare il martello	1	rycka loss	2

peel	2	recidere	1	bryta sig in	2
squash	2	tirare	1	slå upp	2
pull over	2	fare a pezzetti	1	trycka upp	2
trash	2	ridurre in pezzi	1	splittra	2
demolish	2	divellere	1	slita loss	2
dismantle	1	scardinare	1	ta loss	2
barge	1	scassare	1	rätta till	2
barge through	1	scassinare	1	starta	2
break down	1	fermare	1	sätta på	2
break through	1	girare	1	slå genom	2
burst	1	inceppare	1	rycka isär	2
knock through	1	manomettere	1	dela	2
shoulder open	1	stringere	1	dela upp	2
shove	1	svitare	1	riva isär	2
storm	1	testare	1	forcera	1
push into	1	spappolare	1	tackla i	1
shatter	1	vandalizzare	1	borra sönder	1
damage	1	sminuzzare	1	trycka	1
fiddle	1			trycka ner	1
have difficulty	1			dra ur	1
knock out	1			lossna	1
move	1			plocka av	1
set up	1			plocka bort	1
slam	1			sätta igång	1
touch	1			rycks bort	1
work	1			knäppa	1
empty	1			drämma	1
separate	1			ha hål på	1
squeeze	1			slå en slägga	1
hit through	1			slå i	1
make a hole	1			skörda	1
put a hole	1			dra	1
smash through	1			dra isär	1
pick	1			riva av	1
tear apart	1			rycka sönder	1
force open	1			slita	1
lever	1			dyrka upp	1
leverage	1			ta itu	1
prise off	1			få igång	1
pry	1			greja	1
tear open	1			mixtra	1
hold up	1			reparera	1
insert	1			skruva sönder	1
wedge	1			stänga	1
break apart	1			trycka sönder	1
crack open	1			gå bärsärkagång	1
de-shell	1			riva ner	1
peel off	1			trasa sönder	1
flatten	1			vandalisera	1
slap	1			välta ner	1
splash	1			slå ur	1
chip	1			riva upp	1
shred	1				

Appendix E

	Scene	Swedish	Italian	English
1	break chair separating legs from the back	0.89	0.78	0.71
14	break open door	0.87	0.58	0.91
24	break tile with tool	0.84	0.56	0.52
34	break wheel off a toy car	0.90	0.68	0.80
49	computer display not working	0.88	0.88	0.88
59	crack an egg	0.00	0.32	0.38
60	crack mirror with tool	0.68	0.54	0.73
65	crack walnuts	0.29	0.80	0.60
68	create hole in a wooden box	0.74	0.83	0.87
77	detach a branch from a plant	0.82	0.66	0.85
90	divide pencil in two	0.72	0.32	0.55
95	divide rope in two by pulling the extremes	0.90	0.69	0.67
109	divide twig in parts	0.75	0.64	0.55
117	force with a tool to open a safe	0.60	0.74	0.84
122	make a fan non-functional	0.89	0.85	0.78
136	peel peanuts	0.47	0.64	0.85
140	smash egg with a hand	0.61	0.43	0.82
146	smash furniture in a room	0.73	0.54	0.78
157	smash glass bottle on wall	0.49	0.40	0.42
160	smash wooden box with hammer	0.65	0.72	0.85
168	tear paper in four	0.64	0.35	0.59
174	tear paper in two	0.66	0.34	0.71
180	wrestle-breaking arm	0.00	0.44	0.29

Appendix F

Italian Instructions Study 2

Vedrai dei brevi videoclip (3-10 secondi) in cui ci sono dei personaggi che compiono azioni di diverso tipo.
Descrivi cosa succede nei video con una sola frase.
Istruzioni:

- guarda le clip fino alla fine in un dispositivo che consente la corretta visualizzazione.
- Usa questa frase standard: soggetto + verbo + oggetto e complementi che ritieni necessari a descrivere l'azione. Segui gli esempi:

Luca afferra la palla
Luca mescola le verdure

Per i soggetti umani usa nomi propri, come Luca e Maria.

- Preferisci frasi attive, ovvero che iniziano con un soggetto umano
Luca mescola le verdure e non Le verdure sono mescolate
- Non scrivere frasi troppo generiche (ex. Luca fa qualcosa).
- Non dare giudizi o opinioni

Le risposte che non seguono le istruzioni e la struttura della frase standard non saranno accettate.

English Instructions Study 2

Your task in this experiment is to watch short video-clips (from 3 to 11 seconds) and provide a brief caption. Watch the video and write ONE sentence to describe it.

Instructions:

1. Watch the videoclips until the end.
2. Use the following structure to write your descriptions: Subject + verb + object. Add other information if you think it is necessary to describe the action you see. Follow these examples:

Mary takes the book
John drives the car

3. The sentence should describe the physical action the agent does (Mary opens the door). It should not be generic (Mary does something).

Do not think too much about it, just write what is natural to you.
Please follow the instructions carefully!

Swedish Instructions Study 2

Du kommer att se korta videoklipp (3-10 sekunder) som visar människor som utför handlingar. Vänligen följ instruktionerna. Om du inte följer instruktionerna är ditt deltagande inte giltigt.
Instruktioner:
- Titta noga på videorna och beskriv vad du ser med **bara en mening**
- Använd en standardmening: Namn på en person + verb + objekt och övrig information som du tycker är nödvändig för att beskriva videon. Du kan alltid använda samma namn för personerna i videorna. Till exempel Lars och Britta. Följ exemplen:

Lars tar bollen
Lars rör om grönsakerna i pannan

- Var snälla specifik. Undvik meningar som: "Lars gör något"
- Ge inte åsikter eller din personliga bedömning av handlingarna.

Tänk inte för mycket! Skriv det som låter naturligt på svenska.

Italian Instructions Study 3

Iniziamo!

Vedrai dei brevi video della durata di 5-12 secondi. Nei video, vedrai sempre la stessa persona compiere
diversi tipi di azioni. Dovrai guardare i video e descrivere cosa succede con **una sola frase**, iniziando sempre con il nome “Maria”

Per esempio:
Maria mangia una mela.
Maria cammina.

Inizia a descrivere appena hai capito cosa succede nel video. Ti accorgerai che un’azione può essere descritta in più modi. Usa il primo a cui pensi e non fornire un’alternativa.

Adesso iniziamo con due video di prova.
Premi qualsiasi tasto per iniziare

Swedish Instructions Study 3

Instruktioner
Du får se 40 videoklipp (5-12 sekunder) där en och samma person utför några handlingar.
Beskriv vad du ser med **en mening**. Kom ihåg att alltid börja med namn ”Maria”, t.e.:

Maria går
Maria äter ett äpple

Börja så fort du förstått vad som händer.
Det finns fler sätt at beskriva samma handling. Ta den du först tänker på
och ge inget alternativ.

Tryck sedan på valfri tangent på tangentbordet för att gå till nästa videoklipp.

Du kan träna med två videoklipp.

Tryck på valfri tangent för att gå till träningen

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. *Journal of Memory and Language*, *38*(4), 419–439.

Almeida, R. G. De, Nardo, J. Di, Antal, C., & Grünau, M. W. Von. (2019). Understanding Events by Eye and Ear : Agent and Verb Drive Non-anticipatory Eye Movements in Dynamic Scenes. *Frontiers in Psychology*, *10*(October), 1–20. https://doi.org/10.3389/fpsyg.2019.02162

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1

Anderson, J. R., Bothell, D., & Douglass, S. (2004). Eye Movements Do Not Reflect Retrieval Processes: Limits of the Eye-Mind Hypothesis. *Psychological Science*, *15*(4), 225–231.

Anelli, F., Nicoletti, R., & Borghi, A. M. (2010). Categorisation and action: What about object consistence? *Acta Psychologica*, *133*(2), 203–211. https://doi.org/10.1016/J.ACTPSY.2009.11.009

Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, *103*(1), 62–70. https://doi.org/10.1037/H0036799

Arstein, R. (2017). Inter-Annotator Agreement. In N. Ide & J. Pustejovsky (Eds.), *Handbook of Linguistic Annotation* (pp. 297–314). Springer Science and Business Media B.V.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, *11*(3), 211–227.

Barsalou, L. W. (2021). Categories at the interface of cognition and action. In Caterina Mauri, Ilaria Fiorentini & Eugenio Goria (Eds.), *Building Categories in Interaction. Linguistic resources at work* (35–72). Amsterdam: John Benjamins Publishing house.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In Eric Wanner & Lila R. Gleitman (Eds), *Language Acquisition: the state of the art* (pp. 173–218). Cambridge: Cambridge University Press.

Baugh, A., & Cable, T. (1993). *A history of the English Language* (5th ed.). Taylor & Francis Group.

Beavers, J., & Koontz-Garboden, A. (2012). Manner and Result in the Roots of Verbal Meaning. *Linguistic Inquiry*, *43*(3), 331–369. https://doi.org/10.1162/LING_A_00093

Beavers, J., & Koontz-Garboden, A. (2020). *The Roots of Verbal Meaning*. Oxford: Oxford University Press.

Bedny, M., Caramazza, A., Pascual-Leone, A., & Saxe, R. (2012). Typical neural representations of action verbs develop without vision. *Cerebral Cortex, 22*(2), 286–293.

Bedny, M., Dravida, S., & Saxe, R. (2014). Shindigs, brunches, and rodeos: The neural basis of event words. *Cognitive, Affective and Behavioral Neuroscience*, *14*(3), 891–901.

Benedettini, V., Bertinetto, P. M., Lenci, A., & McRae, K. (2019). Multimodal Event Knowledge in Online Sentence Comprehension: The Influence of Visual Context on Anticipatory Eye Movements. *Proceedings of CogSci*, 1395–1401.

Berlin, B., & Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution.* University of California Press, Berkeley.

Biederman, I. (1972). Perceiving Real-World Scenes. *Science*, *177*(4043), 77–80.

Biederman, I., Rabinowitz, J. C., Glass, A. L., & Stacy, E. W. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, *103*(3), 597–600. https://doi.org/10.1037/H0037158

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *10*. https://doi.org/10.1088/1742-5468/2008/10/P10008

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*(3), 355–387. https://doi.org/10.1016/0010-0285(86)90004-6

Bock, J. K., Irwin, D. E., & Davidson, D. J. (2004). Putting first things first. In J. Henderson & F. Ferreira (Eds.), *The integration of language, vision, and action: Eye movements and the visual world* (pp. 249–278).

Bohnemeyer, J. (2007). *How to hammer a shirt apart (and talk about it)*.

Bohnemeyer, J., & Pedersen, E. (Eds.). (2011). *Event Representation in Language and Cognition Event. Cambridge University Press* (Vol. 53, Issue 9). Cambridge University Press.

Borghi, A. M., Parisi, D., & Di Ferdinando, A. (2005). Action and hierarchical levels of categories: A connectionist perspective. *Cognitive Systems Research*, *6*(2), 99–110. https://doi.org/10.1016/J.COGSYS.2004.06.003

Bouveret, M., & Sweetser, E. (2009). Multi-frame Semantics, Metaphoric Extensions and Grammar. *Annual Meeting of the Berkeley Linguistics Society*, *35*(1), 49–59. https://doi.org/10.3765/BLS.V35I1.3597

Bowerman, M. (2005). Why can't you 'open' a nut or 'break' a cooked noodle? Learning covert object categories in action word meanings. In L. Gershkoff-Stowe & D. Rakison (Eds.), *Building Object Categories in Developmental Time: 32nd Carnegie Symposium on Cognition.* (pp. 33–62). Lawrence Erlbaum.

Bunger, A., Skordos, D., Trueswell, J. C., & Papafragou, A. (2016). How children and adults encode causative events cross-linguistically: implications for language production and attention. *Language, Cognition and Neuroscience*, *31*(8), 1015–1037. https://doi.org/10.1080/23273798.2016.1175649

Bunger, A., Trueswell, J. C., & Papafragou, A. (2012). The relation between event apprehension and utterance formulation in children: Evidence from linguistic omissions. *Cognition*, *122*(2), 135–149. https://doi.org/10.1016/J.COGNITION.2011.10.002

Bybee, J. L. (1985). Moprhology. A Study of Relation between Meaning and Form. *Typological Studies in Language*, *9*.

Cacioli, C., & Vernillo, P. (2023). How did you break that? Semantic boundaries of Italian and English action verbs encoding breaking events. *Languages in Contrast*, *23*(1), 93–120. https://doi.org/10.1075/LIC.22004.CAC

Chafe, W. L. (1970). *Meaning and the Structure of Language*. University of Chicago Press.

Chaigneau, S. E., & Barsalou, L. W. (2008). The Role of Function in Categories. *Theoria et Historia Scientiarum*, *8*(1), 33–52. https://doi.org/10.12775/ths.2008.003

Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing Referential Domains during Real-Time Language Comprehension. *Journal of Memory and Language*, *47*(1), 30–49. https://doi.org/10.1006/JMLA.2001.2832

Comrie, B. (1976). *Aspect: an introduction to the study of verbal aspect and related problems*. Cambridge University Press.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*(1), 84–107. https://doi.org/10.1016/0010-0285(74)90005-X

Croft, W. (2012). *Verbs: Aspect and Causal Structure*. Oxford: Oxford University Press.

Croft, W. (2022). On two mathematical representations for "semantic maps." *Zeitschrift Fur Sprachwissenschaft*, *41*(1), 67–87.

Croft, W., & Cruse. D. Alan. (2004). Cognitive Linguistics. Cambridge: Cambridge University Press.

Csárdi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *International Journal of Complex Systems*. 1695 (5): 1--9

Cysouw, M. (2010). Semantic maps as metrics on meaning. *Linguistic Discovery*, *8*(1), 70–96.

Dabrowska, E. (2016). Cognitive Linguistics' seven deadly sins. *Cognitive Linguistics*, *27*(4), 479–491. https://doi.org/10.1515/COG-2016-0059/PDF

Davidson, D. (1967). The logical form of action sentences. In *The Logic of Decision and Action* (pp. 81–95).

Davidson, D. (1969). *The individuation of events*. In Nicholas Rescher (Ed.), Essays in Honor of Carl G. Hempel. Reidel (pp. 216-34). Dordrecht: D. Reidel

de Almeida, R. G., & Manouilidou, C. (2015). The Study of Verbs in Cognitive Science. In Roberto G. de Almeida & Christina Manouilidou (Eds), *Cognitive Science Perspectives on Verb Representation and Processing* (pp. 3-39). Springer.

Delorme, A., Richard, G., & Fabre-Thorpe, M. (2000). Ultra-rapid categorisation of natural scenes does not rely on colour cues: a study in monkeys and humans. *Vision Research*, *40*(16), 2187–2200. https://doi.org/10.1016/S0042-6989(00)00083-3

Devylder, S., & Zlatev, J. (2020). Cutting and breaking metaphors of the self and the Motivation & Sedimentation Model. In Annalisa Baicchi (Ed.), *Figurative Meaning Construction in Thought and Language* (pp. 254–281). John Benjamins Publishing Company. https://doi.org/10.1075/FTL.9.11DEV

Divjak, D., Levshina, N., & Klavan, J. (2016). Cognitive Linguistics: Looking back, looking forward. *Cognitive Linguistics*, *27*(4), 447–463. https://doi.org/10.1515/COG-2016-0095/PDF

Divjak, D., Milin, P., & Medimorec, S. (2020). Construal in language: A visual-world approach to the effects of linguistic alternations on event perception and conception. *Cognitive Linguistics*, *31*(1), 37–72.

Dobel, C., Gumnior, H., Bölte, J., & Zwitserlood, P. (2007). Describing scenes hardly seen. *Acta Psychologica*, *125*(2), 129–143. https://doi.org/10.1016/J.ACTPSY.2006.07.004

Dowty, D. (2001). The Semantic Assymmetry of 'Argument Alternations' (and Why it Matters). *Groninger Arbeiten Zur Germanistischen Linguistik N*, *44*, 171–186.

Dowty, D. R. (1977). Toward a Semantic Analysis of Verb Aspect and the English "Imperfective" Progressive. *Linguistics and Philosophy*, *1*(1), 45–77.

Dowty, D. R. (1979). *Word Meaning and Montague Grammar*. Kluwer Academic Publisher.

Edwards, D. (1991). Categories Are for Talking. *Theoretical Psychology*, *1*(4), 515–542. https://doi.org/10.1177/0959354391014007

Elman, J. L. (2009). On the Meaning of Words and Dinosaur Bones: Lexical Knowledge Without a Lexicon. *Cognitive Science*, *33*(4), 547–582. https://doi.org/10.1111/J.1551-6709.2009.01023.X

Enghels, R. (2018). Extending the notion of near-synonymy. *Languages in Contrast*, *18*(1), 1–5. https://doi.org/10.1075/LIC.00001.ENG

Erkelens, M. (2009). *Learning to categorise verbs and nouns: Studies on Dutch*. Netherlands Graduate School of Linguistics.

Essegbey, J. (2007). Cut and break verbs in Sranan. *Cognitive Linguistics*, *18*(2), 231–239.

Evans, V. (2019). *Cognitive Linguistics. A Complete Guide*. Edinburgh: Edinburgh University Press.

Fernández-Martínez, N. J., & Faber, P. (2019). Who stole what from whom? A corpus-based, cross-linguistic study of English and Spanish verbs of stealing. *Languages in Contrast*, *20*(1), 107–140. https://doi.org/10.1075/LIC.19002.FER

Fillmore, C. J. (1970). *The grammar of hitting and breaking. Readings in English Transformational Grammar.* 9–29.

Fillmore, C. J. (1985). Frames and the semantics of understanding. *Quaderni Di Semantica*, *6*(222–254).

Fillmore, C. J., & Lowe, J. B. (1998). The Berkeley FrameNet Project. *In COLING 1998 Volume 1: The 17th International Conference on Computational Linguistics.*

Flecken, M., & Stutterheim, C. (2014). Grammatical aspect influences motion event perception: findings from a cross- linguistic nonverbal recognition task. *Language and Cognition*, *6*(1), 45–78. https://doi.org/10.1017/langcog.2013.2

Flecken, M., Gerwien, J., Carrol, M., & Von Stutterheim, C. (2015). Analyzing gaze allocation during language planning: a cross-linguistic study on dynamic events. *Language and Cognition*, *7*(1), 138–166. https://doi.org/10.1017/langcog.2014.20

Flecken, M., Von Stutterheim, C., Carrol, M. (2014). Grammatical aspect influences motion event perception: Findings from a cross-linguistic non-verbal recognition task. *Language and Cognition,* 6 (1), 45–78. https://doi.org/10.1017/langcog.2013.2

Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: evidence from Pirahã language and cognition. *Cognition*, *108*(3), 819–824. https://doi.org/10.1016/J.COGNITION.2008.04.007

Frawly, W. (1992). *Linguistic Semantics*. Routledge - Taylor and Francis Group.

Fuji, S. I., Radetzky, P., & Sweetser, E. (2013). A Multi-Frame Semantics of Separation Verbs. In M. Borkent, B. Dancygier, & J. Hinnell (Eds.), *Language and the Creative Mind* (pp. 137–153). CSLI publications.

Ganushchak, L. Y., Konopka, A. E., & Chen, Y. (2014). What the eyes say about planning of focused referents during sentence formulation: A cross-linguistic investigation. *Frontiers in Psychology*, *5*, 1124.

Gärdenfors, P., Veloz, T., Witkowski, O., & Flores, D. (2021). Causal Reasoning and Event Cognition as Evolutionary Determinants of Language Structure. *Entropy 2021, Vol. 23, Page 843*, *23*(7), 843. https://doi.org/10.3390/E23070843

Gast, V. (2012). Contrastive Linguistics: Theories and Methods. In Bernd Kortmann & Johannes Kabatek (Eds.) *Dictionaries of linguistics and communication science: linguistic theory and methodology*, Berlin: Mouton de Gruyter.

Georgakopoulos, T., & Polis, S. (2018). The semantic map model: State of the art and future avenues for linguistic research. *Language and Linguistics Compass*, *12*(2), e12270. https://doi.org/10.1111/LNC3.12270

Georgakopoulos, T., & Härtl, H. (2022). Goal prevalence and situation types: An empirical analysis of differences in Greek and German motion event descriptions. In M. Georgiafentis, G. Giannoulopoulou, M. Koliopulu, & A. Tsokoglou (Eds.), *Contrastive Studies in Morphology and Syntax* (pp. 262–280). Bloomsbury Academic. https://doi.org/10.5040/9781350079212.0024

Georgakopoulos, T., & Polis, S. (2022). New avenues and challenges in semantic map research (with a case study in the semantic field of emotions). *Zeitschrift Fur Sprachwissenschaft*, *41*(1), 1–30. https://doi.org/10.1515/zfs-2021-2039

Georgakopoulos, T., Grossman, E., Nikolaev, D., & Polis, S. (2022). Universal and macro-areal patterns in the lexicon A case-study in the perception-cognition domain. *Linguistic Typology*, *26*(2), 439–487. https://doi.org/10.1515/lingty-2021-2088

Gerwien, J., & Flecken, M. (2016). First things first? Top-down influences on event apprehension. *Proceedings of the 38th Annual Meeting of the Cognitive Science Society (CogSci 2016)*, 2633–2638.

Glanemann, R., Zwitserlood, P., Bölte, J., & Dobel, C. (2016). Rapid apprehension of the coherence of action scenes. *Psychonomic Bulletin and Review*, *23*(5), 1566–1575. https://doi.org/10.3758/S13423-016-1004-Y/TABLES/3

Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, *57*(4), 544–569. https://doi.org/10.1016/J.JML.2007.01.007

Gleitman, L., & Papafragou, A. (2005). Language and thought. In K. Holyoak & B. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 633–661). Cambridge University Press.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*(3), 558–565. https://doi.org/10.3758/BF03196313

Goddard, C., & Wierzbicka, A. (2009). Contrastive semantics of physical activity verbs: "Cutting" and "chopping" in English, Polish, and Japanese. *Language Sciences*, *31*(1), 60–96. https://doi.org/10.1016/j.langsci.2007.10.002

Goldberg, A. E. (1993). Another Look at Some Learnability Paradoxes. In E. V. Clark (Ed.), *The Proceedings of the Twenty-fifth Annual Child Language Research Forum* (pp. 60–75). Center for the Study of Language and Information.

Goldberg, A. E. (2006). *Constructions at Work. The Nature of Generalization in Language*. Oxford: Oxford University Press.

Goldberg, A. E. (2010). Verbs, Frames and Constructions. In M. Rappaport Hovav, E. Doron, & I. Sichel (Eds.), *Syntax, Lexical Semantics and Event Structure* (pp. 39–58). Oxford: Oxford University Press.

Goldberg, A. E. (2019). *Explain Me This. Creativity, Competition and the Partial Productivity of Constructions*. Princeton: Princeton University Press.

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, *306*(5695), 496–499.

Goria, E. (2020). The discursive construction of categories. Categorisation as a dynamic and co-operative process. *Language Sciences*, *81*, 101233. https://doi.org/10.1016/J.LANGSCI.2019.06.003

Griffin, Z. M., & Bock, K. (2000). What the Eyes Say About Speaking. *Psychological Science*. *11*(4), 274–279. https://doi.org/10.1111/1467-9280.00255

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Affectedness and direct objects: The role of lexical semantics in the acquisition of verb argument structure. *Cognition*, *41*(1–3), 153–195. https://doi.org/10.1016/0010-0277(91)90035-3

Györi, G. (2013). Basic level categorisation and meaning in language. *Argumentum*, *9*, 149–161.

Győri, G. (2017). What happens to the basic level in language?. *Cognitive Linguistic Studies*, *4*(2), 171–193. https://doi.org/10.1075/COGLS.00001.GYO

Győri, G. (2019). Categorizing processes, motions, and actions: What is basic about basic level verbs?. *Argumentum*, *15*, 226–240.

Hale, K. L., & Keyser, J. (1987). A View from the Middle. In *Lexicon Project Working Papers*. Lexicon Project, Center for Cognitive Science, Massachusetts Institute of Technology.

Hamilton, A. F. D. C., & Grafton, S. T. (2006). Goal Representation in Human Anterior Intraparietal Sulcus. *Journal of Neuroscience*, *26*(4), 1133–1137. https://doi.org/10.1523/JNEUROSCI.4551-05.2006

Harnad, S. (1987). *Categorical Perception: The Groundwork of Cognition*. Cambridge University Press.

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, *42*(1–3), 335–346. https://doi.org/10.1016/0167-2789(90)90087-6

Harnad, S. (2017). To Cognize is to Categorise: Cognition is Categorisation. *Handbook of Categorisation in Cognitive Science*, 21–54. https://doi.org/10.1016/B978-0-08-101107-2.00002-6

Hassan, W. (2015). The interface between verbal semantics and clause structure in Arabic hitting and breaking verbs. *Białostockie Archiwum Językowe*, *15*, 181–198.

Henderson, J. M., & Hollingworth, A. (1998). Eye Movements During Scene Viewing: An Overview. In G. Underwood (Ed.), *Eye Guidance in Reading and Scene Perception* (pp. 269–293). https://doi.org/10.1016/B978-008043361-5/50013-4

Hintzman, D. L. (1986). "Schema Abstraction" in a Multiple-Trace Memory Model. *Psychological Review*, *93*(4), 411–428. https://doi.org/10.1037/0033-295X.93.4.411

Holmes, K. J., Moty, K., & Regier, T. (2017). Revisiting the role of language in spatial cognition: Categorical perception of spatial relations in English and Korean speakers. *Psychonomic Bulletin & Review*, *24*(6), 2031–2036. https://doi.org/10.3758/S13423-017-1268-X

Holmqvist, K., Nyström, M., Andersson, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye Tracking A Comprehensive Guide to Methods and Measures*. Oxford: Oxford University Press.

Holsanova, J. (2014). Reception of multimodality: Applying eye-tracking methodology in multimodal research. In Carey Jewitt (Ed.), *Routledge Handbook of Multimodal Analysis.* (pp. 285–296). https://doi.org/10.13140/2.1.3790.3041

Hsiao, H. S. (2015). The Role of Force in Mandarin Verbs of Cutting. *Taiwan Journal of Linguistics*, *13*(2), 1–30. https://doi.org/10.6519/TJL.2015.13(2).1

Huisman, J. L. A., Van Hout, R., & Majid, A. (2021). Cross-linguistic constraints and lineage-specific developments in the semantics of cutting and breaking in Japonic and Germanic. *Linguistic Typology*, *27*(1), 1–35. https://doi.org/10.1515/lingty-2021-2090

Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain. *Neuron*, *76*(6), 1210–1224. https://doi.org/10.1016/J.NEURON.2012.10.014

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: The MIT Press.

Jakubíček, M., Kilgarriff, A., Kovář, V., & Rychlý, P. (2013). The TenTen Corpus Family. *7th International Corpus Linguistics Conference*. http://www.sketchengine.co.uk

Johnson, M. (1987). *The body in the mind: the bodily basis of meaning, imagination, and reason*. University of Chicago Press.

Just, A. M., & Carpenter, P. A. (1984). Using Eye Fixations to Study Reading Comprehension. In David E. Kieras & Marcel A. Just (Eds.), *New Methods in Reading Comprehension Research* (pp. 151–182). Routledge. https://doi.org/10.4324/9780429505379-8

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of Syntactic and Semantic Information in Predictive Processing: Cross-Linguistic Evidence from German and English. *Journal of Psycholinguistic Research*, *32*(1).

Katz, J. J., & Postal, P. M. (1964). *And Integrated Theory of Linguistic Description*. Cambridge, MA: The MIT Press.

Knoeferle, P., & Crocker, M. W. (2006). The Coordinated Interplay of Scene, Utterance, and World Knowledge: Evidence from Eye Tracking. *Cognitive Science*, *30*(3), 481–529. https://doi.org/10.1207/S15516709COG0000_65

König, E. (2012). Contrastive linguistics and language comparison. *Languages in Contrast*, *12*(1), 3–26. https://doi.org/10.1075/lic.12.1.02kon

Konopka, A. E. (2019). Encoding actions and verbs: Tracking the time-course of relational encoding during message and sentence formulation. *Journal of Experimental Psychology: Learning Memory and Cognition*, *45*(8), 1486–1510. https://doi.org/10.1037/xlm0000650

Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology*, *73*, 1–40. https://doi.org/10.1016/j.cogpsych.2014.04.001

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, *12*(2), 72–79. https://doi.org/10.1016/J.TICS.2007.11.004

Kwon, I. (2016). How do Koreans break and cut things? A cognitive-semantics approach to break predicates and CUT predicates in Korean. *Linguistic Research*, *33*(1), 65–94. https://doi.org/10.17250/khisli.33.1.201603.004

Labov, W. (1973). The boundaries of words and their meanings. In J. N. Bailey & R. W. Shuy (Eds.), *New ways of analyzing variation in English*. Georgetown University Press

Lakoff, G. (1987). *Women, Fire, and Dangerous things*. University of Chicago Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.

Lalljee, M., Lamb, R., & Abelson, R. P. (1992). The Role of Event Prototypes in Categorisation and Explanation. *European Review of Social Psychology*, *3*(1), 153–182. https://doi.org/10.1080/14792779243000050

Langacker, R. (1987). *Foundations of Cognitive Grammar* (Vol. 1). Stanford: Standford University Press.

Langacker, R. W. (1991). *Foundations of Cognitive Grammar. Descriptive application*. (Vol. 2). Stanford: Stanford University Press.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, *25*(1), 1–18. https://doi.org/10.18637/JSS.V025.I01

Leshinskaya, A., Wurm, M. F., & Caramazza, A. (2020). Concepts of Actions and their Objects. In M. Gazzaniga (Ed.) *The Cognitive Neurosciences* (pp. 757-765). The MIT Press.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. The MIT Press. https://doi.org/10.7551/MITPRESS/6393.001.0001

Levin, B. (1993). *English Verb Classes and Alternation*. The University of Chicago Press. https://doi.org/10.1117/12.789676

Levin, B. (to appear). Pounding up the stairs and slamming into the wall: Hitting verbs in directed motion event descriptions. In *Neglected Aspects of Motion-Event Description: Deconstructing Motion Events* (pp. 1–27).

Levin, B., & Hovav, M. R. (2013). Lexicalised Meaning and Manner/Result Complementarity. *Studies in Linguistics and Philosophy*, *93*, 49–70. https://doi.org/10.1007/978-94-007-5983-1_3

Levin, B., & Rappaport Hovav, M. (2005). *Argument Realization*. Cambridge University Press.

Levinson, S. C., Kita, S., Haun, D. B. M., & Rasch, B. H. (2002). Returning the tables: language affects spatial reasoning. *Cognition*, *84*(2), 155–188. https://doi.org/10.1016/S0010-0277(02)00045-8

Levinson, S. C. (2003). *Space in Language and Cognition*. Cambridge: Cambridge University Press.

Levinson, S. C. (2003). Language and mind: Let's get the issues straight!. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 25–46). Boston Review.

Levshina, N. (2015). *How to do linguistics with R*. Amsterdam: John Benjamins Publishing Company.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(4), 565–572. https://doi.org/10.1037/0096-1523.4.4.565

Majid, A., Boster, J. S., & Bowerman, M. (2008). The cross-linguistic categorisation of everyday events: A study of cutting and breaking. *Cognition*, *109*(2), 235–250. https://doi.org/10.1016/j.cognition.2008.08.009

Malt, B. C., Ameel, E., Imai, M., Gennari, S. P., Saji, N., & Majid, A. (2014). Human locomotion in languages: Constraints on moving and meaning. *Journal of Memory and Language*, *74*, 107–123. https://doi.org/10.1016/J.JML.2013.08.003

Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus Naming: Similarity and the Linguistic Categorisation of Artefacts. *Journal of Memory and Language*, *40*(2), 230–262. https://doi.org/10.1006/JMLA.1998.2593

Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory and Cognition*, *32*(8), 1389–1400. https://doi.org/10.3758/BF03206329/METRICS

Matsumoto, Y. (1996). How Abstract is Subjective Motion? A Comparison of Coverage Path Expressions and Access Path Expressions. In Adele Goldberg (Ed.), *Conceptual Structure, Discourse, and Language.* (pp. 359–373). CSLI Publications.

Mauri, C., & Sansò, A. (2018). Linguistic strategies for ad hoc categorisation: Theoretical assessment and cross-linguistic variation. *Folia Linguistica Historica*, *52*(s39-1), 1–35. https://doi.org/10.1515/FLIH-2018-0001/MACHINEREADABLECITATION/RIS

McKoon, G., & Love, J. (2011). Verbs in the lexicon: Why is hitting easier than breaking? *Language and Cognition*, *3*(2), 313–330. https://doi.org/10.1515/LANGCOG.2011.011

Mereu, L., & Piunno, V. (2019). The argument structure of verbs of hitting and breaking in Italian. *Lingue e Linguaggio*, *18*(1), 143–176. https://doi.org/10.1418/95389

Mervis, C. B., & Rosch, E. (1981). Catergorization of Natural Objects. *Annual Review of Psycholgy*, *32*, 89–115.

Moneglia, M. (1998). Determinazione empirica del senso. *Studi Di Grammatica Italiana*, *XVII*.

Moneglia, M. (1999). Prototypical vs. non-prototypical predicates: ways of understanding and the semantic partition of lexical meaning. *International Conference" Linguistics at the End of the Century" Moscow State University February*.

Moneglia, M., Brown, S. W., Francesca, F., Gagliardi, G., Fahad, K., Monica, M., & Panunzi, A. (2014). The IMAGACT Visual Ontology. An Extendable Multilingual Infrastructure for the Representation of Lexical Encoding of Action. *Proceedings of the 9th International Conference on Language Resources and Evaluation, LREC 2014*, 3425–3432.

Moneglia, M., Panunzi, A., & Gregori, L. (2018). Action Identification and Local Equivalence of Action Verbs: the Annotation Framework of the IMAGACT Ontology. In James Pustejovsky & Ielka van der Sluis (Eds.), *Proceedings of the LREC 2018 Workshop "AREA – Annotation, Recognition and Evaluation of Actions* (pp. 23–30). ELDA.

Morris, M. W., & Murphy, G. L. (1990). Converging operations on a basic level in event taxonomies. *Memory & Cognition*, *18*(4), 407–418. https://doi.org/10.3758/BF03197129/METRICS

Möttonen, T. (2016). *Construal in expression: Intersubjective approach to cognitive grammar.* University of Helsinki.

Murphy, G. (2002). The Big Book of Concepts. The MIT Press. https://doi.org/10.7551/MITPRESS/1602.001.0001

Murphy, G. L., & Smith, E. E. (1982). Basic-level superiority in picture categorisation. *Journal of Verbal Learning and Verbal Behavior*, *21*(1), 1–20. https://doi.org/10.1016/S0022-5371(82)90412-1

Norcliffe, E., & Konopka, A. E. (2015). Vision and language in cross-linguistic research on sentence production. *Attention and Vision in Language Processing*, 77–96. https://doi.org/10.1007/978-81-322-2443-3_5/COVER

Norenzayan, A., Choi, I., & Peng, K. (2007). Perception and cognition. In Harold L. & Miller Jr. (Eds.), *The SAGE Encyclopedia of Theory in Psychology*. https://doi.org/10.4135/9781483346274.N227

Nosofsky, R. M. (1986). Attention, Similarity, and the Identification-Categorisation Relationship. *Journal of Experimental Psychology: General*, *115*(1), 39–57. https://doi.org/10.1037/0096-3445.115.1.39

Oliva, A., Torralba, A., Castelhano, M. S., & Henderson, J. M. (2003). Top-down control of visual attention in object detection. *IEEE International Conference on Image Processing*, *1*, 253–256. https://doi.org/10.1109/ICIP.2003.1246946

Panunzi, A., De Felice, I., Gregori, L., Jacoviello, S., Monachini, M., Moneglia, M., Quochi, V., & Russo, I. (2014). Translating action verbs using a dictionary of images: the IMAGACT

ontology. In *Proceedings of the XVI EURALEX International Congress: The User in Focus* (pp. 1163-1170). Bolzano: EURAC research.

Panzner, M., Gaspers, J., & Cimiano, P. (2019). Modeling the Co-Emergence of Linguistic Constructions and Action Concepts: The Case of Action Verbs. *IEEE Transactions on Cognitive and Developmental Systems*, *11*(3), 435–449. https://doi.org/10.1109/TCDS.2019.2900418

Paradis, C. (2004). Where Does Metonymy Stop? Senses, Facets, and Active Zones. *Metaphor and Symbol*, *19*(4), 245–264. https://doi.org/10.1207/s15327868ms1904_1

Paradis, C. (2005). Ontologies and construals in lexical semantics. *Axiomathes*, *15*(4), 541–573. https://doi.org/10.1007/s10516-004-7680-7

Paradis, C. (2015). Meanings of words: Theory and application. In Ulrike Haß & Petra Storjohann (Eds.), *Handbuch Wort und Wortschatz* (pp. 274–293). https://doi.org/10.1515/9783110296013-012

Pederson, E., Danziger, E., Wilkins, D., Levinson, S., Kita, S., & Senft, G. (1998). Semantic Typology and Spatial Conceptualisation. *Language*, *74*(3), 557. https://doi.org/10.2307/417793

Peelen, M. V., Romagno, D., & Caramazza, A. (2012). Independent representations of verbs and actions in left lateral temporal cortex. *Journal of Cognitive Neuroscience*, *24*(10), 2096–2107. https://doi.org/10.1162/JOCN_A_00257

Preissl, H., Pulvermüller, F., Lutzenberger, W., & Birbaumer, N. (1995). Evoked potentials distinguish between nouns and verbs. *Neuroscience Letters*, *197*(1), 81–83. https://doi.org/10.1016/0304-3940(95)11892-Z

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews. Neuroscience*, *6*(7), 576–582. https://doi.org/10.1038/NRN1706

Pulvermüller, F., Lutzenberger, W., & Preissl, H. (1999). Nouns and Verbs in the Intact Brain: Evidence from Event-related Potentials and High-frequency Cortical Responses. *Cerebral Cortex*, *9*(5), 497–506. https://doi.org/10.1093/CERCOR/9.5.497

Pulvermüller, F., Shtyrov, Y., & Hauk, O. (2009). Understanding in an instant: neurophysiological evidence for mechanistic language circuits in the brain. *Brain and Language*, *110*(2), 81–94. https://doi.org/10.1016/J.BANDL.2008.12.001

Pye, C. (1996). K'iche' Maya Verbs of Breaking and Cutting. *Kansas Working Papers in Linguistics*, *21*, 87–98.

Pye, C., Loeb, D. F., & Pao, Y. Y. (1996). The acquisition of cutting and breaking. In E. V. Clark (Ed.), *The Proceedings of the Twenty- seventh Annual Child Language Research Forum* (pp. 227–236). Stanford: Center for the Study of Language and Information.

Radden, G., & Dirven, R. (2007). *Cognitive English Grammar*. Amsterdam: John Benjamins Publishing Company.

Radvansky, G. A., & Zacks, J. M. (2011). Event perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(6), 608–620. https://doi.org/10.1002/WCS.133

Rappaport Hovav, M., & Levin, B. (2010). Reflections on manner/result complementarity. In E. Doron, M. Rappaport Hovav, & I. Sichel (Eds.), *Syntax, Lexical Semantics, and Event Structure* (pp. 21–38). Oxford: Oxford University Press.

Ratneshwar, S., Barsalou, L. W., Pechmann, C., & Moore, M. (2001). Goal-Derived Categories: The Role of Personal and Situational Goals in Category Representations. *Journal of Consumer Psychology*, *10*(3), 147–157. https://doi.org/10.1207/S15327663JCP1003_3

Rayner, K. (2009). Eye Movements in Reading: Models and Data. *Journal of Eye Movement Research*, *2*(5), 1. https://doi.org/10.16910/jemr.2.5.2

Rayner, K., & Reingold, E. M. (2015). Evidence for direct cognitive control of fixation durations during reading. *Current Opinion in Behavioral Sciences*, *1*, 107–112.

Rifkin, A. (1985). Evidence for a basic level in event taxonomies. *Memory & Cognition*, *13*(6), 538–556. https://doi.org/10.3758/BF03198325/METRICS

Romain, L. (2021). Putting the argument back into argument structure constructions. *Cognitive Linguistics*. 33(1), 35-64. https://doi.org/10.1515/COG-2021-0021

Rosch, E. (1978). Principles of categorisation. In E. Rosch & B. B. Loyd (Eds.), *Cognition and Categorisation* (pp. 27–48). Hillsdale: Lawrence Erlbaum.

Rosch, E., & Lloyd, B. L. (1978). *Cognition and Categorisation*. John Wiley & Sons Inc.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605. https://doi.org/10.1016/0010-0285(75)90024-9

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382–439. https://doi.org/10.1016/0010-0285(76)90013-X

Sakarias, M., & Flecken, M. (2019). Keeping the Result in Sight and Mind: General Cognitive Principles and Language-Specific Influences in the Perception and Memory of Resultative Events. *Cognitive Science*, *43*(1), e12708. https://doi.org/10.1111/COGS.12708

Schulze, R. (2009). Categorisation in linguistics. *Zeitschrift Fur Anglistik Und Amerikanistik*, *57*(3), 219–231. https://doi.org/10.1515/ZAA.2009.57.3.219

Shapiro, K. A., Moo, L. R., & Caramazza, A. (2006). Cortical signatures of noun and verb production. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(5), 1644–1649. https://doi.org/10.1073/PNAS.0504142103

Slobin, D. I. (1996). From "thought and language" to "thinking for speaking." In J. Gumperz & S. C. Levinson (Eds.), *Rethinking Linguistic Relativity* (pp. 70–96). Cambrdige: Cambridge University Press.

Slobin, D. I., Ibarretxe-Antunanõ, I., Kopecka, A., & Majid, A. (2014). Manners of human gait: A crosslinguistic event-naming study. *Cognitive Linguistics*, *25*(4), 701–741. https://doi.org/10.1515/COG-2014-0061/MACHINEREADABLECITATION/RIS

Smith, L. B., & Samuelson, L. K. (1997). Perceiving and remembering: Category stability, variability and development. In K. Lamberts & D. R. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 161–195). Cambridge, MA: The MIT Press. https://psycnet.apa.org/record/1997-36563-005

Spalek, A. A. (2015). Spanish change of state verbs in composition with atypical theme arguments: Clarifying the meaning shifts. *Lingua*, *157*, 36–53. https://doi.org/10.1016/J.LINGUA.2014.12.007

Stathi, K. (2023). *Granularity in the verbalization of events and objects: a cross-linguistic study*. Amsterdam: John Benjamins Publishing Company

Staub, A., Abbott, M., & Bogartz, R. S. (2012). Linguistically guided anticipatory eye movements in scene viewing. *Visual Cognition*, *20*(8), 922–946. https://doi.org/10.1080/13506285.2012.715599

Suchomel, V., & Pomikálek, J. (2012). Efficient Web Crawling for large Text Corpora. *Proc Seventh Web as Corpus Workshop (WAC7)*, 39–43.

Takashima, A., Konopka, A., Meyer, A., Hagoort, P., & Weber, K. (2020). Speaking in the Brain: The Interaction between Words and Syntax in Sentence Production. *Journal of Cognitive Neuroscience*, *32*(8), 1466-1483.

Talmy, L. (1985). *Lexicalisation patterns: semantic structure in lexical forms*. Cambridge: Cambrdige University Press.

Talmy, L. (2000). *Towards a Cognitive Semantics*. The MIT Press.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634. https://doi.org/10.1126/SCIENCE.7777863

Taylor, J. R. (1989). *Linguistic Categorisation* (Second Edition). Oxford: Oxford University Press.

Tversky, B., & Hemenway, K. (1983). Categories of environmental scenes. *Cognitive Psychology*, *15*(1), 121–149. https://doi.org/10.1016/0010-0285(83)90006-3

Ünal, E., Richards, C., Trueswell, J. C., & Papafragou, A. (2021). Representing agents, patients, goals and instruments in causative events: A cross-linguistic investigation of early language and cognition. *Developmental Science*, *24*(6), e13116. https://doi.org/10.1111/DESC.13116

Underwood, G., & Everatt, J. (1992). The Role of Eye Movements in Reading: Some Limitations of the Eye-Mind Assumption. *Advances in Psychology*, *88*(C), 111–169. https://doi.org/10.1016/S0166-4115(08)61744-6

Van De Velde, M., Meyer, A. S., & Konopka, A. E. (2013). Message formulation and structural assembly: Describing 'easy' and 'hard' events with preferred and dispreferred syntactic structures. *Journal of Memory and Language*, *71*, 124–144. https://doi.org/10.1016/j.jml.2013.11.001

Van Der Auwera, J. (2008). In defense of classical semantic maps. *Theoretical Linguistics*, *34*(1), 39–46. https://doi.org/10.1515/THLI.2008.002/MACHINEREADABLECITATION/RIS

Van Gelderen, E. (2012). *A history of the English Language* (Revised). Amsterdam: John Benjamins Publishing Company.

Van Valin, R. D., & LaPolla, R. J. (n.d.). *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge University Press.

Vendler, Z. (1957). Verbs and Times. *The Philosophical Review*, *66*(2), 143. https://doi.org/10.2307/2182371

Vendler, Z. (1967). *Linguistics in Philosophy*. Cornell University Press.

Ty Verhagen, A. (2007). Construal and Perspectivization. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 48–81). Oxford: Oxford University Press.

Verhagen, A. (2010). Usage, structure, scientific explanation, and the role of abstraction, by linguists and by language users. In Kasper Boye & Elisabeth Engberg-Pedersen (Eds), *Language Usage and Language Structure* (pp. 45–72). https://doi.org/10.1515/9783110219180.1.45/HTML

Verkuyl, H. J. (1972). On the Compositional Nature of the Aspects. Springer Dordrecht. https://doi.org/10.1007/978-94-017-2478-4

Viberg, Å. (1985). Hel och trasig. En skiss av några verbala semantiska fält i svenskan. In *Svenskans beskrivning 15* (pp. 529–554). Göteborg: Göteborgs universitet.

Viberg, Å. (2007). Whole and broken. Breaking and cutting in Swedish from a crosslinguistic perspective. *In Elisabeth Ahlsén et Al. (Eds.), Communication – Action – Meaning. A Festschrift to Jens Allwood, 17–42. Gothenburg University: Department of Linguistics.*

Viberg, Å. (2013). Posture verbs: A multilingual contrastive study. *Languages in Contrast*, *13*(2), 139–169. https://doi.org/10.1075/LIC.13.2.02VIB

Viberg, Å. (2020). Contrasting semantic fields across languages. In Renata Enghels , Bart Defrancq & Marlies Jansegers (Eds.), *New Approaches to Contrastive Linguistics: Empirical and Methodological Challenges* (pp. 265–312). Berlin, Boston: Mouton De Gruyter.

Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, *48*(4), 422–488. https://doi.org/10.1016/J.COGPSYCH.2003.09.001

Warglien, M., Gärdenfors, P., & Westera, M. (2012). Event structure, conceptual spaces and the semantics of verbs. *Theoretical Linguistics*, *38*(3–4), 159–193. https://doi.org/10.1515/tl-2012-0010

Watson, C. E., Cardillo, E. R., Ianni, G. R., & Chatterjee, A. (2013). Action Concepts in the Brain: An Activation Likelihood Estimation Meta-analysis. *Journal of Cognitive Neuroscience*, *25*(8), 1191–1205. https://doi.org/10.1162/JOCN_A_00401

Wheeldon, L. R., & Konopka, A. (2023). Grammatical Encoding for Speech Production. *Elements in Psycholinguistics*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009264518

Wierzbicka, A. (1988). *The Semantics of Grammar*. Amsterdam: John Benjamins Publishing Company.

Winskel, G. (1982). Event structure semantics for CCS and related languages. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *140 LNCS*, 561–576. https://doi.org/10.1007/BFB0012800/COVER

Wittgenstein, L. (1953). *Philosophical Investigations. [Philosophische Untersuchungen]*. Macmillan.

Wurm, M., Caramazza, A. & Lingnau, A. (2017). Action Categories in Lateral Occipitotemporal Cortex Are Organized Along Sociality and Transitivity. *Journal of Neuroscience, 37(3),* 562-575. https://doi.org/10.1523/JNEUROSCI.1717-16.2016

Yates, T. S., Sherman, B. E., & Yousif, S. R. (2023). More than a moment: What does it mean to call something an "event"? *Psychonomic Bulletin & Review*, *30*(6), 2067–2082. https://doi.org/10.3758/S13423-023-02311-4

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3–21. https://doi.org/10.1037/0033-2909.127.1.3

Zhuang, T., & Lingnau, A. (2022). The characterization of actions at the superordinate, basic and subordinate level. *Psychological Research*, *86*(6), 1871–1891. https://doi.org/10.1007/s00426-021-01624-0

Zlatev, J., Jacobsson, G., & Paju, L. (2021). Desiderata for metaphor theory, the Motivation & Sedimentation Model and motion-emotion metaphoremes. In A. Soares da Silva (Ed.), *Figurative Language – Intersubjectivity and Usage* (pp. 41–73). Amsterdam: John Benjamins Publishing Company.