SCUOLA
NORMALE
SUPERIORE

Classe di Lettere e Filosofia
Corso di perfezionamento (Ph.D.) in Filosofia
XXXIII ciclo

Implicit indefinite objects at the syntax-semantics-pragmatics interface: a probabilistic model of acceptability judgments

Settore Scientifico-Disciplinare: L-LIN/01 Glottologia e linguistica

Candidata
dott.ssa Giulia Cappelli

Relatori
prof. Pier Marco Bertinetto
prof. Alessandro Lenci

A.A. 2021/2022

Colophon
The data, code, and materials supporting the findings of this dissertation are available within the thesis (and its supplementary materials) or in GitHub repositories referenced in the text. This document was typeset with the help of KOMA-Script and LaTeX using the kaobook class by Federico Marotta.

First version
Submitted in September 2022 in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Ph.D.) in Linguistics.

All models are wrong, but some are useful.

– George E. P. Box

Abstract

Optionally transitive verbs, whose Patient participant is semantically obligatory but syntactically optional (e.g., *to eat, to drink, to write*), deviate from the transitive prototype defined by Hopper and Thompson (1980). Following Fillmore (1986), unexpressed objects may be either indefinite (referring to prototypical Patients of a verb, whose actual entity is unknown or irrelevant) or definite (with a referent available in the immediate intra- or extra-linguistic context). This thesis centered on *indefinite* null objects, which the literature argues to be a gradient, non-categorical phenomenon possible with virtually any transitive verb (in different degrees depending on the verb semantics), favored or hindered by several semantic, aspectual, pragmatic, and discourse factors. In particular, the probabilistic model of the grammaticality of indefinite null objects hereby discussed takes into account a continuous factor (semantic selectivity, as a proxy to object recoverability) and four binary factors (telicity, perfectivity, iterativity, and manner specification).

This work was inspired by Medina (2007), who modeled the effect of three predictors (semantic selectivity, telicity, and perfectivity) on the grammaticality of indefinite null objects (as gauged via Likert-scale acceptability judgments elicited from native speakers of English) within the framework of Stochastic Optimality Theory. In her variant of the framework, the constraints get floating rankings based on the input verb's semantic selectivity, which she modeled via the Selectional Preference Strength measure by Resnik (1993, 1996). I expanded Medina's model by modeling implicit indefinite objects in two languages (English and Italian), by using three different measures of semantic selectivity (Resnik's SPS; Behavioral PISA, inspired by Medina's Object Similarity measure; and Computational PISA, a novel similarity-based measure by Cappelli and Lenci (2020) based on distributional semantics), and by adding iterativity and manner specification as new predictors in the model.

Both the English and the Italian five-predictor models based on Behavioral PISA explain almost half of the variance in the data, improving on the Medina-like three-predictor models based on Resnik's SPS. Moreover, they have a comparable range of predicted object-dropping probabilities (30-100% in English, 30-90% in Italian), and the predictors perform consistently with theoretical literature on object drop. Indeed, in both models, atelic imperfective iterative manner-specified inputs are the most likely to drop their object (between 80% and 90%), while telic perfective non-iterative manner-unspecified inputs are the least likely (between 30% and 40%). The constraint re-ranking probabilities are always directly proportional to semantic selectivity, with the exception of Telic End in Italian. Both models show a main effect of telicity, but the second most relevant factor in the model is perfectivity in English and manner specification in Italian.

Contents

List of Figures

List of Tables

Introduction 1

1.1 Overview

1.1.1 Relevance of this thesis

This thesis is about the omission of direct objects from predicates headed by verbs with two semantic participants, i.e., an Agent (in the syntactic subject position) and a Patient (in the syntactic object position). These "optionally transitive" verbs, deviating from the transitive prototype defined by Hopper and Thompson (1980), appear in a wide variety of contexts cross-linguistically, and are licensed by different semantic, aspectual, pragmatic, and discourse factors. Within this broad area of interest, I will focus on *indefinite* null objects, corresponding to what Fillmore (1986) called "indefinite null complements". These omitted objects, as shown in (1-a), refer to something that is "unknown or a matter of indifference" (Fillmore 1986, p. 96). Indeed, what matters in the example is that Giulia is writing something, and in particular, something that is usually written. The actual product of the writing event, be it a novel, a shopping list, or a doctoral dissertation, is irrelevant. On the contrary, the referent of *definite* null objects (which I am not studying in this thesis) "must be retrieved from something given in the context", as in (1-b). In this case, the context is provided in the first sentence in the example, where a reference is made to grad school. Thus, the omitted object in the second sentence can be understood to refer to a doctoral thesis (the thing one wants to defend soon when in grad school) rather than, say, the title of Olympic champion or the national borders.

(1) a. Giulia is writing $\varnothing_{\text{dObj}}$.
b. Grad school is hard. Giulia hopes to defend $\varnothing_{\text{dObj}}$ soon.

The available literature suggests indefinite object drop to be possible with different types of verbs to varying extents (for instance, change-of-state verbs such as *to kill* are much more resistant to object drop than incremental-theme verbs such as *to eat*), and for any given verb, to be more likely under some specific semantic, aspectual, and pragmatic circumstances. For instance, a direct object can only participate in the implicit indefinite object construction if it is recoverable from the meaning of the verb itself, and transitive verbs are much more likely to be used without a direct object when they are used in imperfective or iterative contexts than in perfective, single-occurrence contexts.

While many pages have been written about the role of several linguistic factors in facilitating or blocking indefinite object drop, as I will detail in the first section of this thesis, way fewer attempts have been made to understand the nature of this phenomenon via experimental means, modeling the joint effect of several predictors of object drop. Medina (2007) made a substantial step in this direction in her (linear) Stochastic

Optimality Theoretic model of indefinite object drop in English, taking into consideration the joint effect of object recoverability, telicity, and perfectivity on the grammaticality of indefinite null objects (as gauged via gradient acceptability judgments elicited from native speakers) occurring with 30 transitive verbs. This model shows that:

- indefinite object drop is a gradient, non-categorical phenomenon;
- it is possible with virtually any transitive verb, but in different degrees depending on the verb semantics;
- for any given verb, different aspectual features may favor or hinder object drop.

In the experiments I will perform to study the implicit indefinite object construction, I inherit Medina's methodology and employ the same variant of Stochastic Optimality Theory she devised, with several updates I will illustrate in Section 1.1.2 and, in more detail, in the experimental section of this thesis.

1.1.2 Main goals and elements of novelty

This thesis is meant as an expansion upon the original model of indefinite object drop designed by Medina (2007). I will collect acceptability judgments following her same experimental design and model the data thus collected within the bounds of the same framework (her variant of Stochastic Optimality Theory). In doing so, I add several elements of novelty to the study:

- I will model implicit indefinite objects both in English (like Medina did) and in Italian (which is included in such a probabilistic model for the first time), analyzing language-specific differences in the way several factors facilitate object drop;
- in addition to using Resnik's (1993) Selectional Preference Strength measure to quantify semantic selectivity (as a proxy to object recoverability), following Medina (2007), I will also define a novel computational measure based on distributional semantics (Computational PISA, presented in Cappelli and Lenci (2020)) and a behavioral measure (Behavioral PISA) meant to improve on Medina's Object Similarity;
- in addition to the three predictors included by Medina in her model (semantic selectivity, telicity, and perfectivity), I will also add iterativity and manner specification as predictors to find out how they affect indefinite object drop and whether a more complex, five-predictor model actually provides a more accurate view on this phenomenon than the original three-predictor model;
- in order to make it easier for future research to build on my results (possibly applying my method to other languages, or to the same languages with different predictors) or to replicate them, I intend to share my materials and document my methods (as well as my Python scripts), as detailed in Section 1.2.2.

Assuming that my probabilistic models of the gradient grammaticality of indefinite object drop are solid, this thesis will be an additional cobblestone on the well-trodden road of studies about implicit indefinite objects, omitted arguments, and transitivity-related phenomena. More in general, it will add to the understanding of the ways several linguistic factors give rise to phenomena that, despite appearing to happen arbitrarily on a lexically-determined basis, are quite systematic in their behavior. Thus, looking at this thesis from a much broader perspective, it can also be argued to be a contribution to research about the systematicity (i.e., rule-abiding behavior) of human language and cognition, and about the interaction of semantic, aspectual, world-knowledge, pragmatic, and discourse factors in determining the way we translate our communicative intentions into syntactically well-formed utterances.

1.2 Contents within and without

1.2.1 Chapters of the thesis and their structure

This thesis is divided into two main parts, one devoted to the review of the literature on object drop, Optimality Theory, and gradient models of indefinite null objects (from Chapter 2 to Chapter 5), and another devoted to my own experiments and the results thereof (from Chapter 6 to Chapter 9). Let us consider each Chapter in more detail.

Theory and literature review In Chapter 2 (*Indefinite object drop*) I will define the *indefinite* implicit object construction as a deviation from the transitive prototype (see Section 2.1) and in contrast with *definite* object drop (see Section 2.2), based on the literature. In Section 2.3 I will argue that there is virtually no reason why a transitive verb should not be able to participate in the implicit indefinite object construction (provided favorable aspectual, semantic, and discourse conditions), and that the implicit object is understood to be the prototypical Patient for a given sense of a given verb. In Section 2.4 I will argue that optionally transitive verbs should only be taken to have a single entry in the lexicon, realized syntactically either with an overt or with an implicit object, rather than having two separate lexical entries. In Section 2.5 I will provide the perspective on indefinite object drop I adopt in my experiments and throughout this thesis.
In Chapter 3 (*Factors allowing indefinite object drop*) I will discuss the role played by semantic factors (recoverability, Agent affectedness, and manner specification, in Section 3.1), aspectual factors (telicity and perfectivity, in Section 3.2), and pragmatic factors (routine[1] , iterativity, habituality, and discourse factors, in Section 3.3) in facilitating or blocking object drop with optionally transitive verbs. After some considerations in Section 3.4 on the reasons why corpus frequency is not included among the relevant factors, I conclude the Chapter in Section 3.5 with the reasoning behind my choice of predictors of object drop to be used in the experimental section of this thesis.
Chapter 4 (*Towards a Stochastic Optimality Theoretic account of indefinite object drop*) will explain the main tenets of Optimality Theory relative to syntax (in Section 4.1), the limits of standard Optimality Theory as a model of the implicit indefinite object construction, and, therefore, why it is advisable to resort to probabilistic models of grammar that are able to account for the gradient grammaticality shown by indefinite object drop, such as Stochastic Optimality Theory (as argued in Section 4.2).
In particular, in this thesis I will adopt the variant of Stochastic Optimality

1: In the sense intended by Glass (2013, 2020, 2022).

Theory specifically designed by Medina (2007) to model indefinite object drop, which I describe and discuss in Chapter 5 (*Medina's (2007) model*). The contents of the input and the output will be presented in Section 5.1. In Section 5.2 I will discuss the implementation of the three predictors the author used in her model (semantic selectivity, telicity, and perfectivity). The probabilistic ranking of the constraints, which I will introduce in Section 5.3, will be defined in a top-down perspective (from constraint ranking as a function of semantic selectivity to object-drop probability as gradient grammaticality) in Section 5.4, and finally implemented in a bottom-up perspective (from the acceptability judgments to the estimation of parameters of the linear functions) in Section 5.5.

Experiments and results Chapter 6 (*Linguistic factors used as predictors*) opens the experimental part of this thesis. I will present five facilitating factors (a continuous factor and four binary factors) of object drop I will use as predictors in my Stochastic Optimality Theoretic model, picked among the ones introduced in Chapter 3. The continuous factor is object recoverability, which I will model via three different measures of semantic selectivity described in Section 6.1, namely Resnik's (1993) Selectional Preference Strength (following Medina (2007)), Computational PISA (a novel measure based on distributional semantics I contributed to define in Cappelli and Lenci (2020)), and Behavioral PISA (a similarity-based measure inspired by Computational PISA and Medina's Object Similarity measure). The four binary factors are telicity in Section 6.2, perfectivity in Section 6.3, iterativity in Section 6.4, and manner specification in Section 6.5.
In Chapter 7 (*Collecting acceptability judgments: materials and methods*) I will present the materials and methods employed in the behavioral experiments to collect acceptability judgments from native speakers of English and Italian relative to the implicit indefinite object construction. In particular, in Section 7.1 I will describe how I built the experiment with PsychoPy, how I ran it on Pavlovia, and how I recruited the participants via Prolific. Finally, I will present my 30-verb target dataset in Section 7.2, the experimental design in Section 7.3, the stimuli in Section 7.4, and the experimental setting in Section 7.5.
A first analysis of the data collected with these behavioral experiments will be provided in Chapter 8 (*Exploring the acceptability judgments*), where I will describe the structure of the Python script I devised to perform the analysis and to compute the models, as well as the procedures of data preprocessing employed (see Section 8.1), before discussing the separate and joint effects of semantic selectivity and the four binary predictors on the acceptability judgments in English and in Italian (see Section 8.2 and Section 8.3, respectively). In Section 8.4, I will argue that the five factors facilitating indefinite object drop are able to predict, to a non-negligible extent, the likelihood a transitive verb will appear without an overt object in a statistical (linear mixed-effects) model, and I will also explain why Medina's variant of Stochastic Optimality Theoretic is a more linguistically-motivated way of modeling these results than the linguistically-naive statistical model.
I will define and discuss my Stochastic Optimality Theoretic models of indefinite object drop in English and Italian in Chapter 9 (*Predicting the grammaticality of implicit objects*). In Section 9.1, I will describe and evaluate my 18 models, stemming from the union of three measures

of semantic selectivity, three increasingly more complex constraint sets (Medina's basic set, another with the addition of iterativity, and a full set with manner specification too), and two target languages. In Section 9.2, I will discuss the theoretical aspects and computational implementation of the two full models of object drop in English and Italian where semantic selectivity is modeled via Behavioral PISA. I will then compare my models with Medina's model and with regression models in Section 9.3.
Finally, I will provide my conclusions and propose some possibile future directions for research about modeling the implicit indefinite object construction in Chapter 10 (*Conclusions and open questions*).

1.2.2 Supporting materials

With an eye to the Open Science environment, I used open source software and programming languages to collect and analyse data for this thesis whenever possible, and I am sharing my data, scripts and results on GitHub. The interested reader will find my data, i.e., the stimuli for each experiment and the raw results I got from participants, in a dedicated GitHub repository[2] . In more detail, this repository contains:

- 30 target verbs and 10 filler verbs both for English and for Italian, used in all the computational (see Section 6.1) and behavioral (see Section 6.1.3 and Chapter 7) experiments, as in Appendix A;
- full lists of the direct objects of each target verb as extracted from the ukWaC corpus for English and from itWaC for Italian, both raw and manually cleaned (as detailed in Section 6.1.2);
- stimuli, full judgments elicited from 25 participants per language on a 7-point Likert scale, and final scores obtained in the Behavioral PISA experiment (see Section 6.1.3), also provided in Appendix B;
- each verb tagged with its features relative to the verb-specific predictors of object drop, i.e., telicity, manner specification, and semantic selectivity, as in Appendix C;
- stimuli and full judgments elicited from 30 participants per language on a 7-point Likert scale in the main behavioral experiment of this thesis (see Chapter 7), aimed towards creating a Stochastic Optimality Theoretic model of object drop in English and Italian (see Chapter 9), as in Appendix D.

As for the data processing, analysis of results, computational implementation of experimental designs, and creation of stimuli, I coded several Python scripts and documented their usage on GitHub. In detail, they are as follows:

- Quantify the polysemy of words in a list[3] using WordNet (Wu-Palmer Similarity), as in Section 7.2;
- Behavioral PISA[4] , a (behavioral) measure of Preference In Selection of Arguments to model verb argument recoverability, as in Section 6.1.3. The script takes care both of creating the stimuli for the experiment and of generating Behavioral PISA scores based on the Likert-scale acceptability judgments provided by human participants;
- PsychoPy Builder source code[5] of my behavioral experiments to collect acceptability judgments relative to the implicit indefinite

2: https://github.com/giuliacappelli/dissertationData

3: https://github.com/giuliacappelli/checkPolysemy

4: https://github.com/giuliacappelli/behavioralPISA

5: https://github.com/giuliacappelli/psychopy_exps object construction from native speakers of English and Italian, described in Chapter 7;

- Psychopy-to-Medina converter[6], to convert the output of my PsychoPy behavioral experiment (see Chapter 7) into a suitable input for my scripts to analyse the results (see Chapter 8) and create Stochastic OT models of the implicit object construction following Medina (2007) (see Chapter 9);
- Modeling the grammaticality of implicit objects[7] based on Medina (2007)'s variant of Stochastic Optimality Theory, as in Chapter 8 and Chapter 9;
- Generate mock Likert-scale acceptability judgments[8] based on factor levels specified in the input, to test the above Stochastic Optimality Theoretic model on ideal data before running the experiment proper.

6: https://github.com/giuliacappelli/PsychopyToMedina

7: https://github.com/giuliacappelli/MedinaStochasticOptimalityTheory

8: https://github.com/giuliacappelli/generateMockLikertGrammaticalityJudgments

1.2.3 Published work and outreach

Relevant parts of the experimental section of this thesis have been shared with the scientific community, both in written form and during conferences. The original distributional measure of Preference In Selection of Arguments (Computational PISA) presented in Cappelli and Lenci (2020) and discussed here in Section 6.1.2, tested on large sets of transitive verbs and Instrument verbs in English, was presented at:

- *SEM 2020, 9th Joint Conference on Lexical and Computational Semantics, December 12-13th 2020, online due to the Covid-19 pandemic (originally in Barcelona, Spain);
- LSA 2021, 95th Annual Meeting of the Linguistic Society of America, January 7-10th 2021, online due to the Covid-19 pandemic (originally in San Francisco, California);
- CLiC-it 2020, 7th Italian Conference on Computational Linguistics, March 1-3rd 2021, online due to the Covid-19 pandemic (originally in Bologna, Italy).

The results of the main behavioral experiment of this thesis (detailed in Chapter 7 and Chapter 8), especially the ones pertaining to Italian, were presented at:

- SyntOp 2022, Syntactic Optionality in Italian, July 4-5th 2022, Venice (Italy).

Theory and literature review

Indefinite object drop | 2

As mentioned in Chapter 1, this thesis is about implicit indefinite objects. What is "indefinite" about them? In what sense can they be considered "implicit"? And ultimately, what is objecthood itself? This Chapter will answer these questions in reverse order, from the most general to the most specific one. In Section 2.1 I will make reference to the transitivity continuum, as defined by Hopper and Thompson (1980) and further explored by later literature. In Section 2.2 a crucial distinction will be made between definite and indefinite object drop, following Fillmore (1986) and subsequent works. The nature of *indefinite* object drop will be finally described in Section 2.3 and Section 2.4, and a working definition (for the purposes of this thesis) will be provided in Section 2.5.
Before delving into the main contents of the Chapter, a terminological clarification is in order. Throughout this thesis, I refer to an "intransitive" use of transitive verbs to intend the absence of a possible overt syntactic object for verbs which semantically take an Agent and a Patient argument, e.g., *John broke (the window)*. Crucially, I am *not* referring to senses of such verbs where the subject is non-Agentive (e.g., *The ball broke the window*), or to their unaccusative (Perlmutter 1978), anticausative uses[1] (e.g., *The window broke*), nor am I referring to verbs that have two semantic participants (an Agent and a Patient) but can only express the internal argument[2] (corresponding to the Agent participant) syntactically, such as *to dine*.

1: It is important to make this distinction explicit, since some authors pair together anticausatives and object-less uses of optionally transitive verbs. For instance, Huddleston et al. (2002, pp. 216–217) labels them both "dual-transitivity verbs".

2: Refer to Engelberg (2002) for more considerations on the difference between such verbs and the ones exhibiting optional object drop.

2.1 Transitivity as a prototype

School kids everywhere are used to call "transitive" the verbs which take an overt direct object. In a traditional semantic definition, a clause is deemed "transitive" if it describes an event where the action performed by an Agent "passes over"[3] to a Patient, which usually undergoes some kind of transformation.

3: Hence the name of the concept of transitivity, from Latin *transire* 'to go over'.

Going beyond these naive definitions, but still capturing their spirit, Hopper and Thompson (1980) first proposed an account where transitivity is interpreted as a scalar concept whose strength depends on several parameters, or, to use more modern terminology, as a prototype category (Næss 2007). In particular, they identified ten parameters (Hopper and Thompson 1980, p. 252), reported almost *verbatim* in Table 2.1.

These parameters are potentially active in all languages, but languages may differ from one another with respect to the actual subset of parameters they select as necessary criteria for transitivity. This depends on the "recursivity" (Næss 2007, p. 29) of prototypical concepts, which assign membership in a category (in this case, transitive clauses) on the basis of attributes which are prototype concepts themselves (J. R. Taylor 1995, p. 61).
Parameters A, B, E, F, and G from Table 2.1 are self-explanatory. Parameter

		high transitivity	low transitivity
A.	**Participants**	2+ (Agent and Object)	1 participant
B.	**Kinesis**	action	non-action
C.	**Aspect**	telic	atelic
D.	**Punctuality**	punctual	non-punctual
E.	**Volitionality**	volitional	non-volitional
F.	**Affirmation**	affirmative	negative
G.	**Mode**	realis	irrealis
H.	**Agency**	A high in potency	A low in potency
I.	**Affectedness of O**	O totally affected	O not affected
J.	**Individuation of O**	O highly individuated	O non-individuated

Table 2.1: Hopper and Thompson (1980, p. 252) defined transitivity as a prototype concept determined by ten parameters.

C (telicity) will be discussed in more detail in Section 3.2.1. Parameter D (punctuality) refers to the phase between inception and completion of an action, which is non-existent in verbs like *to kick* and obviously present in verbs like *to carry*. Parameter H (agency) separates animate and inanimate subjects. Parameter I (affectedness of the object) determines that sentences like *I drank up the milk* are more transitive than sentences like *I drank some of the milk*, since the milk is only partially affected by the drinking in the latter sentence. Finally, parameter J (individuation of the object) refers to the distinctness of the object both from the Agent and from the background, as summarized in Table 2.2 (Hopper and Thompson 1980, p. 253).

individuated	non-individuated
proper	common
human, animate	inanimate
concrete	abstract
singular	plural
count	mass
referential, definite	non-referential

Table 2.2: Hopper and Thompson (1980, p. 252) defined transitivity as a prototype concept determined by ten parameters.

The individuation parameter is the most controversial among the ten proposed ones, as observed by Comrie (1989, p. 128) and later on by Næss (2007, p. 18). According to what has come to be known as "Comrie's generalization", in prototypical transitive clauses both animacy and definiteness are high in the Agent and low in the Patient, *contra* Hopper and Thompson (1980). The weak argumentation Hopper and Thompson (1980) provide in favor of the individuation parameter is that speakers would be more likely to focus on the Patient in *I bumped into Charles* than in *I bumped into the table*, since bumping is more likely to affect human beings than tables. Comrie (1989) makes a much more compelling point with reference to cross-linguistic typology, basing the generalization on the animacy hierarchy and on referential case-marking (which I will not discuss here, since it would lead me too far from my topic). Later literature (Kardos 2010; Kemmer 1993; Næss 2007, 2009) reinforced this point by assuming that prototypical transitive events are described by verbs whose subject and object are maximally distinct from a semantic point of view.

To sum up, the terse summary by Næss (2007, p. 15) clearly shows the relation between the naive definitions of transitivity and the ten-parameter account by Hopper and Thompson (1980). A prototypical transitive clause is understood to describe an event such that:

- a volitional Agent (E, H)
- performs an action (B)
- with a tangible, lasting effect on a Patient (A, I, J),
- and it is presented as real and completed (C, D, F, G).

Lorenzetti (2008, p. 78) provides a tighter cluster of parameters, arguing that only a subset of the ones proposed by Hopper and Thompson (1980) are truly relevant. In particular, the author ditches the criteria relative to the transitive event being real and completed (C, D, F, G), and only keeps agentivity, affectedness, and individuation of the object among the other groups of criteria.

2.2 Definite *vs* indefinite drop

In Section 2.1 I introduced transitivity as a prototype concept depending on a cluster of parameters and, specifically, involving an Agent acting on a Patient. But what about utterances where events of this kind are expressed without an overt syntactic object? In this Section I will provide an account of the literature on the matter.

2.2.1 Either definite or indefinite: discrete accounts

Introduction Verbs behaving intransitively are, using words by Rutherford (1998, p. 191), "a mixed bag". Consider, for instance, the examples in (1). The sentence in (1-a) features a typical intransitive verb, describing an event where the subject is *not* performing an action with effects on some Patient. The sentence in (1-b) also describes an event where the subject is not volitionally acting on a Patient, but it is nevertheless clear that there has to exist something that John knows (unlike in (1-a), where there cannot be something that John sleeps). The sentence in (1-c) has an Agent acting volitionally on a Patient, which is however not instantiated syntactically as an overt direct object.

(1) a. John slept.
b. John knew.
c. John ate.

There is a clear similarity between (1-b) and (1-c), as opposed to (1-a). They both require a Patient/Theme semantically (Somers 1984, p. 510), and they both surface as object-less syntactically. Quoting Yasutake (1987, p. 48), "they are different from pure intransitives in that the action will not be complete without some lexically implied (but unspecified) object". Oddly, some literature (Bourmayan and Recanati (2013) and D. Liu (2008) a.o.) does not interpret such verbs as transitive-become-intransitive verbs via the omission of the direct object, but as intransitive-made-transitive verbs. Such an interpretation is totally counter-intuitive and it goes against the generally-accepted tenet that a core feature of so-called "intransitive verbs" is that they have no object slot available in the syntax.
There is, however, a crucial difference between (1-b) and (1-c). Native speakers of English insist that they have to be provided some context in order to understand (1-b) —what is it exactly that John knew? On

the contrary, (1-c) can be interpreted to mean that John had a meal at a certain moment in time, without specifying any additional context. This distinction was captured and defined by Fillmore (1986) (building upon Allerton (1975) and Fillmore (1969)), the established seminal work on the distinction between so-called "definite" and "indefinite" object drop (here represented by (1-b) and (1-c), respectively).

Fillmore's account Fillmore (1986, p. 96) distinguishes between Indefinite Null Complements (hence, INC) and Definite Null Complements (hence, DNC) by testing "whether it would sound odd for a speaker to admit ignorance of the identity of the reference of the missing phrase". So, making reference to (1) again, there would be no issue with saying "John ate. I wonder what he ate.", but it would be quite odd to say "John knew. I wonder what he knew.". Thus, the missing object in (1-b) is a DNC, while the missing object in (1-c) is an INC. Fillmore than splits INCs into two sub-groups based on whether the omitted object is "of considerable generality" or "requiring the specification of various degrees of semantic specialization". The examples in (2) (Fillmore 1986, pp. 96–97) show increasing degrees of, using his words, "semantic specialization". In (2-a), the subject cannot perform the very act of eating or drinking, regardless of the actual ingested item. In (2-b) something specific was eaten, but its specific nature is irrelevant inasmuch the speaker is referring to eating as the act of having a meal. In (2-c) the omitted object is referring not just to any drinkable liquid, but to alcohol specifically. Finally, in (2-d) something very specific was baked by the subject, but this information is backgrounded[4] to focus on the activity itself (I will come back to this in Section 2.4.2).

4: Refer to David (2016) for more observations on implicit objects and other omissible elements usually being "the ground in a figure-ground relation".

(2) a. When my tongue was paralyzed I couldn't eat or drink.
b. We've already eaten.
c. I've tried to stop drinking.
d. I spent the afternoon baking.

However, as it will be shown throughout this Chapter, this secondary division of INCs into subgroups does not have to be a binary theoretical distinction. On the contrary, it follows from several finer-grained considerations on the nature of INCs and the factors allowing them. Moreover, this binary division actually opens the door to discussions about a DNC-INC continuum (more details on this in Section 2.2.2). Specifically, how is "semantic specialization" different from the "knownness" of the object in DNC constructions (Eu 2018, p. 525)? An answer can be found in Allerton (1975, p. 218), where the case is made that semantically specialized INCs (just like any INC) refer to a category of individuals, while DNCs refer to specific instances of a given category.
Going back to the main distinction between INCs and DNCs, finally, Fillmore (1986) formally defines the former as objects whose "referent's identity is unknown or a matter of indifference" and the latter as objects whose referent "must be retrieved from something *given* in the context". This context "has either to be given linguistically, in the preceding context, or extralinguistically, in the situational context" (Stark and Meier 2017, p. 13).

Other accounts Several other researchers made use of the distinction between definite and implicit indefinite objects brought to the fore by Fillmore (1986), providing slightly different definitions which capture different aspects of the phenomenon. Allerton (1975), then further developed by Fillmore, distinguishes between "contextual omission" and "lexical omission", which respectively refer to Fillmore's DNCs and INCs. Cummins and Roberge (2004) distinguish between "internally-licensed null objects" (INCs) and "referential null objects" (DNCs). According to Ruppenhofer (2005) and Pethõ and Kardos (2006, p. 30), INCs "receive an *existentially quantified* interpretation", while DNCs are "interpreted anaphorically and must therefore have an appropriate antecedent in context to make sense" (Fillmore (1986) himself referred to this in the title itself, "Pragmatically Controlled Zero Anaphora"). The anaphoric status of DNCs is also central in Keller and Maria Lapata (1998). Medina (2007, p. 13), the foundational work upon which I am basing my own model of the indefinite object construction, sees DNCs as "implicit objects whose particular meaning can be recovered from the preceding discourse or disambiguating physical context" and INCs as "implicit objects whose meaning is recoverable only from the verb in the sentence". Here the focus is all on recoverability, and the author goes on to show that semantic recoverability can be a reliable predictor of object drop in INC sentences. D. Liu (2008, p. 293), following García-Velasco and Muñoz (2002), takes the shift away from lexical semantics and onto aspectual territory. In particular, the point is made that INCs involve a change of focus "from the object in the transitive use to the activity (the verb) itself in the intransitive use" (an idea that I will discuss in full detail in Section 2.4.2), while DNCs do not determine such a shift.
The accounts provided so far are not in conflict with Fillmore's formulation of the problem at hand, nor are they in conflict with the view I am adopting in this thesis in order to provide a probabilistic model of the implicit object construction. Other accounts, on the other hand, are more challenging and deserve a more thorough clarification. Let us consider the most relevant ones for my argumentation.
Tonelli and Delmonte (2011, p. 55) argue that, while INCs are "constructionally licensed, in that they apply to any predicate in a particular grammatical construction", DNCs are "lexically specific, in that they apply only to some predicates". Later in this Section (on Page 13) and in Section 2.2.2 I will bring evidence in support of the opposite point of view, which is in favor of seeing DNCs as (extra- and intra-linguistically) contextually, not lexically, determined. Moreover, Tonelli and Delmonte's (2011) account is in direct conflict with Fillmore (1986, p. 95), who argues that INCs are "limited to particular lexically defined environments" (such as the object slot of *to eat* and *to read*). In this regard, I side with Tonelli and Delmonte (2011). My probabilistic model of INCs (the results thereof are discussed in Chapter 8 and modeled in Chapter 9) will provide strong evidence in support of the idea that any transitive verb can participate in INC constructions, provided the right aspectual, semantic, and pragmatic features. Indeed, as noted by Huddleston et al. (2002, p. 216), transitivity is better thought of as a property of verb *use*, rather than a feature of verbs themselves.
In a pragmatic (in particular, not lexical) perspective, AnderBois (2012, p. 44) and Melchin (2019, pp. 53–54) both stay true to Fillmore's original interpretation of DNCs as "pragmatic anaphoras", arguing that DNCs

corefer with other referents in the discourse. Moreover, they maintain that INCs lack the possibility of having coreferential interpretations. Fillmore himself (Fillmore 1986, p. 97) made this point with example (3), where (3-b) cannot be considered a proper answer to the question in (3-a). This is taken to mean that there is no co-reference between the sandwich and the implicit object of *to eat* in (3-b).

(3) a. What happened to my sandwich?
 b. *Fido ate.

However, examples can be provided in support of the opposite. Groefsema (1995, pp. 142–144) makes use of sentences such as the one in (4) to argue that INCs can indeed refer to specific individuals, provided sufficient linguistic context. As Scott (2006, p. 168) observes, omissions of this kind "allow the speaker's meaning to hover between the definite and indefinite readings", so that the interpretation the hearer applies to the omitted object is "specific yet indefinite". I will come back to other blurred distinctions between definite and indefinite object drop in Section 2.2.2.

(4) John picked up the glass of beer and drank.

Nevertheless, this account does not disrupt Fillmore's foundation. As explained by Eu (2018, p. 527), not even in sentences like (4) do INCs force the identification of a specific referent. What happens, instead, is that native speakers processing INCs in a flexible context of this kind can be induced to understand the missing object as if co-referring to the previously mentioned one. Thus, INCs can grammatically dissociate the mentioned referent from the one implied by the missing object, while this possibility is not active for DNCs (which are always co-referential, regardless of the context). Considering eventualities like this, it really is no wonder that Cote (1996, p. 110), with respect to implicit objects in English, spoke of "murky water" in relation to the distinction between lexically-provided information and context available via world knowledge.

Genre-based implicit objects: a special case of definite object drop I will now discuss genre-based implicit objects, a type of DNCs whose very existence goes in favor of DNCs being virtually possible with any verb, provided it appears in a discourse context that is conducive to object omission (*contra* Tonelli and Delmonte (2011, p. 55), *pro* Goldberg (2001)). To quote Ruppenhofer and Michaelis (2010, p. 175), "argument omission can but need not be licensed by a lexeme". This possibility was first observed by Fillmore (1986, p. 95), who acknowledged that in "certain kinds of highly restricted mini-genres" (e.g., instructional imperatives in recipes) the omission of objects and other non-subject complements is not lexically determined (see also Haegeman (1987, p. 237)). Crucially, it is not the case that there is a special grammar of recipe contexts that supersedes the actual grammar of the language the recipe uses (Cote 1996; Culy 1996). On the contrary, recipes and other specialized genres just serve to provide an encompassing discourse and world-knowledge context to the listener/reader. Sigurðsson and Maling (2008, pp. 30–31) observe that this strong link between genre-licensed DNCs, discourse context, and communication goals may be the reason why this type of object drop is much more common cross-linguistically than other types. Not only that, but genres are so intertwined with argument omission that it is sometimes possible to evoke a genre just by performing the right kind of DNC, as Ruppenhofer and Michaelis (2010, p. 159) exemplify by making reference to the title of a novel by Cynthia P. Lawrence, "Chill ∅ before Serving ∅: A Mystery Novel for Food Lovers" (a clear reference to instructional imperatives found in recipes).

Many linguistic analyses of DNCs licensed in "mini-genres" focus on recipe contexts (Ahringberg 2015; Bender 1999; Culy 1996; García-Velasco and Muñoz 2002; Massam 1992; Massam and Roberge 1989; Megitt 2019; Paul and Massam 2021; Ruda 2014). In particular, Culy (1996) performed a multiple regression analysis on diachronic sets of contemporary and historical recipes with several predictors, finding that the style of a recipe book and discourse factors are the most important predictors of recipe DNCs.

Other authors provided accounts pertaining to a broader spectrum of genres. For instance, in addition to recipes, Cote (1996) also considers "telegraphese", i.e., the telegraphic register used in telegraphs, memos, and signs. A. Weir (2017) focuses on what he calls "reduced written register" in English, i.e., the absence of objects in recipes, instructional/directive imperatives[5] , diaries, text messages, internet-based communication, and similar contexts. The presence of DNCs in text messages and internet-based communication is further explored by Stark and Meier (2017) and Stuntebeck (2018), focusing on Whatsapp messages. D. Liu (2008, p. 304) mentions instructional languages, such as that found on manuals, warning signs, and product labels. Paesani (2006) provides a thorough account of object (and subject) drop in special registers (such as recipes, diaries, and headlines), noting clear similarities between DNCs in recipes and the broader phenomenon of Topic drop (Paesani 2006, p. 165). Focusing instead on non-contemporary language found in historical texts, it is possible to find studies by Almeida (2009) on Middle English medical texts, and by Korkiakangas (2018) on Early Medieval documentary Latin. In an unconventional account of football language, Bergh and S. Ohlander (2016) argue that verbs licensing DNCs are "monotransitives" (Quirk et al. 1985, p. 54) in this sublanguage, since they can only take one argument. Let us consider the examples in (5).

5: As noted by Ruppenhofer and Michaelis (2010, p. 162), non-instructional imperatives cannot participate in DNC constructions, as shown in their example *Take *(the money) and run*.

(5) a. Iniesta passed (the ball) and Messi finished (the attack) clinically.
b. John passed *(the salt) and finished *(his steak).

In (5-a), the direct objects can be omitted because the two footballers are performing acts that need no further explanation in the football community. In this game, you can only pass balls and finish attacks. Incidentally, Dvořák (2017b, p. 266) notes something similar about some verbs in Czech (e.g., *smeknout* 'to uncap, to tip' the hat one is wearing, *zaparkovat* 'to park' the vehicle one is driving) having "idiomatized meanings [...] limited to a particular jargon or slang" and, crucially, allowing "only one particular entity in the role of an internal argument". Moreover, given the presence of a single ball against many players, match reports are much more likely to DNC the ball rather than the footballers, as observed by Ruppenhofer and Michaelis (2010, p. 167) and Ebeling (2021). On the contrary, in the probable context of a dinner in (5-b), it is not possible to say that John just "passed" or "finished", let alone clinically. Bergh and S. Ohlander (2016, p. 22) explain the existence of DNCs in football reports (and, more generally, object omission) as a manifestation of the "principle of least effort" (Zipf 1949) and also of the Gricean pragmatic maxim of quantity, which compels speakers to avoid being more informative than necessary. However, as Ruppenhofer and Michaelis (2010, p. 166) observed before, "genre-based omissions are never obligatory", since the maxim of quantity (favoring implicit objects) is counterbalanced by the need for informativeness (favoring overt objects).

2.2.2 Neither definite nor indefinite: continuous accounts

The account of genre-determined DNCs offered in Section 2.2.1 opens the door to a broader discussion of Fillmore's (1986) distinction between definite and indefinite omitted objects. As argued in Ruppenhofer and Michaelis (2010, p. 165) and Bergh and S. Ohlander (2016, p. 24), the main factor allowing for an object to be omitted is its recoverability (refer to Section 3.1.1 for a full discussion), which depends on linguistic aspects as well as on contextual and discourse factors, and on world knowledge too. Focusing on recoverability makes it possible to go beyond the binary distinction between INCs and DNCs provided by Fillmore (1986) and many others, and also beyond the need to postulate verb-specific object-dropping capabilities. In particular, it paves the way for a non-binary account of object drop, where no clear-cut distinction between two types of omission have to be postulated (something that, in essence, was already implicit in Hopper and Thompson's (1980) assumptions).
If recoverability is the cornerstone of object-droppability, and if it is a scalar, or even graded, feature of objects, then it stands to reason that object-droppability itself is a graded phenomenon. Taking a small step forward in this direction, AnderBois (2012) posits the existence of "flexible implicit arguments" to explain sentences like *The Giants won* ∅, whose implicit object has a referent known to the reader (as with DNCs), which is, crucially, known because of world knowledge[6] and not because of the presence of a linguistic antecedent (as with INCs). Cummins and Roberge (2005) provide what they call a "modular account" of null objects in French, stemming from the intersection of several syntactic, semantic, pragmatic and discourse factors (a similar account of object drop, still abiding to Fillmore's binary distinction, is provided by Cennamo (2017)). A more cogent, continuous account is offered by Glass (2013), who acknowledges that there is "plenty of middle ground" between minimum recoverability (an object has to exist but it is unknown) and maximum recoverability (the specific identity of the object is known). In particular, she argues that, in order to be omitted, an object just has to be sufficiently recoverable for speakers to communicate felicitously, and that community- or genre-specific sublanguages are more prone to certain kinds of object omission simply because those smaller contexts favor object recoverability. Moreover, she explicitly argues against an INC-DNC distinction (Glass 2013, p. 1). A pioneering attempt to bring evidence in favor of the intuition that recoverability is the key in object omission is found in Resnik (1993, 1996), an information-theoretic account of selectional constraints testing, among other things, the relationship between transitivity and discourse context (more on Resnik's method in Section 6.1.1).

6: In this case, world knowledge about American football.

As I will argue later in Section 2.5, I stay agnostic with respect to the binary or continuous nature of object-droppability in my account of indefinite null objects. Following binary accounts, such a study would simply be a matter of considering those factors which are known to favor the emergence of INCs. On the other hand, under continuous-droppability assumptions, it would be a more complex matter of modeling both contexts and linguistic factors determining any kind of object drop, trying to position implicit null objects in a specific portion of the object-droppability spectrum.

2.3 Defining the indefinite

In this Section I will delve into a detailed discussion on implicit indefinite objects, the focus of this dissertation. In Section 2.2 I reported a series of both now-classic and more recent accounts of the differences between so-called definite and indefinite null objects. Let us now comment on the nature of the latter, which were given several labels in the literature (objects of "detransitive verbs" in Yasutake (1987, p. 46), "implicit objects" in Glass (2013) and Pethõ and Kardos (2006, p. 29), or "pseudo-intransitive", "labile", "ambitransitive", "null complements", "understood arguments", "unspecified objects", "null instantiations" in other authors).

2.3.1 Which verbs?

Traditionally (*contra* this thesis and Tonelli and Delmonte (2011, p. 55), among others), indefinite null objects are taken to only be possible with a restricted set of activity verbs. For instance, Rizzi (1986, p. 510) provides an argument in favor of indefinite object drop being "lexically governed" in English on the basis that in some pairs of semantically related verbs (e.g., *to eat* and *to devour*) one member of the pair allows for object drop, while the other does not. Haegeman (1987, p. 236) does not hesitate to define this account "convincing", and similar notes are also found in Fillmore (1986), Gillon (2012), Mittwoch (2005), and Rice (1988). I will come back on the theory referring to specific case of these "semantic minimal pairs" in Section 3.1.3. The important aspect, here, is that traditional or traditionally-leaning literature[7] has trouble motivating implicit indefinite objects on the basis of meaning alone, but it also needs it to be lexically determined. With that said, which verbs does then the literature identify as allowing indefinite object drop?

First of all, the verbs under consideration drop *syntactic* arguments, crucially, not *semantic* ones. In other words, indefinite object drop are obligatory semantic arguments of such verbs (Cote 1996, p. 120), while they do not surface syntactically (more on the syntax of implicit indefinite objects in Section 2.4.2). Jackendoff (2003, p. 134) specifically observes that it is quite inaccurate to say that such verbs "license an optional argument", since this definition "conflates semantic and syntactic argument structure". He illustrates this point by comparing *to eat* and *to swallow* in (6). Both show identical syntactic behavior, but while it is possible to swallow without swallowing anything[8], it is not possible to eat without eating something.

7: Such as Ruppenhofer and Michaelis (2014), arguing that object drop is "an aspect of argument realization licensed by lexemes, which may differ from one another in idiosyncratic ways".

8: If one does not consider saliva.

(6) a. Bill ate (the food).
b. Bill swallowed (the food).

Another key point in traditional literature on implicit indefinite objects concerns the difference between change-of-state verbs (such as *to break, to harden, to open*) and pseudo-transitive verbs (such as *to eat, to write, to sweep*). Verbs belonging to the former class are prototypically transitive (Hopper and Thompson 1980; Kardos 2010; Lemmens 2006), since they feature two maximally distinct arguments (a volitional Agent subject and a non-volitional Patient object), while verbs belonging to the latter class exhibit both transitive and intransitive features (Armstrong 2011). Thus, only pseudo-transitives can license indefinite null objects in this dichotomy. They also appear to be a semantically rich class, comprising verbs of creation (e.g., *to cook, to write, to knit*), verbs of ingestion or consumption (e.g., *to eat, to drink*), and verbs of surface contact (e.g., *to sweep*). While syntactically they have their ambivalent behavior in common, semantically they share the fact that their objects all are "incremental themes"[9] , a term originally proposed by Dowty (1991) to refer to verbs showing homomorphism between the physical extent of their object and the temporal progress of the event. Thus, *to eat* is an incremental-theme verb because the Patient gets progressively smaller while ingested by the Agent, *to write* because the Theme gets progressively more wordy while the Agent types or pens it, and so on. On the contrary, *pace* Dowty's attempt to apply this analysis to change-of-state verbs (Dowty 1991, p. 568), *to close* is not an incremental-theme verb because sentences like *Matt closed the door half-way* do not entail that half the door was closed (Rappaport Hovav and Levin 2005, p. 279). Since incremental-theme verbs can behave both transitively and intransitively, in Levin (1993, p. 33) they are said to participate in the "unspecified object alternation". The author also provides a list of more than 40 verbs allowing for indefinite object drop, an event which Dvořák (2017a, p. 116) praises as a major breakthrough after previous literature only focusing, "somewhat disturbingly", on the sole verb *to eat*[10] .

9: Please refer to Rappaport Hovav and Levin (2005, p. 279) and Kardos (2010, p. 4) for an extensive account of incremental themes.

10: An unfortunate choice for a single example of the whole category of object-dropping verbs, as we will see later in Section 2.4.2.

As I will show with the probabilistic model of indefinite null objects I define in this dissertation (final results presented in Chapter 9), object drop is possible both with change-of-state verbs and with incremental-theme verbs, the difference being a matter of degrees (determined by several linguistic factors, see Chapter 3 and Chapter 6), not a binary feature as in traditional accounts.

2.3.2 Which objects?

While syntactically unexpressed and semantically unspecified (at least with respect to a specific entity), implicit indefinite objects of the verbs allowing them still have to refer to *something*. To what, though? Is it possible to generalize the answer?

Omitting *something* This *something* that object-dropping verbs refer to has been interpreted quite literally in traditional literature on the issue. Katz and Postal (1967) and Fraser and Ross (1970) distinguish between the deletion of *it* (expressed by the constructions which Fillmore (1969, 1986) made known as "definite null complements") and the deletion of *something* (expressed by indefinite null objects). A similar consideration is also found, in passing, in Pethõ and Kardos (2006, p. 30) and in Ahringberg (2015, p. 6).

However, several objections can be made to this idea. Historically, the first came from Mittwoch (1982, 2005), who argued that the omitted object cannot be *something* because "this would be incompatible with the atelic nature of the resulting sentence". I will come back to telicity, and the somewhat different approach I will embrace in the next Chapters, in Section 3.2.1. Other authors (Condoravdi and Gawron 1996; Dvořák 2017a; Jerry Fodor and Janet Fodor 1980; Gillon 2006, 2011, 2012; Lasersohn 1993; Martí 2015; Melchin 2019) argue instead that implicit indefinite objects have to be interpreted as "weak indefinites" (Melchin 2019, p. 55) as bare masses and plurals, instead of the indefinite pronoun *something*, because only the former have obligatory narrow scope with respect to other quantifiers in the sentence (a behavior shown by indefinite null objects)[11] . Let us consider the examples in (7), taken from (Melchin 2019, p. 55). In (7-a) only the universal quantifier in the subject can take wide scope and the implicit object has to take the lowest scope, while in (7-b) either the universal quantifier in the subject or the existential quantifier in the object can take wide scope.

11: Please refer to Carnie (2012), a thorough handbook of generative syntax, for a gentle introduction to scope-taking constituents and the use of logic operators in syntax.

(7) a. Everyone ate. $\forall > \exists$ / # $\exists > \forall$
b. Everyone ate something. $\forall > \exists$ / $\exists > \forall$

Let us unpack this notation. This means that in (7-a), for every (universal quantifier $\forall$) person, there exists some edible entity that was eaten. In (7-b) this interpretation is possible too, as well as the wide-scope-existential interpretation. In this second reading, which appears to be less perspicuous than the other, for some edible item (existential quantifier $\exists$), every person ate it.

The observation that implicit indefinite objects can only take low scope also holds with respect to other logic operators than quantifiers. For instance, Gillon (2012, p. 316) provides example (8) about negation, where the operator is shown to take obligatory wide scope over the sentence.

(8) Bill did not read. $\neg\exists x\ Rbx$ / # $\exists x\ \neg Rbx$

This means that the only possible interpretation of this sentence is that Bill did not read anything (the first proposed truth condition of the sentence), not that there exists something that Bill did not read (the second truth condition, marked as improper).

Therefore, implicit indefinite objects have to be interpreted as weak indefinites (bare masses or plurals), not as the pronoun *something*. However, it is clear that not *any* weak indefinite can be interpreted as the omitted object of a given verb. Which ones are the right ones? I am now going to provide some answers based on the literature.

Prototypical objects Van Valin and LaPolla (1997, p. 122) coined the term "inherent arguments" to describe the implicit indefinite objects occurring with activity verbs, based on the idea that they denote a facet of the meaning of the verb, characterizing the action itself rather than a participant. I will devote some space to a detailed discussion of intransitivization as a means to focus on the activity itself on Page 23. Resorting once again to the concept of linguistic prototype, which is indeed central in this discussion of transitivity (Section 2.1) and object drop, much literature agrees on indefinite null objects being understood as prototypical[12] arguments of the verb (Bresnan 1978; Dvořák 2017b; Levin 1993; Lorenzetti 2008; Melchin 2019; Mittwoch 2005; Næss 2007; Quirk et al. 1985; Rice 1988). However, what is a prototypical object of a given transitive verb? Rice (1988, p. 204) provides the examples in (9).

12: Note that while all referenced works appeal to the notion of prototypical argument, not all of them phrase this idea exactly in these terms. Some refer to "implied arguments", "default interpretations", "standard objects", "understood object", and other labels in the same vein.

(9) a. John smokes (cigarettes / *Marlboros / *a pipe / *SMOKING MATERIALS).
b. John drinks (alcohol / *gin / *water / *coffee / *LIQUIDS).
c. When he goes to Boston, John drives (a car / *a Toyota / *a motorcycle / *A VEHICLE).
d. Each afternoon, John reads (a book / *Ulysses / *the newspaper / *PRINTED MATTER).

Examples in (9-a), (9-c), and (9-d) all appeal to our world knowledge, in particular, to our knowledge of what is the most probable choice of the average Joe (or John, in these examples). People usually smoke cigarettes which do not have to be necessarily Marlboros, they drive differently branded cars in their trips out of town, and they like to read generic books in the afternoon[13] .

13: Apparently, for many people, especially the ones still liking their news to be printed on paper, reading the news is a leisurely activity to be specifically enjoyed in the mornings while having breakfast. This habit had to be even more common in the late ’80s than today.

As noted by Næss (2007, p. 125), however, example (9-b) poses a challenge to the protypicality-enabled omission theory. The most typical liquid one can drink is usually water, not alcohol. And if one was to understand that the omitted substance is alcohol due to its very omission, would not this argumentation become circular?[14] The problem alcohol poses for linguists (or better, for linguistic theory) can be explained from different angles. I, for one, would appeal to the gricean maxim of relevance, in that water is indeed the typical liquid we drink, but it is so much typical (being necessary for good health and even life) that it would be actually weird to mention water-drinking in casual conversation. No one would bat an eye at John drinking water, so it would make little sense to utter (9-b) implying water-drinking. Indeed, we only refer to the act of drinking water when it becomes relevant, for instance during hot summers (*Remember to drink!*) or on a Sunday morning (*I am so glad I drank water before going to bed.*). Thus, since the only socially relevant, statistically likely, choice of a drink for John in (9-b) is alcohol, that is what we intend as a prototypical, omissible object for the verb *to drink*. This perspective is also echoed by Newman and Rice (2006, p. 14), who ascribe the intransitive use of *to drink* to "the prominence of alcohol consumption as a topic of discourse". Another possible account is the one by Goldberg (2005a, pp. 21–28), where "taboo verbs" are argued to facilitate object drop due to our culturally-induced shame in mentioning that which is perceived as unmentionable in polite society (such as bodily fluids or, in this case, enjoying alcoholic drinks).

14: See also Mittwoch (2005, p. 20) on the issue of circularity.

Changing perspective, when presented with puzzling verb behavior such as the one expressed in (9-b), Huddleston et al. (2002, pp. 303–305)[15] needlessly assume that such verbs participate in two different patterns of object-droppability, i.e., "specific category indefinites" (where the omitted liquid would be interpreted as being of the alcoholic variety) and "normal category indefinites" (where the omitted liquid would be interpreted as being water). This account is flimsy at best, since it puts labels on

15: Following Fillmore (1986, pp. 96–97), as discussed in Section 2.2.

a given state of affairs without actually providing an explanation for this epiphenomenal dichotomy. Næss (2007, p. 141) offers a much more compelling account, which is both descriptive and explanatory, where the "specialized meanings" *to drink* and other ingestion verbs take get explained by the affectedness of their Agent[16] . In other words, omitting the object highlights the effect the action described by the verb has on the Agent (e.g., getting the Agent inebriated) by backgrounding the effect it has on the Patient. The fact that intransitive *to drink* elicits a drink-alcohol reading much more readily than intransitive *to eat* elicits a eat-a-meal reading (also noted by Newman and Rice (2006, p. 14)) is explained by the author (and again later in Næss (2011, p. 420)) with reference to world knowledge and social norms. In particular, the case is made that intoxication by means of alcohol not only has a direct effect on the imbibing Agent (who takes on the drinking endeavor with this precise goal), but it also has an indirect, sometimes unintended effect, i.e., making the Agent appear visibly drunk, and thus disrespectful of several unwritten societal rules. This doubly-affected-Agent reading gives then rise to the highly specialized reading of intransitive *to drink*.

16: More on Agent affectedness on Page 25.

Interestingly, Yasutake (1987, pp. 48–50) suggests a three-way graded account of the different types of objects which can participate in implicit indefinite object constructions where the prototypicality of the omitted object is taken to be a rather flexible requirement for omission. In fact, the omissibility-as-prototypicality accounts I presented so far in this Section all made reference to the omitted object being somewhat "typical" of the verb, so that less typical Patients of the same verb are less likely to be omitted (or require quite the flight of fancy to be accounted for, as seen in the proposal by Huddleston et al. (2002) about the verb *to drink*). Yasutake's perspective integrates the prototypicality intuition with other accounts based on world knowledge, envisioning these three types of implicit object:

- typical objects (e.g., *to read*);
- socially-understood objects (e.g., *to drink, to shave, to drive*);
- semantically unspecified objects of highly specialized activities (e.g., *to steal*).

Most importantly, rather than saying that optionally transitive verbs can omit their prototypical Patient, it would be best to say that object drop is licensed by the recoverability of the prototypical Patient of a given *sense* of a verb (Fillmore 1969, p. 100). This is evident, for instance, in example (10) by Iten et al. (2005).

(10) I applied.

Indeed, as she argues, the verb in the example is acceptable when used intransitively only if it is understood to refer to the job-seeking sense, not to the bandage-application sense.

2.4 How many lexical entries?

In Section 2.3, I discussed some traditional views on the characteristics a verb has to express in order to license implicit indefinite objects and I also presented different views on the semantics of the omitted object. Now, another question arises about the nature of object-dropping verbs. When a verb participates in the so-called "implicit object alternation", is the transitive form of the verb actually distinct from the intransitive form in the lexicon, or are they different syntactic expressions of a single lexical entry? I will devote this Section to possible answers to this dilemma, to use Gillon's (2012) word. The issue of having two separate lexical entries or just a single entry for the transitive and the intransitive use of these verbs is not only relevant to develop a theoretical account of indefinite object drop, but it also has considerable effects on applied uses of this knowledge. For instance, McShane (2005, p. 118) observes that the choice between one or two lexical entries would have direct consequences on machine translation systems, in all cases when a transitive verb can be used intransitively in the source language but not in the target language (e.g., Russian *mešat'* 'to bother', which has to be translated in English as 'to get in the way' when objectless). In these cases, one should either posit two entries in the target language (and a rule to favor one or the other according to the presence or absence of a direct object), or have a single entry enriched with semantic information.

2.4.1 Two meanings, two verbs: the naive account

The problem of having a single verb exhibiting two syntactically different behaviors (transitivity and intransitivity) was first identified by Jerry Fodor and Janet Fodor (1980) and Dowty (1981), a reply to the former paper. Both treat verbs allowing for implicit objects as ambiguous between two different lexical entries, one transitive and one intransitive. This view is shared by other traditional literature on the matter (Brisson 1994; Cote 1996; Fellbaum and Kegl 1989; Mittwoch 1982; Van Valin and LaPolla 1997) and by more recent accounts (Bourmayan and Recanati 2013; Pethõ and Kardos 2006).

Such an interpretation is clean on the surface, as clear-cut binarisms often are, but it does little to describe the complexity of reality —again, as binarisms often do. In a broader theory of semantics, the problem of the two uses of a single verb mirrors the well-known problem of deciding, for instance, whether *bank* is a polysemous noun with two interpretations ("financial institution" and "river bank") or whether it has two homonymic interpretations. Say we go for the second, safer, account, since the only factor keeping the two senses together, i.e., etymology, is not transparent to native speakers of English nowadays. On the other hand, we would be much more keen to ascribe a polysemous interpretation to the different senses of the noun *man*, which depending on the context can be used to mean "human being", "male human being", or "adult male human being". Crucially, the different senses of *man* are all facets of the same entity, while the different meanings of *bank* are not. Going back to the issue at hand, i.e., the distinction between transitive and intransitive senses of a given verb, it would indeed seem that these senses capture different facets of the same action performed by the Agent, instead of being two totally different meanings. This interpretation, fully consistent with the hypothesis that transitivity is a prototype (refer back to Section 2.1), is further explored in Section 2.4.2 with reference to relevant literature.

2.4.2 One verb, two meanings: the state-of-the-art account

As just argued at the end of Section 2.4.1, transitive verbs admitting object drop are better interpreted as a single lexical entry with two different meanings, rather than two separate entries in the lexicon (an old-fashioned perspective that Lorenzetti (2008, p. 60) defines "counterintuitive and inappropriate"). In particular, far from being a "lexical quirk" of a restricted class of verbs, indefinite object drop appears to be "a syntactic detransitivisation mechanism" used to express events which do not embody the transitive prototype (Næss 2007, p. 134). Let us now discuss this behavior in more detail.

The syntax of indefinite null objects A brief syntactic detour is in order. While this dissertation is much more concerned with the effect that semantics (and, to a lesser extent, pragmatics) has on indefinite object drop, it is still important to take a position with respect to the syntactic nature of the omitted object. Is it absent from the syntax, as many used to argue (as seen at the beginning of this Section and later on Page 23)? Or is there a syntactic slot available for the omitted object, even though it has no phonological representation? Convincing arguments brought forth by the literature on the matter, as shown throughout this Section, make a strong case for the second hypothesis. For a discussion in favor of the syntactic representation of implicit arguments, touching topics that go beyond the scope of this dissertation, the interested reader can refer to I. Landau (2010).

In syntactic theory, Roberge (2002) (further explored by Cummins and Roberge (2004, 2005)) proposed an internal-argument equivalent of what the Extended Projection Principle (EPP) (Chomsky 1982) is for subjects, called "Transitivity Requirement". The EPP is the requirement for a subject position in the clause[17] , provided by Universal Grammar, which then gets filled in by lexical material in non-*pro*-drop languages (such as English) or by the empty category *pro* in *pro*-drop languages (such as Italian). Likewise, the Transitivity Requirement posits a direct object position in the clause, provided by Universal Grammar too, which accounts for implicit objects just like EPP accounts for null subjects. The only difference between the two is that, as noted by Cummins and Roberge (2004), "recoverability for the EPP is morphologically based, as is evident in null-subject languages, while recoverability involving the TR may also be semantically and pragmatically based". Null-subject languages[18] have "morphological recoverability" for their null subjects in that the required information is stored in verb morphology, as shown in my examples in (11) relative to Italian.

17: The subject position is spec-TP in languages where the subject raises to spec-TP after getting base-generated in spec-VP, or spec-VP in languages where no raising happens.

18: Let it be noted that here I am not referring to null-subject languages where the omission of the subject is a kind of Topic drop, not licensed by morphological recoverability, such as Chinese.

(11) a. Corr-o.
run.1.SG.PRS
'I run.'

b. Corr-iamo.
run.1.PL.PRS
'We run.'

Semantic and pragmatic recoverability of the implicit object licensed by the Transitivity Requirement has been partially shown in Section 2.3.2 and will be further discussed in the rest of this Section.

Focus on the activity Many authors (Ahringberg 2015; Fillmore 1986; García-Velasco and Muñoz 2002; Goldberg 2005a; Levin 1993; D. Liu 2008; Yasutake 1987) argue that the implicit indefinite object construction is used to focus on the activity itself, "downgrading the referential status of the object" (García-Velasco and Muñoz 2002, pp. 7–8). I provided an example of this on Page 19, where the verb *to drink* used intransitively was shown to refer to the habit of drinking alcoholic beverages. Thus, the focus of such utterances is not on the actual drink the subject is imbibing, but rather on the activity itself.
The word "focus" is not used idly here. Indeed the distinction between topic (the known, background information) and focus (the new, foreground information), central in pragmatic and discourse-oriented accounts of human language, also applies to the problem at hand. As argued by Lorenzetti (2008, p. 66), given that most sentences require at least one focus, and that the focus is by its very nature new, pragmatically non-recoverable information, it stands to reason that omitted objects (which are recoverable[19] and prototypical[20] , hence, known) cannot be the focus. Thus, the focus in such utterances has to be on the activity itself, as in her examples in (12).

19: Refer to Section 3.1.1.

20: Refer to Section 2.3.2.

(12) a. I thought you said your dog doesn't bite ∅!
b. Religion integrates ∅ and unifies ∅.

Goldberg (2005a) formalized this intuition via her Principle of Omission under Low Discourse Prominence, which states that the Patient argument of a transitive verb is possible when it is "de-emphasized/unprofiled in the discourse" (i.e., neither topical nor focal) and when the action, on the other hand, is "particularly emphasised". This shift in meaning, granted by the omission of the direct object, has been shown (Greene and Resnik 2009, p. 507) to trigger a reduced sentiment response in native speakers presented with pairs of sentences like the one in (13).

(13) a. At the same time, we should never ignore the risks of allowing the inmate to kill again.
b. At the same time, we should never ignore the risks of allowing the inmate to kill someone again.

Some authors (Groefsema 1995; Hall 2009; Iten et al. 2005; Recanati 2002; D. Wilson and Sperber 2000) even took advantage of this focus-on-the-activity interpretation to state that the omitted object is absent in the syntax just as it is absent phonologically, being instead pragmatically provided. Such an account is tempting and not inconsistent with Goldberg's omission principle, but (as I am going to argue right away) it is not convincing in the light of further linguistic evidence.

Indefinite object drop as noun incorporation Not only do (indefinite) object-dropping transitive verbs describe activities, they specifically describe "conventional, name-worthy, institutionalized, habitual activities" (Dvořák 2017b, p. 119). As observed by Bourmayan and Recanati (2013), Martí (2010, 2015), and Yasutake (1987), this is exactly the case of verbs having undergone noun incorporation[21] , a linguistic process "traditionally understood as the compounding of a noun stem with a verb stem to form a new verb stem" (Mithun 2009, p. 5). Syntactically, this can be seen as a kind of head-to-head movement, as depicted in Figure 2.1 (taken from Carnie (2012, p. 495), a simplified account of the account provided by Baker (1988)).

21: Noun incorporation is found in several polysynthetic languages, but it is not a requisite for polysynthesis, just as polysynthesis is not a requisite for noun incorporation (Mithun 2009).

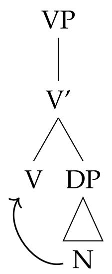

Figure 2.1: Portion of syntax tree illustrating the head-to-head movement involved in noun incorporation, from Carnie (2012, p. 495).

Without bringing systematically noun-incorporating languages into the discussion, such as the ones from the Iroquoian family (Mithun 2009), Frisian, or West Greenlandic (Martí 2015), it is possible to find such behavior in now-lexicalized object-verb compounds in English too. I illustrate the case of *to babysit* (also valid for *to birdwatch, to fingerprint*, and other such compounds) in Figure 2.2.

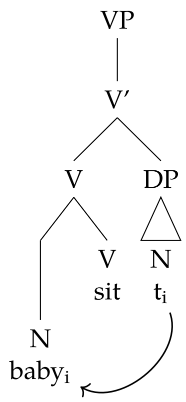

Figure 2.2: Portion of syntax tree illustrating the object-verb compound *to babysit* in English as a result of noun incorporation.

In Martí's (2015) account, then, the only difference between indefinite null objects and incorporated nouns such as *baby* in *to babysit* would be that the former are phonologically null, while the latter are not. Noun incorporation gives rise to several effects, which Martí (2015, pp. 455–456) reports based on evidence from West Greenlandic (an ergative language):

- incorporated nouns must be bare, i.e., with no preceding article/determiner (just like indefinite null objects were shown to be "weak indefinites" in Section 2.3.2);
- the subject of a noun-incorporating verb is marked with absolutive case like the subjects of intransitive (unergative) verbs and the objects of transitive verbs, while the subjects of transitive verbs are marked with ergative case (mirroring the intransitive behavior of object-dropping transitive verbs);
- incorporated nouns always precede the verb in the linear word order, regardless of the position a full-fledged direct object would take in the sentence;

- incorporated nouns are interpreted indefinitely and non-specifically (just like indefinite null objects);
- incorporated nouns take narrow scope with respect to other operators in the sentence (as argued about indefinite null objects in Section 2.3.2);
- verbs undergoing noun incorporation usually refer to "nameworthy, typical activities" (like indefinite null objects, as shown on Page 23);
- verbs undergoing noun incorporation tend to have "conventionalized meanings", as noted before on Page 19 relatively to the imbibe-alcohol conventionalized meaning of the verb *to drink* used intransitively.

Martí (2015, pp. 461–463) then proceeds to test the hypothesis that verbs allowing indefinite object drop and verbs allowing noun incorporation share common properties with evidence from Frisian, a Germanic language having both indefinite null objects and noun incorporation. Her analysis demonstrates that transitive-made-intransitive verbs and incorporated-into verbs indeed belong to the same class. A crucial property both types of verbs share is, for instance, that only verbs selecting for a Patient object and having a volitional subject can participate in these constructions (making it only possible for verbs such as *to notice, to hate, to know* in English to participate in *definite*, not *indefinite*, object drop).
Purely pragmatic accounts of implicit indefinite objects, as revealed on Page 23, fail to take into account the cluster of properties shared by noun-incorporating and object-dropping verbs alike. On a side note, Mittwoch (2005, p. 249) comments on a construction sharing similar properties with noun incorporation, i.e., the *out*-verb formation (e.g., *I don't think they can outbuild us)*, which is a productive process where the original object gets omitted to put the focus on the activity[22] and the "resulting form selects for an object that belongs to the same class as the subject" (typical of low-transitivity utterances, as per Hopper and Thompson (1980)).

22: In particular, on a competitive perspective about the activity, which is typical of commercials.

Affected Agents, effected Patients I will now tackle an aspect of indefinite objecthood originating directly from the account of transitivity as a prototype concept by Hopper and Thompson (1980) and later literature[23] , i.e., the need for the subject and the object of transitive sentences to be maximally distinct in their semantic behavior. This requirement was formalized by Næss (2007, p. 30) in the Maximally Distinguished Arguments Hypothesis. This observation led some authors[24] to posit change-of-state verbs such as *to break* as obligatorily transitive, since they feature[25] a volitional, unaffected Agent and a non-volitional, affected Patient. As argued time and again in this Chapter, indefinite object drop is far from being prototypical behavior for transitive verbs. Under this lens, then, it would make sense to find that indefinite null objects are more common with verbs having affected Agents (i.e., Agents being the endpoint of the event) and/or effected Patients (i.e., Patients brought into existence by the event the verb refers to). Such an analysis is discussed in detail in Næss (2007).
Let us consider affected Agents first. Tenny (1994, p. 158) calls them "measuring arguments", in that they delimit the event "by undergoing a change of state that marks the temporal end of the event". For instance, the event

23: Refer back to Table 2.1 and further considerations in Section 2.1.

24: Refer back to Section 2.3.1.

25: Ignoring uses where the subject is non-Agentive such as *The ball broke the window* and anticausative uses such as *The window broke*.

described by *John ate* would be delimited by the affectedness of the Agent (in this case, the feeling of fullness), rather than by the affectedness of the unmentioned Patient. On the opposite, if one wanted to focus on the affectedness of the Patient, they would then have to resort to a transitive use of the verb *to eat* (Næss 2007, p. 80). An affected-Agent interpretation of intransitive *to eat*[26] , which may or may not convince the reader yet by means of this example in English, becomes quite more persuasive in the account by Næss (2007, pp. 61–63) of the same verb in Yucatec, a Mayan language spoken in the Yucatán Peninsula. In this language, due to its Agent being affected, *to eat* patterns aspectually and morphologically with change-of-state verbs, not with activity verbs (as one would expect). Crosslinguistically (Næss 2007, p. 126), ingestive verbs (*to eat*, *to drink*, but also *to learn*) consistently show characteristics typically belonging to intransitive verbs (Amberber 1996, 2009), leading Marantz (1981) and subsequent decades of literature on indefinite object drop to consider them "class representatives" of the typical behavior of verbs allowing for indefinite object drop. Taking the affected-Agent interpretation to the extreme, in some languages (such as Korean and Turkish) *to eat* is even used as a grammaticalized marker of Agent affectedness, e.g., as an auxiliary, as a light verb, in constructions where it expresses undergoing or adversativity (Næss 2007, p. 75), and in antipassive constructions[27] (Næss 2011, p. 414). Nicolas (2019) offers an interesting account of the features null objects in English share with antipassive constructions. Moreover, an affected-Agent account can be easily employed to explain linguistic behavior that would otherwise remain unmotivated, such as the resistance of the verb *to lock* to object drop as opposed to the ease one finds in using *to eat* without an overt object, as noted by Pethõ and Kardos (2006, p. 30). They argue that the opposite behavior of these two verbs with respect to indefinite object drop "does not become clear", on the basis of their selection restrictions[28] having comparable extent (i.e., the objects of *to eat* are edible items, the objects of *to lock* are items provided with a lock). The reason for this difference, however, becomes quite clear when one considers that the subject of *to eat* is an affected Agent, while this does not hold true for *to lock*. However, it should be noted that while this analysis works, allegedly, for English, it is not universally valid for other languages. For instance, *chiudere (a chiave)*, the Italian equivalent of English *to lock*, can also be used intransitively, at least in spoken language (e.g., *Hai chiuso?* 'Did you lock?', asked to someone leaving their house). Isingoma (2020) observes that object drop is also possible with *-siba*, the equivalent of English *to lock* in Rutooro[29] . Thus, Agent affectedness is a relevant facilitator of object drop, but its role has to be put in perspective (refer also to Section 3.1.2 for more considerations on this).

It is also important to note that while Agent affectedness is inherent to the semantics of some verbs (such as ingestion verbs), it may also be activated by verb-external elements of a clause. Let us consider the example sentence *John murdered for the money* (Næss 2007, p. 136). In this case, the affected-Agent interpretation is fostered by the purpose clause *for the money*, since the Agent's motive for the homicide is a direct gain, i.e., something that positively affects the Agent. Finally, the affected-Agent account can also explain constructions of the type *have a drink, have a lick, have a bite* (Wierzbicka 1982, pp. 758, 771). The author argues that in such constructions, *have* has a detransitivizing function in that it backgrounds the object while focusing on the Agent. Not only that, but also, *have a*

26: Such an account was first offered, *in nuce*, in Haspelmath (1994), Nedjalkov and Jaxontov (1988), Starosta (1978), and Wierzbicka (1982).

27: Used in ergative languages to convey certain aspectual and modal nuances. They are called "antipassives" because, much like in passive sentences the object of the active sentence becomes the subject and the subject of the active sentence is deleted, in antipassive constructions the object is (usually) deleted and the subject changes case from ergative to absolutive case.

28: More on selectional restrictions and object recoverability in Section 3.1.1.

29: A language of the Bantu family, spoken in western Uganda.

[verb] events deviate from the transitive prototype because their Agent is affected by it (typically, by enjoying the activity the verb refers to), while the Patient is minimally affected.

Let us now discuss the other kind of argument deviating from the transitive prototype, i.e., effected Patients. These are non-affected objects that come into existence thanks to the very action described by the verb, and only if this action is brought to completion, e.g., the letter in *John is writing a letter* or the cake in *John is baking a cake*. Such constructions tend to feature indefinite null objects crosslinguistically and to have unaffected Agents, making it necessary to provide a different analysis than before in this Section (Næss 2007, pp. 127–128). What affected-Agent and effected-Patient constructions have in common is that they both show low semantic distinctness between Agent and Patient (making them the optimal environment for felicitous indefinite object drop, based on Hopper and Thompson (1980)), and they both largely depend on the semantics of the verb itself. As Næss (2007, p. 127) and Næss (2011, p. 421) observe, the low-distinctness of effected Patients is so embedded in the verb semantics that the very intransitive use of an effected-Patient verb evokes the non-referentiality of the object. This is most evident in imperfective contexts[30] (e.g., *John was writing*), where the effected Patient is not presented as fully effected yet, and thus it is even less prominent in the discourse. On the contrary, perfective contexts tend to block indefinite object drop with effected-Patient verbs (e.g., *? John had written*), while the same does not usually hold for affected-Agent verbs (e.g., *John had eaten*). Crucially, an effected-Patient account can be used to explain why, as first observed by Fillmore (1986, p. 96), intransitive *to bake* in English (e.g., *I spent the afternoon baking*) can only be understood to refer to the act of baking "bread or pastries, but not potatoes or ham". Næss (2007, p. 135) easily explains this linguistic fact by observing that the bread-or-pastries interpretation features an effected Patient, while the potatoes-or-ham interpretation features an *affected* Patient, which makes the verb prototypically transitive and, thus, resistant to indefinite object drop. On the same note, intransitive *to paint* (e.g., *He paints*) is interpreted to refer to the act or habit of painting pictures, not house walls.

30: I will come back to the role of (im)perfectivity as a factor determining indefinite object drop in Section 3.2.2.

2.5 A working definition of "indefinite object drop"

This dissertation is about indefinite null objects, whose nature I discussed throughout this Chapter with reference to traditional and recent literature alike. It is time to end this Chapter with some more detail on what this thesis is going to focus on, and also on what is going to be ignored in my experimental account of indefinite object drop.

As I mentioned in Section 2.2.2, I assume there is a distinction between *indefinite* object drop, motivated by linguistic factors (see Chapter 3 and Chapter 6) of varying nature, and other kinds of object drop, such as definite null objects depending on Topic drop, genre-based null objects, and implicit objects depending on extra-linguistic context. This choice is motivated by lexical, semantic, and pragmatic accounts of the indefinite object drop discussed in this Chapter. Crucially, while I assume that verbs participating in the indefinite object drop construction show two surface sentence structures while being single lexical entries (Section 2.4), I will not take a stance with respect to the problem of whether indefinite and definite object drop form part of a continuum or are two discrete, binary phenomena (Section 2.2). Indeed, one could interpret the distinction between the different kinds of object drop as stemming from a situation like the one in Figure 2.3, where virtually any verb can participate in object-dropping constructions, but in different ways depending on the underlying factor.

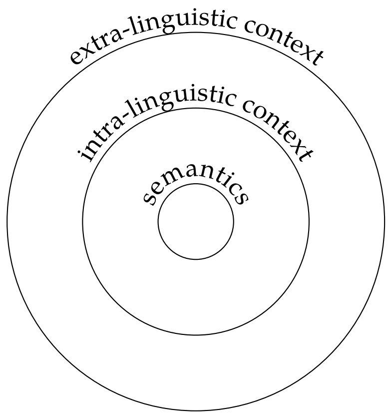

Figure 2.3: A concentric view of the factors determining the possible continuum between definite and indefinite object drop.

Let us interpret the examples in (14) in the light of Figure 2.3. In (14-a), the referent of the omitted object (gas) is supplied by pragmatics —if one wanted to echo Glass (2020), it is recoverable because in the sub-community of car-drivers it is customary to say cars "drink" to refer to them burning fuel. In (14-b), the definite null object is immediately recoverable from the linguistic context provided in the sentence, and no additional extra-linguistic context is needed to interpret this utterance. Instead, the indefinite null object in (14-c) is interpretable with no context whatsoever, given that the semantics of the verb (and of the Agent, as argued on Page 25) are sufficient for its recoverability.

(14) a. It drinks ∅ a lot!
(in the social context of someone speaking about a car)
b. # My milk has been opened, who drank ∅?
c. He drinks ∅.

One could, naturally, have qualms with respect to the full grammaticality of (14-b), since English, after all, is not typically a language allowing definite null objects outside genre-specific environments. Moreover, one may want to keep verb-specific semantic affairs separate from extra- and intra-linguistic contextual information. In this case, the concentric, continuous view depicted in Figure 2.3 would not correspond anymore to the theory, and one would have to adopt a strictly binary definite-or-indefinite perspective. Such an account easily handles cases where the same verb can appear in indefinite object drop constructions when proper situational context is provided, but not in zero-context environments, as in (15).

(15) a. Do you even lift ∅, bro? (Glass 2020, p. 9)
(common among people training for strength in gyms)
b. *John had lifted ∅.

What is important, here, is that the same factors acting on no-context indefinite null objects are also active in context-rich scenarios, regardless of the continuous or binary view one adopts. What matters is that additional context (such as the one granted by a sub-genre) makes it possible for a verb to license implicit indefinite objects more freely than in no-context utterances. Thus, by studying the factors determining indefinite object drop in no-context utterances, I am still providing useful information to understand context-depending object drops, although without committing to a specific interpretation of the distinction between implicit indefinite objects and other null objects.
To sum up, my stance is that:

- it is possible to characterize *indefinite* null objects, whether they form part of a continuum having definite null objects at the other end or not (e.g., in a binary account of this distinction), as discussed in Section 2.2;
- indefinite object drop is possible both with change-of-state verbs and with incremental-theme verbs (refer to Section 2.3.1), to different extent, and the understood omitted object is not *something* but the most prototypical Patient for a given verb sense in a given context (refer to Section 2.3.2);
- rather than positing two separate lexical entries for the overt-object and null-object uses of a transitive verb allowing indefinite object drop, a better account would have a single entry in the lexicon for the verb, which would then admit a null object under specific circumstances (refer to Section 2.4);
- the same semantic and aspectual factors allowing for indefinite object drop in no-context utterances are active in context-rich utterances, where context can be provided by linguistic means or via community-specific world knowledge (refer to Section 2.5).

I will explore the main semantic, aspectual, and pragmatic factors playing a role in indefinite object drop in Chapter 3. Clearly, all these perspectives have to be interpreted in the light of unavoidable lexical idiosyncrasies, such as the difference (noted on Page 26) between English *to lock*, blocking object drop, and its equivalents in Italian and Rutooro, allowing object drop. In order to compare my models of indefinite object drop in English and in Italian, I will design my experiments in such a way as to minimize such semantic differences between my target verbs in English and their Italian counterparts (see Chapter 7).

Factors allowing indefinite object drop

3

In Chapter 2 (*Indefinite object drop*) I presented indefinite object drop as a marked construction deviating from the transitive prototype. Indefinite object drop has been argued to be binarily distinct from definite object drop, to be possible with some or all transitive verbs, to imply an unsaid *something* or a prototypical object different for each verb, to make the verbs participating in this construction have an additional, intransitive entry in the lexicon or just the transitive one. Several answers were proposed to these conundrums in the literature, and I provided my own perspective.
In this Chapter, I am going to focus on the main intra-linguistic (semantic, aspectual, and pragmatic) factors allowing a transitive verb to participate in the indefinite object drop construction, based on literature on this topic. I will return to the subject of recoverability, manner specification, telicity, perfectivity, and iterativity in Chapter 6 (*Linguistic factors used as predictors*), where I will define them as predictors of indefinite object drop in my Stochastic Optimality Theoretic model.

3.1 Semantic factors

3.1.1 Recoverability

An intuitive notion of recoverability As observed several times in Chapter 2, recoverability is the *sine qua non* of object omission. Cote (1996) even went as far as to identify it as "the only absolute constraint on null arguments". Intuitively, this notion can be used to tell apart definite implicit objects (whose meaning is recoverable from context, be it extra- or intra-linguistic) and implicit indefinite objects (whose meaning is recoverable from the semantics of the verb itself). Even authors who cautiously (Olsen and Resnik 1997; Resnik 1993, 1996) or openly (Glass 2013, 2020, 2022) reject a distinction between definite and indefinite object drop still maintain a certain notion of recoverability as a fundamental requirement for object drop. Let us discuss this in some more detail.
Decades of literature on the matter teem with pre-theoretical definitions of object recoverability. The oldest is in Jespersen (1927, p. 321), where the author argues that "the omission of an obvious object probably produces more intransitive uses of transitive verbs than anything else". Later on, U. Ohlander (1943, p. 105) observed that object-less utterances "may appear complete enough" by virtue of the fact that "the element to be understood or supplied is so self-evident that the gap is mentally filled in by the audience more or less unreflectingly". Like many others, Ohlander makes reference to the notion of recoverability without using this exact wording. Similarly, Hickman, J. Taylor, and Raskin (2016) resort to the idea of "conceptual defaultness", i.e., the property of unmentioned direct objects whose omission does not depend on an informational failure on the part of the speaker/writer —on the contrary, defaults are omitted to

comply with the Gricean maxim of quantity, since their very mention "would be unnecessary and, perhaps, awkward" (Hickman, J. Taylor, and Raskin 2016, p. 516).
Interestingly, recoverability appears to be such an intuitive notion as to become a major determinant of argument omission (not limited to indefinite object drop) in the early stage of grammar acquisition crosslinguistically, even in context and languages where adult grammar would normally prohibit it (Allen 2000; Ingham 1993; Medina 2007; O'Grady, Yamashita, and Cho 2008; Pérez-Leroux, Pirvulescu, and Roberge 2011, 2018; Pérez-Leroux, Pirvulescu, Roberge, and Castilla 2013; Rasetti 2003; Ratitamkul, Goldberg, and Fisher 2004; Sopata 2016). Young children are shown to omit objects (and other arguments) of verbs they are exposed to, both in transitive and intransitive utterances, if extra-linguistic context makes them sufficiently recoverable.

Between lexical and contextual recoverability In Kardos's (2010) words, omitted objects are recoverable "either through lexical stereotypes or based on the context" (Kardos 2010, p. 7). In a sense, it is not even necessary to postulate a binary distinction between the two facilitators of recoverability, since "lexical stereotypes" (i.e., the selectional preferences of a verb) descend from world knowledge, situational context is where meaningful conversations happen, and narrower context enables more object omissions without contradicting world knowledge (more on this in Chapter 2, and in Glass (2013, 2020, 2022), where recoverability is intended as a matter of degree). A prime example of this is (1), taken by Bergh and S. Ohlander (2016, p. 24), whose full interpretation depends on additional context. Indeed, this sentence in isolation has no unique interpretation. *What* did they play beautifully? Was it an instrument or a game? And what kind, exactly? However, while context would make it possible to know the referent of this implicit indefinite object, the semantics of the verb (in particular, its selectional restrictions) still provide us relevant information without the need for additional context. Indeed, we know that they played either a game, or a musical instrument, or a role in a theater piece. This would not be possible, for instance, with a selectionally un-restricted (or, better, very loosely restricted) verb such as *to make*.

(1) They played beautifully.

This also holds true for verbs with much stricter selectional preferences than *to play*, e.g., the verb *to eat*, as discussed in Chapter 2. As noted by Cote (1996, p. 149) among others, intransitive *to eat* tends to refer to a meal (which is the prototypical item humans eat), but it does not have to. In context-rich utterances, the actual referent may be different, depending on "the underlying context and intentional structure of the discourse structure at the time of utterance". Crucially, context may make it obvious that the omitted Patient is some specific kind of edible item (e.g., pasta, hamburgers, or even something as extravagant as Hawaiian pizza), but it cannot make it deviate from the basic selectional preferences of the verb[1] —the omitted Patient has to be something edible.
Semantically-licensed recoverability also interacts with world knowledge, specifically with societal norms, in Goldberg (2005a). In this case, the author argues that politeness is a driving factor in the omission of direct

1: Unless, of course, the verb takes a specialized meaning in a specific sub-genre, e.g., *to eat* in a game of chess would mean *to capture enemy pawns*.

objects occurring with verbs of bodily emission, since they are very imageable (hence, recoverable), but also very taboo in usual social contexts. On the other hand, it would be easy to imagine that in contexts where such verbs are not taboo (e.g., in medical or adult-only literature, as a future corpus study could ascertain) the main driver of object omission would be contextual inferability, rather than misguided concerns for politeness. Either way, context and world knowledge (about the verb itself, but also about proper customs) comply with verbal semantics.

Consistently with all the previous observations on recoverability, Mittwoch (2005) and Glass (2013) also observe that the referent of implicit indefinite objects corresponds to the literal meaning of the verb, rather than to metaphorical or idiomatic meanings. For instance, intransitive *to read* refers to "written or printed material rather than, say, the stars or coffee grounds" (Mittwoch 2005, p. 2). Likewise, when a verb has strict selectional preferences (e.g., *to eat* selects for edible items) and one of its near-synonyms has broader preferences[2] (e.g., *to devour* selects for edibles, but also for metaphorical items such as books), direct objects are much more likely to be dropped with the former than with the latter (Glass 2013, p. 5). Clearly, this mechanism is in place to maximize recoverability, since the literal selectional preferences of a verb are known and predictable (based on lexical semantics and world knowledge), while its metaphorical or idiomatic behavior is largely arbitrary and unpredictable. More in general, the more an object is semantically dependent from a verb, the more likely it is to be omitted (Rice 1988, pp. 203–204).

2: An in-depth discussion of such verb pairs will be tackled in Section 3.1.3.

Semantic selectivity as a proxy to recoverability Let us now look in more detail into semantic selectivity as the main verb-internal, context-independent factor allowing for object recoverability. So far in this Section, I made the case that knowing the specific type of objects a verb favors in its selectional preferences is the first step towards object recoverability and, consequently, felicitous object omission. World knowledge and context shape the way we process object-less transitive verbs when accessing their selectional preferences. However, as hinted before, there is a close correlation between indefinite object drop and the breadth of a given verb's selectional preferences, regardless of the actual items or family of items it tends to occur with (García-Velasco and Muñoz 2002; Glass 2020; D. Liu 2008; Maouene, Laakso, and L. B. Smith 2011; Medina 2007; Olsen and Resnik 1997; Resnik 1993, 1996). The intuition behind the use of a verb's selectional preferences as a means to gauge the recoverability of its objects stems from the observation that implicit indefinite objects "are clearly understood because they are inferred from a very narrow, if not exclusive, range of possibilities" (García-Velasco and Muñoz 2002, p. 4). This is the reason why native speakers of English are more likely to find indefinite object drop grammatical with *to read* than with *to know*, since there are way fewer readable than knowable things in our lives (D. Liu 2008, p. 302).

Resnik (1993, 1996) was the first to provide a more data-grounded definition of recoverability, by means of a computational model of a verb's selectional preferences. In particular, his Selectional Preference Strength taxonomy-based measure is shown to be inversely proportional to the semantic narrowness of a verb's selectional preferences, so that a verb will receive a higher score if its direct objects are semantically similar (e.g., *to eat*, *to read*), and a lower score if they are semantically different (e.g., *to make, to know*). I will present the mathematical details of Resnik's measure, discuss the implications of such an approach, and propose my own distributional semantics-based alternatives in Section 6.1, where I expand upon both my Preference In Selection of Arguments (PISA) computational measure presented in Cappelli and Lenci (2020) and on my behavioral variant of Computational PISA, inspired by Medina's (2007) Object Similarity measure. An important implementation-related aspect to note here, common to all these measures of semantic selectivity used as proxies to object recoverability (Resnik's SPS, Medina's OS, my own PISAs), is that recoverability is modeled as being gradient, and these models capture the semantic narrowness/breadth of the semantic categories[3] the potential direct objects of a verb belong to, rather than focusing on the specific objects themselves.

Most importantly, with his computational experiment Resnik (1993, p. 88) could conclude that recoverability, as quantified via gradient semantic selectivity, is a necessary (albeit insufficient) condition for object omission. This conclusion, consistent with pre-theoretical intuitions about recoverability and previous theory-informed, non-experimental statements about the role of selectional preferences in determining object recoverability, stems from the observation that object-less transitive verbs never receive low semantic selectivity scores[4] in Resnik's experiment.

3: This is less true for semantic similarity measures based on behavioral judgments, such as Object Similarity and Behavioral PISA, but still not completely off the mark. Please refer to Section 6.1 for a full account of such measures.

4: However, some verbs high in semantic selectivity fail to license indefinite object drop. Resnik explains this apparent failure of his measure with reference to the aspectual properties of such verbs. I will come back to this in Section 3.2 and Chapter 6.

3.1.2 Agent affectedness

Back on Page 25, I argued with plenty of references to relevant literature (first and foremost, Næss (2007)) that verbs whose Agent is in some ways affected by the action (e.g., *to eat, to learn*) described by the verb tend to be more likely to participate in the implicit indefinite object construction than unaffected-Agent transitive verbs (e.g., *to kill, to break*). This is a direct consequence of the need for the arguments of prototypical transitive verbs to be maximally distinct from a semantic point of view, captured in Hopper and Thompson's (1980) parameters H (agency) and I (affectedness of the object)[5] , and later on in Næss's (2007) Maximally Distinguished Arguments Hypothesis. Malchukov (2006, p. 335) captures the same intuition in his Relevance Principle, stating that Hopper and Thompson's (1980) transitivity parameters have to be marked on the relevant constituent (e.g., volitionality on the Agent, affectedness on the Patient) in prototypical transitive clauses.

5: Refer back to Table 2.1.

I will not go again over the affected-Agent account of object drop, since I already discussed it in Chapter 2. For the purposes of the review of the main factors facilitating indefinite object drop provided in this Chapter, I will just mention that Agent affectedness can manifest in two ways. One is inherent to the semantics of the verb, as it happens with ingestion verbs such as *to eat*, *to drink*, and, in a sense, *to learn*. The other is instead context-dependent, as in (2), where the verb *to murder* gets an affected-Agent interpretation and thus participates in felicitous object drop due to the purpose clause *for the money*.

(2) John murdered for the money. (Næss 2007, p. 136)

It is important to note, as seen before in this Chapter and Chapter 2, that the context enables verb behavior that is already possible, virtually, thanks to the semantics of the verb itself. After all, does a murder not affect the murderer even without mention of the cause? Stating it explicitly puts the focus on the motive, putting in the background both the Patient (which was already backgrounded, due to being unmentioned) and the murdering activity itself (which would be the focus of cause-less *John murdered*). Crucially, I would like to point out that it is not possible to use purpose clauses to induce an affected-Agent reading on object-less transitive verbs lacking this possibility in their semantics, e.g., the verb *to build* in (3).

(3) *John built for the money.

Indeed, (3) would only be considered grammatically acceptable provided it is inserted in a larger context (just like plain **John built*). However, one could object to this account of the verb *to build* making reference to the effected-Patient account I discussed in Chapter 2 together with the affected-Agent account —the object of *to build* comes into existence via the act itself of building it, unlike the Patient of *to murder* (which exists before the murder, and ceases to do so due to it). Why does the verb *to build* not allow for indefinite object drop, even though it is a handbook effected-Patient verb just like *to bake*? Based on Goldberg (2001, p. 512) and Næss (2007, p. 139), *to bake* and *to build* are actually more different than it would seem at first glance, since verbs like the latter (e.g., *to break*) refer to events whose interpretation strictly depends on the Patient itself. In other words, while it is possible to imagine a baking event without having a precise baked good[6] in mind, it is impossible to picture a breaking or building event without having a precise broken or built object in mind. Without going into idiom territory, where one could "break a bank note" to get change or "build one's hope", it is clear that the act of breaking a glass is quite different from the act of breaking a leg, just as building a sand castle is quite different from building an airplane. Thus, these examples go to show that recoverability (introduced in Section 3.1.1) is the preminent factor in determining object drop, and neither Agent affectedness nor Patient effectedness cannot overpower it.

6: A baked good which, we will remember from Chapter 2, is much more likely to be bread or some pastry rather than rotisserie chicken.

3.1.3 Manner specification

Introduction Manner specification is a tricky semantic predictor of indefinite object drop to define, due to the different interpretations the concept of "manner" received[7] from different authors. This word is used, fundamentally, in two different ways:

- to refer to "semantically marked" counterparts of "semantically neutral" verbs, e.g., *to devour, to nibble* with respect to *to eat* (Fellbaum and Kegl 1989; Næss 2007; Rice 1988);
- in contrast with "result", to separate "manner" activity verbs such as *to sweep* from "result" causative verbs such as *to break*[8] (Beavers 2013; Beavers and Koontz-Garboden 2012, 2017; Levin and Hovav 2008; Melchin 2019; Rappaport Hovav and Levin 1998, 2005, 2010).

Keeping these two senses apart is important to avoid drawing forced conclusions about the nature of verbs exhibiting "manner" components

7: For an extensive discussion on the concept of "manner", touching several aspects that are beyond the scope of this dissertation, refer to Stosic (2019, 2020).

8: The concept of "manner", here discussed in relation to transitive verbs, is also central in studies about motion verbs (Beavers, Levin, and Wei Tham 2010; Cennamo and Lenci 2019; Iwata 2002).

in their meanings. García-Velasco and Muñoz (2002, p. 7) run exactly into this problem when they argue that "manner-of-ingesting verbs may be the exception to the rule", namely, the fact that manner-of-action verbs allow for object drop in Rappaport Hovav and Levin (1998) while result verbs do not. Indeed, they acknowledge that "both Rice (1988) and Fellbaum and Kegl (1989) suggest that the presence of a manner component in manner-of-eating verbs accounts for the impossibility of omitting the object", but they fail to recognize that these authors are using "manner" in a very different sense from Rappaport Hovav and Levin. In a way, these two senses are so different to become almost opposites, given that manner-specified verbs being semantically marked counterparts of other verbs would be the least manner-y of all in the account interpreting "manner" as opposed to "result", since they also encode a result component (as I will show in this Section).
In this thesis, as I will also argue in Section 6.5, I am only interested in the first sense of the word "manner". However, since the two senses overlap in significant ways, despite their fundamental difference, I will also comment now the second sense in some detail.

"Manner" as opposed to "result" With respect to indefinite object drop, Levin and Hovav (2008) and Rappaport Hovav and Levin (1998, 2005, 2010) argue that verbs expressing manner in their meaning, such as *to eat*, are much more likely to allow for object drop than verbs expressing result, such as *to devour*. In particular, *to devour* is considered a result verb because it entails complete consumption of the Patient[9] by the Agent, unlike *to eat*, at least in an unmarked, uninterrupted scenario (Melchin 2019; Piñón 2008; Smollett 2005).
This idea, which Goldberg (2001) and Onozuka (2007) oppose on the basis of the aforementioned Principle of Omission under Low Discourse Prominence, is famously exemplified by Rappaport Hovav and Levin (1998) with the examples in (4). The rationale behind this account is that result verbs specify scalar change (see (4-b)), while manner verbs specify non-scalar change (see (4-a)). Crucially, the entity changing along the scale specified by result verbs (i.e., the Patient object) is argued to be ungrammatical to omit, giving rise to a test for result-lexicalization used by Beavers and Koontz-Garboden (2012) and Rissman (2016).

(4) a. Phil swept.
 b. *Tracy broke.

In such an account, manner and result are to be considered complementary, in that a verb can only lexicalize one of them. However, Levin and Hovav (2008) also observe that there is some understood manner component in many result verbs (e.g., the result of *to clean*[10] is achieved by acting in a specific manner), and likewise, some understood result component in many manner verbs (e.g., *to scrub* is a manner of cleaning, that will likely generate cleanliness of a surface as a result). This particular observation serves to bridge the gap between this account, where lexicalized manner leads to felicitous object drop, to the other account, where an overt, specified manner component blocks object drop. Beavers and Koontz-Garboden (2012, p. 5) even make the case that so-called "poison verbs", a sub-class of manner-of-killing verbs identified by Levin (1993, pp. 230–233) in opposition to "murder verbs", actually entail both manner

9: This analysis would have manner specification be collinear, or at least highly correlated, with telicity as Olsen (1997 [2014]) (whose account is used in the experimental setting by Medina (2007), which is also mine) intends it. In Section 6.5 I will demonstrate that this is not the case at all.

10: Melchin (2019, p. 52) actually argues that "dynamic verbs" such as *to clean* are specified for neither manner nor result. However, delving into this debate would not bring my own argumentation further.

and result, *contra* Levin and Hovav (2008) and Rappaport Hovav and Levin (1998). Crucially, Melchin (2019, pp. 71, 89) adds to this by arguing that also *to devour* entails both manner and result, because the Agent acting in a very specific manner brings forth the "manner" interpretation, while the "scalar change affecting another participant" brings forth the "result" interpretation (which was also in Levin and Rappaport Hovav's original proposal).

Introduction to "manner" as "semantic narrowness" The other interpretation of the concept of "manner", which is the one I employ in my experiments (Chapter 6 and Chapter 7) and probabilistic model of object drop (Chapter 8 and Chapter 9), is offered by Rice (1988), Fellbaum and Kegl (1989), and Næss (2007), among others. These authors argue that the impossibility of omitting the object with verbs like *to devour* is explained by the presence of a manner component in their meaning (see also García-Velasco and Muñoz (2002) for further considerations). In other words, while *to eat* (which Rice (1988) calls a "semantically neutral" verb) is a base verb referring to a general activity, *to devour* (which Rice (1988) calls an "action-plus-manner" verb) has an additional manner specification in that it refers to a particular *manner* of eating. Most importantly, Melchin (2019, pp. 49–50) shows that this distinction holds crosslinguistically with examples in French, Dutch, and Arabic. The same also goes, for instance, for *to guzzle, to chug* with respect to basic *to drink*, and moving from transitive to motion verbs, for *to saunter, to stride* with respect to basic *to walk*.

Links between manner specification and Agent affectedness Næss (2007, p. 139) provides an intriguing link between manner specification and the affected-Agent account discussed in Chapter 2 and Section 3.1.2. In particular, she observes that indefinite object drop is infelicitous with manner-specified verbs because they typically refer to the way in which the Patient (crucially, not the Agent) is affected, in true prototypical transitive behavior. However, manner specification in the verb root stops being an obstacle to object drop if proper context is provided to imply Agent affectedness, as in her example in (5).

(5) The dinner was delicious, but Jane had no appetite and only nibbled.

Bringing Agent affectedness into the equation can also solve a decades-old conundrum by Fellbaum and Kegl (1989). Why do *to mush, to nosh, to graze* allow for implicit indefinite objects while *to gobble, to gulp, to devour* do not, despite them all being manner-specified troponyms[11] of *to eat*? Based on everything I observed about manner specification so far, both the first and the second group of verbs should block indefinite object drop on the basis of their manner component. In their taxonomic account, Fellbaum and Kegl (1989) explain this issue by positing two lexical entries for *to eat*[12] , one meaning roughly "to eat a meal" and another meaning "to ingest food". The first entry would be intransitive, and its manner-specified troponyms (*to mush, to nosh, to graze*) are too. The second entry would instead be strictly transitive, and its manner-specified troponyms (*to gobble, to gulp, to devour*) are too. As I argued in Section

11: I am using terminology from Fellbaum and Kegl (1989). Troponymy is a relation among verbs akin to what hyponymy is for nouns, "although the resulting hierarchies are much shallower" (G. A. Miller 1995).

12: Refer back to Section 2.4 for a discussion of whether or not to have two separate lexical entries for transitive verbs used transitively and intransitively.

2.4, positing two separate lexical entries does little more than restate the problem, without actually providing substantial explanatory power to the discussion. I would instead explain the difference in transitivity between these two groups of manner-specified troponyms of *to eat* with reference to the affected-Agent analysis. In particular, verbs like *to mush, to nosh, to graze* are activity verbs with a clear focus on the way the action affects the Agent (just like plain *to eat*), and indeed they do allow for their object to be dropped. On the contrary, verbs like *to gobble, to gulp, to devour* tend to highlight the affectedness of the Patient, making it necessary to express it overtly with a direct object in the syntax.
A link between manner specification and Agent affectedness also emerges from Lemmens's (2006) corpus analysis of verbs of killing in English, such as *to kill, to murder, to execute, to assassinate, to massacre*. He finds that, while *to kill* is not unlikely to be used intransitively in several corpora, the same does not hold for the other verbs of killing, which never occur with null objects. He attempts an explanation by observing that manner-specified verbs of killing may have "a stronger Patient-orientation, as they incorporate a more salient reference to Patients that are considered important in some socio-economical context (*to assassinate*) or to a high number of Patients (*to massacre*)". Flipping this perspective, *to kill* could be argued to be more likely to license object drop because it projects Agent affectedness more strongly than Patient affectedness, if compared to the other verbs of killing. Even disregarding this possible explanation of Lemmens's (2006) findings, they still confirm the relevance of manner specification in the implicit indefinite object construction.

Recoverability explains failures of a manner-based account Interestingly, Rice (1988, p. 207) notes in passing that "verbs that are very neutral, but that furthermore sustain a wide variety of complements, tend always to require objects", considering the ungrammaticality of intransitive *to love* as an example. Once again, object recoverability (as an effect of a transitive verb's semantic selectivity) is shown to be prominent with respect to other drivers of indefinite object drop. Recoverability also accounts for some examples Jackendoff (2003, p. 134) cites as "immediate counterexamples" to the idea that manner specification is relevant for the issue of argument drop, such as *serve/give the food to Sally* as opposed to *serve/*give the food*, and *insert/put the letter in the slot* as opposed to *insert/*put the letter*. These examples feature the omission of a Recipient/Goal instead of the omission of a Patient (which I am focusing on), but the principle holds in this case too. Indeed, *to serve* is a manner-specified troponym of *to give* ("a more specific form of giving", in Jackendoff's words) and *to insert* is in the same relation with respect to *to put*, but their hypernyms select for a much wider range of arguments, making it much more difficult for a speaker to recover them if they are unexpressed. Jackendoff also brings two other examples to his argumentations, this time relative to direct objects, i.e., *juggle (six balls), flirt (with Kim)*. The two verbs are shown to be grammatical both in their transitive and in their intransitive use. Since the author calls them "highly specific verbs", it seems that he is conflating both semantic selectivity and manner specification into the same label of "semantic specificity". I argue that these two properties have to be kept separated instead, and that whenever manner-specified verbs allow for their object (or other internal argument) to be dropped, they do so by virtue of the high recoverability of the intended object/argument (stemming from the high semantic selectivity of the verb).

3.2 Aspectual factors

3.2.1 Telicity

Vendler's aspectual classes Telicity, as a component of "lexical aspect", is a well-known predictor of indefinite object drop. As shown in Section 2.1, Hopper and Thompson (1980) included it in their ten-parameter account of transitivity as a prototype concept, associating telic aspect with high transitivity and atelic aspect with low transitivity. Telicity is one of the facets of so-called "lexical aspect", first organized by Vendler (1957) into four *Aktionsarten* ("types of action") as in Table 3.1, to which C. S. Smith (1991) then added a fifth class of "semelfactives"[13] .

Table 3.1: *Aktionsarten* as defined by Vendler (1957), plus semelfactives.

	punctual	durative
telic	achievement (e.g., *to find*)	accomplishment (e.g., *to build*)
atelic	semelfactive (e.g., *to knock*)	activity (e.g., *to run*)
stative	-	state (e.g., *to know*)

13: The term was actually coined by Comrie (1976, p. 42) to refer to "a situation that takes place once and once only", such as "one single cough", abiding by Latin etymology (*semel*, 'once'). He observes in a footnote that in Slavic linguistics the equivalent of this term was used to refer both to proper semelfactives and to "clearly iterative" utterances such as *He coughed five times*. More recently, the label "semelfactive" became widespread in linguistics in the imprecise sense attributed to it by C. S. Smith (1991), i.e., referring to atelic punctual events such as *to knock, to cough*. However, the actual etymological meaning of semelfactive would fit for any event that occurs only once, independently of its duration.

To simplify a very complex issue, one might observe the following. The first distinction to be made is between states, which cannot be used in progressive aspects (e.g., **John is knowing*), and the other categories. The two dimensions along which non-stative verbs vary are durativity and telicity. Durativity, which is quite self-explanatory, was also included among the ten transitivity parameters by Hopper and Thompson (1980), with punctual verbs being high in transitivity and durative verbs being instead low in transitivity. The other dimension, i.e., telicity, is defined as the property of having an endpoint of some kind. Crucially, literature on telicity tends to envision it as a property of *predicates*, not just of verbal heads (Hopper and Thompson 1980, p. 270). Thus, a durative activity such as *John is running* can be made into an accomplishment by specifiying a terminal point for the event, as in *John is running home*[14] , and *vice versa*. Under this account, then, intransitive uses of transitive verbs would just be transforming accomplishments (durative and telic) into activities (durative and atelic), as in *John is smoking (a cigarette)*. Such an account was explored by Mittwoch (1982) with respect to the intransitive uses of transitive verbs.

14: Please refer to Cennamo and Lenci (2019) and Cappelli, Bertinetto, and Lenci (2019) for considerations on the argumenthood of the added locative phrase to motion-verb activities in Italian.

One could be tempted to say that only activities are involved in the implicit indefinite object construction, considering that, based on three pieces of evidence I discussed so far,

- object drop turns accomplishments into activities (as argued just now in this Section);
- intransitivization is a mechanism employed to focus on the activity (refer to Page 23);
- activities are the only *Aktionsart* to bear two low-transitivity features in Hopper and Thompson's (1980) account (i.e., atelicity and durativity).

However, as Vendler (1957, p. 151) himself observes, this is not the case. He makes an interesting point relative to what he calls a "habit-forming" semantic behavior of some verbs, which also takes us back to the observations about noun incorporation and the focus on the activity in Section 2.4.2. In particular, while it is true that some activities are "habit-forming"[15] (such as *to smoke* in *Do you smoke?*), this behavior is also shown by accomplishments (e.g., a writer is someone who writes books for a living, just like a cabdriver is someone who drives a cab to earn money) and achievements (e.g., dogcatchers catch dogs for a living).

15: All examples here are from Vendler (1957, p. 151).

(Non-) "Inherently telic" verbs Moving closer to the account of telicity I am going to employ in my experimental setting (following Medina (2007)), Van Valin and LaPolla (1997, p. 112) propose a distinction between "inherently telic" verbs such as *to kill* and *to break* on one hand, and activity verbs made into accomplishments, such as *to eat*, on the other hand. This account still understands telicity as a feature of predicates, but it also acknowledges that it can be somewhat embedded in the meaning of a verb. I will return to this issue later in this Section. Crucially, as observed by Newman and Rice (2006, pp. 5–6), the activity use of *to eat* is considered the "basic" meaning of the verb in Van Valin and LaPolla (1997). The presence of a direct object in sentences featuring such verbs begets telicity depending on the nature of the object itself (Dowty 1991; Filip 2004; Tenny 1994), so that a verb is telic if its "measuring argument" is delimited (i.e., a quantized[16] object), atelic otherwise (i.e., bare plurals and mass nouns, as noted by Verkuyl (1972, 1989)). Adapting an example from Tenny (1994, p. 24), *to eat* is telic in (6-a) because the apple is a delimited measuring argument —the eating event ends when the apple is gone. On the contrary, in (6-b) *to eat* is atelic, because there is no fixed quantity of ice-cream for Chuck to consume. If he happened to live in a universe blessed with neverending, incredibly cheap ice-cream, his eating act could go on forever.

16: The standard view that maximally affected quantized objects determine telicity (Tenny 1994; Verkuyl 1972) is challenged by Smollett (2005) and Piñón (2008). For an extensive discussion of this issue, which is not within the scope of these pages, the reader is referred to Melchin (2019).

(6) a. Chuck eats an apple.
    b. Chuck eats ice-cream.

This behavior, far from being expressed by *to eat* alone, is shown by all incremental-theme verbs (already mentioned in Section 2.3.1). It is also consistent with the classic *in/for* telicity test[17] , as shown in (7).

17: I will delve into more detail about telicity tests in Section 6.2.1.

(7) a. Chuck ate an apple in an hour / *for an hour.
    b. Chuck ate ice-cream *in an hour / for an hour.

Næss (2007, 2011), Ruda (2017), and Willim (2006) note that intransitive *to eat* is compatible both with a telic reading, as in (8-a), and with an atelic reading, as in (8-b). With reference to her affected-Agent account of object drop, which I discussed extensively in Section 2.4.2, Næss (2007, pp. 78–79) argues that the telic reading is granted by Agent affectedness, as if the Agent itself worked as a measuring argument in this case. The atelic reading, instead, is argued to be typical of an event "leading to a result state, but which is in principle independent of this result state" (much like Vendler's (1957) "habit-forming" verbs), e.g., *to cook*. However, it has to be noted that the Agent is indeed affected in some measure in (8-b), making Agent affectedness a non-decisive factor.

(8) a. I ate in five minutes, then rushed off to work.
    b. We ate all evening.

Crucially, and consistently with the particular perspective adopted by Medina (2007) (and, following her, by me) following Olsen (1997 [2014]), the atelic interpretation of intransitive *to eat* seems to be more easily attained than the telic interpretation, since it does require less processing effort. In other words, both interpretations imply a focus on the activity (rather than on the Patient object, which is missing altogether), but the telic interpretation also requires that one understands the sentence as if putting additional focus on the way the Agent is affected by the activity. Quoting Olsen and Resnik (1997, p. 4), implicit indefinite objects need to appear "in the appropriate context" in order to get a telic interpretation. Næss (2007, p. 79) actually leverages the bivalent behavior of intransitive *to eat* with respect to telicity to weaken the use Hopper and Thompson (1980) make of this parameter to determine transitivity, at least in the specific case of this verb. However, since a rule specific to a single verb would not make for a strong grammar, I argue that telicity as a factor determining (or blocking) object drop is there to stay.

A note on telicity and (in)definite object drop "Inherent telicity" has interesting consequences on the theory of object drop, and in particular on the distinction between definite and indefinite object drop discussed in Section 2.2, as observed by Olsen and Resnik (1997, pp. 3–4) with reference to Allerton (1975), Mittwoch (1982), and Olsen (1997 [2014]). What they note is that implicit objects tend to receive indefinite interpretations with atelic verbs (unless they appear in particularly favorable contexts as to license a definite interpretation, as seen in (8-a)) and definite interpretations with telic verbs. An example of this argumentation is provided in (9), adapted from Olsen and Resnik (1997, p. 3). In (9-a), the inherently telic verb *to win* is shown to require a definite interpretation for the missing object, while it can be interpreted as indefinite when occurring with an inherently atelic verb such as *to eat* in (9-b).

(9) a. Benjamin won, #but I don't know what.
    b. Benjamin ate, but I don't know what.

However, such a telicity-as-definiteness account is not bulletproof. In this Paragraph I already made reference to the possibility of inducing a definite interpretation for implicit objects of inherently atelic verbs, provided sufficient context. I argue that it is also possible to induce an indefinite interpretation for implicit objects of inherently telic verbs, e.g., by presenting the action as iterative or habitual (Goldberg 2001, pp. 507–509), as in (10). Example (10-a) would be ungrammatical if the missing object was given a definite interpretation, but it can actually be considered at least partially acceptable (even though *to kill* is inherently telic) under a habitual reading. Indeed, the Joker is a notorious fictional villain from the Batman universe known for his ruthlessness, so it would be quite easy to imagine killing to be a frequent habit of his. This effect is much more evident in (10-b), where the iterative reading is made explicit by the addition of the adverb *again*, without need for extra-linguistic context.

(10) a. # The Joker killed.
b. The Joker killed again.

I will expand more on this in Section 3.3.2 and Section 6.4. For now, suffice it to say that implicit objects occurring with both inherently telic and inherently atelic verbs can be made to yield either a definite or an indefinite interpretation based on context. In particular, with relevant implications for my own model of object drop, both telic and atelic verbs can occur felicitously with implicit indefinite objects.

Olsen (1997)'s account of telicity Let us now comment on the specific account of telicity I am going to employ in my experimental setting (see also Section 6.2 for details on the implementation). Since I intend my probabilistic model of the implicit indefinite object construction as an expansion upon the original model by Medina (2007), I am going to base my interpretation of telicity on the same source she chose for her study, i.e., Olsen (1997 [2014]).
In her "semantic and pragmatic model" of aspect, Olsen (1997 [2014]) interprets telicity as a privative feature. This means that a verb can either have or not have the [+telic] feature. This feature, assigned to achievements and accomplishments, denotes in her words "the existence of an end or result to which a situation naturally will lead, not necessarily the actual attainment of such an end". The interpretation of the attainment of this end depends, in her view, not only on telicity, but also on perfectivity and tense (as I will discuss in more detail in Section 3.2.3). Crucially, the [+telic] feature is "semantic", in Olsen's words, and cannot be canceled by additional constituents. Consider, for instance, her examples in (11). She notes that "although durative adverbials are supposed to turn accomplishments into activities, (11-a) and (11-b) represent iterative accomplishments".

(11) a. Eli won for years.
b. Eli ran a mile for years.

On the contrary, she considers the [-telic] feature[18] to be "a cancelable conversational implicature", as exemplified in (12). This means that atelic verbs can receive a telic interpretation by adding a measuring argument, such as a bounded object in (12-a) or a Goal in (12-b). Indeed, "progressive forms of atelic verbs are said to entail the corresponding perfect form" (e.g., *Eli is running* entails *Eli has run*), but neither sentence in (12) obeys this requirement (e.g., *Eli is running a mile* does not entail *Eli has run a mile*). This particular state of affairs is known as the "imperfective paradox" (Dowty 1979 [2012]; White 1993), stating that progressive aspect overrides the result entailment (refer to Copley and Harley (2015), Dvořák (2017b, p. 115), and Melchin (2019) for more on this issue).

18: Actually, Olsen indicates the absence of telic denotation with [0 Telic], rather than with [-telic]. However, I will use the latter notation here for clarity, and also for consistency with my later use of this feature throughout my dissertation.

(12) a. Eli is running a mile.
b. Eli is running to the store.

This is consistent with other views of telicity discussed before in this Section, where implicit objects were argued to be more easily accepted with atelic verbs than with telic verbs, since an overt object is usually required by telic verbs (for which it works as an explicit endpoint[19] or, in other words, a "measuring argument"), while it is *admitted*, but not required, by atelic verbs.

Following Olsen, Medina (2007) uses telicity as a binary predictor of object drop. Crucially, she assigns the [+telic] or [-telic] feature to the target verbs themselves, rather than to the predicates they head. Thus, a verb in her experimental setting can either be inherently telic or inherently atelic, on the basis of rigorous tests she performed beforehand (refer to Section 6.2 for more details on the tests).

19: This traditional account (Tenny 1994) gets re-interpreted by Smollett (2005), who argues that quantized objects do not strictly delimit the event, but they just make the delimiting endpoint contextually available. Either way, Olsen's account still holds.

3.2.2 Perfectivity

Introduction Perfectivity, as a component of "grammatical aspect" or "viewpoint aspect", is a property assigned to a verb based on the perspective the speaker has on the temporal constituency of the event the verb describes (Comrie 1976). Thus, an event seen as complete[20] , having an initial and a final point, will be encoded by a verb in the perfective aspect, while an event seen as ongoing, having neither an initial nor a final point, will be encoded by a verb in the imperfective aspect. For instance, the perfective/imperfective opposition can be seen in (13), relative to English *to write*, which is perfective in (13-a) and imperfective in (13-b).

20: Not necessarily *completed*, as Comrie (1976, p. 18) observes with examples from several languages.

(13) a. John had written a thesis.
b. John was writing a thesis.

The grammatical relevance of the property of perfectivity varies crosslinguistically. In Slavic languages, such as Russian, perfectivity is embedded in the lexicon itself, so that sentences like the ones in (13) would actually require two different lexical entries to refer to the same writing event (i.e., *napisat'* for the perfective form and *pisat'* for the imperfective form). In other languages, such as English and Italian, (im)perfectivity is encoded with morphological means on a single verb lexical entry, and both perfective and imperfective aspects are compatible with any tense[21] (more on the relation between aspect and tense in Section 3.2.3).

Crucially, while telicity is a property of verbs themselves[22] (as argued in Section 3.2.1 on the basis of Olsen (1997 [2014]) and Medina (2007)) or of predicates, perfectivity is a property of events, which gets encoded on verbal heads in different ways (morphologically, in the two languages I am interested in). As discussed in Chapter 6 and shown in Chapter 7, this will have important consequences on my experimental setting.

21: This is not true of every language marking grammatical aspect morphologically. For instance, in Latin it is only possible to encode the perfective/imperfective distinction in the past tense.

22: Which gets fully realized only when interpreted within the full linguistic context.

Perfectivity and object drop Grammatical aspect has been argued to play a role in licensing or blocking indefinite object drop, even though it appears to be a path less trodden than lexical aspect. Medina (2007, p. 30) explains the lesser attention devoted to perfectivity as a predictor of object drop with reference to the fact that for many years scholars interpreted object omission as a verb-specific phenomenon (refer back to Chapter 2 for a full commentary on this). Nevertheless, literature on the matter (Cote 1996; Dvořák 2017b; Lorenzetti 2008; Næss 2007; Tsimpli and

Papadopoulou 2006) agrees that imperfective aspect (encoding duration) is more likely than perfective aspect (encoding completion) to license implicit objects. Compare, for instance, (14-a), where *to write* appears in perfective aspect and blocks object drop, and (14-b), where it appears in imperfective aspect and allows for indefinite object drop.

(14) a. *John had written.
b. John was writing.

Tsimpli and Papadopoulou (2006, p. 1609) explain this by observing that "perfectivity is understood as involving an endpoint", which is made explicit by the use of an overt object.
This different behavior expressed by perfective and imperfective clauses holds crosslinguistically. For instance, Tsimpli and Papadopoulou (2006, p. 1597) observe that while null objects are acceptable both with perfective and with imperfective verbs in Greek, they tend to be favored more by imperfective aspect. They also note (Tsimpli and Papadopoulou 2006, p. 1601) that the strict ungrammaticality of indefinite null objects occurring with perfective verbs in Russian (and Polish, as found in Sopata (2016, p. 89)) is not found in Greek. As for the languages I will base my model on, Medina (2007) provides experimental evidence in support of imperfective clauses being more prone to favor object drop than perfective clauses in English, while Cennamo (2017) comments on Italian sentences reaching the same conclusion. Moving to typologically different languages, Næss (2007, p. 118) observes that in Kalkatungu[23] variation in grammatical aspect is accompanied by changes in case-marking, i.e., ergative-absolutive in perfective clauses and absolutive-dative in imperfective clauses. This becomes relevant for the role of perfectivity in the implicit indefinite object construction when compared to previous observations about ergative languages made in Section 2.4.2, where the case was made that subjects of intransitive verbs and transitive verbs used intransitively are in the absolutive case (as they are in imperfective clauses in Kalkatungu), while subjects of transitive verbs used transitively are in the ergative case (as they are in perfective clauses in Kalkatungu). To put it more simply, since object drop is favored by imperfective aspect, it stands to reason that in such languages case gets assigned accordingly. The limitation of ergative constructions[24] to perfective environments is also noted in Hopper and Thompson (1980, p. 271), specifically with evidence from Hindi and Georgian, and references to literature about other languages.

23: Kalkatungu is a language belonging to the Pama–Nyungan family, spoken in Australia.

24: It has to be noted that this account is only valid for languages exhibiting a kind of split ergativity that is conditioned by the grammatical aspect.

3.2.3 Interactions among telicity, perfectivity, and tense

Telicity and perfectivity There are several comparisons to be drawn between telicity and perfectivity, and also between tense and these two facets of aspect (Yousefi and Mardian 2019, pp. 394–397). I will come back to this in Section 6.3.1, focusing on observations bearing direct consequences for my experimental design. Here, I will focus instead on broader concerns.
Lazard (2002, p. 162) notes "an affinity between the incompleteness of the process and the low individuation of the object", which is one of the ten parameters of prototypical transitive clauses in the account by

Hopper and Thompson (1980) (see Section 2.1, and Table 2.1 in particular). According to Lazard, an event can be construed as incomplete when the verb is in "an incompletive aspect" (progressive, habitual, imperfective...) or when "the object is only partly affected" (which is typical of atelic verbs and indefinite objects, which are interpreted as non-measuring plurals or mass nouns, as seen in Section 2.3.2). Relatedly, Næss (2007, p. 118) puts a particular focus on the association between the concept of delimitedness and the concept of affectedness. Thus, Lazard shows that atelicity and imperfectivity share the property of favoring indefinite object drop by construing the event as incomplete. Moreover, similarly to what Olsen (1997 [2014]) concludes about telicity (telic aspect being uncancelable, unlike atelic aspect), Dvořák (2017b) notes that perfective aspect is the marked form and imperfective aspect is the unmarked form. The close relation between telicity and perfectivity is also found in Hopper and Thompson (1980) themselves, who explicitly state they use these two terms "interchangeably" (Hopper and Thompson 1980, p. 270). They justify this choice, which nowadays would be untenable (Bertinetto 2001; Bertinetto and Delfitto 2000; Civardi and Bertinetto 2015), by acknowledging the poverty of the literature on the matter up to their time of writing —indeed, they note that it would be "risky to infer a distinction between the two types of aspect when none is explicitly discussed". However, consistently with previous observations about telicity and perfectivity provided in this Chapter, Hopper and Thompson recognize that telicity "can be determined generally by a simple inspection of the predicate" while "perfectivity is a property that emerges only in discourse".

Preminence of telicity over perfectivity Tsimpli and Papadopoulou (2006, p. 1598) observe that both imperfective and perfective activity verbs in Greek receive an atelic reading when combined with bare plurals (e.g., *Helen painted / was painting portraits*), hinting towards a more preminent role of (a)telicity than (im)perfectivity in determining indefinite object drop, even though the two are related in multiple ways. Indeed, if imperfectivity played a bigger role than atelicity, data would show that only imperfective verbs, regardless of their telicity feature, could occur with bare plurals (which, I insist, are the only possible interpretation for indefinite null objects together with mass nouns). The hypothesis of the preminence of telicity on perfectivity with respect to their role in the implicit indefinite object construction is also consistent with experimental evidence from English and Italian I will provide in Chapter 8. Even more strongly, Stoica (2017) finds that native speakers of Romanian are equally avoidant of indefinite null objects both in perfective and in imperfective contexts.

The preminence of telicity on perfectivity is found not only in adult grammar, but also in the early stages of grammar acquisition by children. In particular, three-year-olds and younger children have been shown to assign preferably imperfective aspect to atelic verbs and perfective aspect to telic verbs both in production and in comprehension (Medina 2007; Olsen, Weinberg, et al. 1998; Wagner 2001). However, this account of L1 acquisition, which is known as "aspect-first hypothesis" in the field (Antinucci and R. Miller 1976), needs to be taken with a pinch of salt. An alternative view exists in the literature claiming that children, instead of first acquiring aspect as entangled with Aktionsart and tense (i.e., atelic=imperfective=present *vs* telic=perfective=past), actually extract the relevant information out of the morphological structure of their L1 instead of relying on a pre-built strategy (Bertinetto, Freiberger, et al. 2015; Bertinetto, Pacmogda, and Lenci 2021).

Tense and aspect As noted by Medina (2007, p. 68), tense and aspect are independent but interrelated properties, "to the extent that one encourages certain interpretations of the other". In adult grammar (Comrie 1976) and especially in child grammar (Wagner 2001), for instance, past tense may induce a perfective interpretation of the event. I will return on the relation between past tense and perfective aspect in Section 6.3.1.
With specific regard to the implicit indefinite object construction, Dixon (1992), Glass (2020), and Goldberg (2005b) note that verbs in the past tense tend to block object drop. On the contrary, García-Velasco and Muñoz (2002, p. 9) note that present tense, interpreted as an expression of imperfectivity, favors object drop.

3.3 Pragmatic factors

In this Section, I use "pragmatic factors" as an umbrella term for several factors (neither verb-specific nor aspect-related) involved in the implicit indefinite object construction. The term covers not only purely pragmatic factors (such as a routine interpretation), but also phenomena related to intra-linguistic context (such as iterativity and habituality) and discourse factors (such as emphasis and contrastive focus). These factors, which Medina (2007) did not include in her novel model of indefinite object drop, are nevertheless crucial in a comprehensive analysis of this construction. Indeed, as DeLancey (1987, p. 54) observes in his cognition-oriented account of transitivity parameters, the interpretation of utterances in actual language use is based on real-world context or, failing such, on discourse context. This also echoes the conclusion, reached by Prytz (2016, p. 176), that the "structural sides of linguistic meaning" go hand in hand with the "contextual, pragmatic, and encyclopedic sides of meaning".

3.3.1 Purely pragmatic factors: routine

Routine is described by Glass (2020, p. 2) as "a series of recognized, conventional actions within a community", whose association with a given verb is shown by experimental evidence to vary gradiently across different communities of speakers. The author observes, with no dearth of experimental evidence from Reddit communities, that transitive verbs describing routines facilitate object drop and, likewise, object drop induces hearers to imagine scenarios where the described event is routine for its performers. She also makes an explicit connection between this account and the object-drop-as-noun-incorporation account I discussed in Section 2.4.2.
Crucially, routine cannot be construed as yet another semantic factor, because it is strongly rooted in extra-linguistic context (unlike semantic factors such as the ones I discussed in Section 3.1, which are strictly related to the very meaning of the verb itself). In particular, speakers encode a lot of world knowledge in their utterances when they use a

verb to convey a routine interpretation, and hearers (or readers) of such utterances have to be aware of the specific context they stem from in order to make sense of the intended meaning. For instance, example (15) from Glass (2020, p. 9) would only be correctly interpreted as "you are not in the habit of lifting weights to exercise" if one knew that it was uttered in the context of a conversation about fitness.

(15) You don't lift.

The role of world knowledge in facilitating object drop is also noted by Eu (2018, p. 528), where reference is made to "contextually established semantic specialization". Here, "context" is taken to refer both to "immediate context" such as the time and place of utterance, and to "general context" such as world knowledge and the life habits of the speaker. Similarly, the routine-licensed account of object drop contributes to explaining the "specialized readings" of verbs such as *to eat* (i.e., "to eat a meal") and *to drink* (i.e., "to drink alcohol")[25] noted, among others, by Næss (2011, p. 420).

25: Which I discussed extensively in Chapter 2.

3.3.2 Linguistic-context factors: iterativity and habituality

Mittwoch (2005, p. 237) observes that "the omissibility of unspecified objects is for many verbs subject to contextual factors". For instance, she notes that habitual contexts, "where the lexicon interacts with more general properties of the sentence", are more likely to license object drop than episodic contexts. Mittwoch uses the term "habitual" in a way that closely resembles the account provided by Glass (2020) (refer back to Section 3.3.1). Indeed, she understands the "imbibe alcoholic beverages" interpretation of intransitive *to drink* to be a habitual reading of the verb, since it refers to a habit the Agent is shown to have. She also applies the term, admittedly "rather freely", to sentences such as (16-a), where the habitual reading is not inferred by means of verb semantics and world knowledge (as in the case of *to drink*), but instead via extra-linguistic context —the Agent is obviously not reading the Iliad in his sleep, but he is in the habit of doing so while awake. Dispositional properties and professions appear to be another class of broadly-defined habits giving rise to felicitous object drop, as Mittwoch shows in (16-b) and (16-c), respectively. Similar considerations are also found in Fellbaum and Kegl (1989), Levin (1993, p. 39), Goldberg (2001, p. 518), and Pethõ and Kardos (2006, p. 29).

(16) a. He is reading the Iliad at the moment.
(said about somebody who is asleep)
b. Fido bites.
c. She directs (films), produces (films), conducts (music), dyes (textiles), programmes (computers).

Goldberg (2001, p. 518) argues that these "characteristic property" examples, where some typical transitive verbs can be used intransitively to elicit the interpretation that the action is somewhat characteristic of the agent, get easily explained by her principle of Omission under Low

Discourse Prominence (which I introduced on Page 23). However, she also observes that the characteristic-property interpretation is not strictly required —I would say, the characteristic property can be interpreted as either permanent or temporary, as long as it is implied in some way. She provides an example of this in (17).

(17) That dog has been known to occasionally bite, but he is generally very loving.

In addition to habitual contexts, Mittwoch (2005, p. 248) also discusses other examples of pluractionality (Lasersohn 1995 [2013]), i.e., contexts fostering event plurality, as in (18)[26] . With respect to pluractionality, Bertinetto and Lenci (2012) observe that habituality, where "the resulting habit is regarded as a characterizing property of a given referent", is closely related to iterativity (both being expressions of pluractionality) but distinct from it (since habituality, but not iterativity, belongs to the class of "gnomic imperfectives", i.e., constructions having "a characterizing function" expressed in the imperfective aspect in languages with explicit aspectual marking).

26: It should be noted that Lavidas (2013), a lone voice, argues that in English indefinite null objects of the kind illustrated by (18) "are not accepted by all native speakers". He even argues that only definite implicit objects are possible in modern English, but I would note that decades of literature on the matter, as well as my own contribution in these pages, demonstrate that indefinite object drop is indeed a possibility in modern English.

(18) a. They murdered, raped, and plundered.
b. [International tribunals] are valuable, she argues, because when they punish criminals, they also affirm, condemn, purge, and purify.

Most importantly, it should be noted that (18-a) is not a good example of habituality belonging to "gnomic imperfectives", since Simple Past is aspectually neutral in English. Let us consider the two possible translations of Mittwoch's example in Italian, provided in (19). Since both (19-a) (with verbs in the *passato remoto* tense, aspectually perfective) and (19-b) (with verbs in the *imperfetto* tense, aspectually imperfective) are grammatical, then, it is pluractionality and not grammatical aspect making (18-a) (as well as the Italian equivalents in (19)) acceptable. As a consequence, one should distinguish two independent factors:

- pluractionality, either aspectually expressed, as in (19-b), or contextually driven, as in (18-a) and (19-a);
- imperfectivity, to be specifically intended as progressivity rather than habituality, as in (20).

(19) a. Uccisero, stuprarono, saccheggiarono.
b. Uccidevano, stupravano, saccheggiavano.

(20) John was writing.

Mittwoch (2005, p. 250) further observes that habitual contexts favor object drop with verbs that would usually block it, such as the verbs of destruction used in (21).

(21) a. They usually demolish rather than restore.
b. They fell indiscriminately.

I briefly mentioned iterative contexts in (10), here reported again in (22), where the verb *to kill* was shown to admit an implicit indefinite object more easily when used iteratively, as in (22-a), than in an episodic context, as in (22-b).

(22) a. The Joker killed again.
b. # The Joker killed.

Glass (2013, p. 5) explains this difference between iterative and episodic sentences by claiming that more information is lost when object drop happens in the latter than in the former. In other words, "in these sentences describing iteration, it becomes less likely that interlocutors' communicative purposes would be thwarted" by object drop. In an interesting account of the transitivity of iterative constructions in Warrungu[27] , Tsunoda (1999, pp. 4–5) notes that in this language the iterative suffix, typically having imperfective readings (e.g., iterative and habitual), tends to combine with verbs that are very low on the transitivity scale, consistently with everything I observed about iterativity so far in this Section. The facilitating role of iterativity with respect to the implicit indefinite object construction, amply discussed by Goldberg (2001), will be further explored in Section 6.4 relatively to my experimental setting.

27: Warrungu is a language of the Pama–Nyungan family, spoken in Australia.

3.3.3 Discourse factors: emphasis, coordination and contrastive focus

According to Mittwoch (2005, pp. 251–252), "the most permissive contexts for object drop involve pairs of verbs that stand in some sort of semantic contrast", even when these verbs are "some of the poorest candidates for object drop" (e.g., *to break* and other prototypical change-of-state verbs). Contrastive focus is also discussed in these terms by Dixon (1992). Examples (23-a) and (23-b) are from Mittwoch, while examples (23-c) and (23-d) (featuring coordination instead of constrastive focus) are from Cote (1996, pp. 112, 143).

(23) a. This one creates, that one destroys.
b. A few people bought, most just looked.
c. You wash and I'll dry.
d. Bert pushed and Ernie pulled.

In these cases, where the verbs "prop each other up", object drop "is clearly a rhetorical device" used to put all the focus on the verb itself (refer to Page 23 for more on this).
In addition to constrastive focus, Goldberg (2006, pp. 196–197) mentions several other pragmatic factors licensing[28] object drop, or, as she calls it, the "Deprofiled Object construction". These additional factors are "repeated action" (which Rissman (2016) terms "x-and-x construction"), as in (24-a), and "strong affective stance", as in (24-b). In particular, Rissman (2016) interprets repeated-action contexts to target agentive meaning, i.e., to highlight "an atelic event in which the agent repeatedly performs an action". This, consistently with the affected-Agent account I discussed on Page 25, would appear to favor the implicit indefinite object construction.

28: She actually claims that "the underlying motivation for the expression of arguments is at root pragmatic", in a much more stronger account of the role of pragmatics than the one I am arguing for here. Ahringberg (2015) observes that Goldberg's claim is in direct constrast with the account provided by Fillmore (1986). A mild critique of all-pragmatic accounts of object drop is also found in Németh (2014).

(24) a. Pat gave and gave but Chris just took and took.
b. He murdered!

Goldberg (2001, pp. 506–514) also identifies additional discourse factors facilitating object drop, i.e., generic statements, as in (25-a), and infinitives, as in (25-b). Once again, object drop is shown to be favored by pluractionality. She notes (Goldberg 2001, p. 507) that in these cases "atelicity could supply the appropriate constraint", given that "repeated actions are often construed as atelic or temporally unbounded events", and atelicity (as discussed in Section 3.2.1) strongly facilitates object drop.

(25) a. Tigers only kill at night.
b. The singer always aimed to please/impress.

To all these discourse factors, Glass (2020, p. 3) also adds modal statements, as in (26).

(26) Dresses I would murder for.

Lorenzetti (2008, p. 66) uses "structural omission" as a cover term for the discourse factors discussed here. Finally, Cummins and Roberge (2005, p. 46) provide an interesting review of not-so-recent pieces of literature in French language from the late ’80s to the year 2000 about the role of pragmatics (i.e., contextual, discursive, constructional, and intention-related factors) in favoring object drop. Together with Goldberg (2001, 2005a,b, 2006), Groefsema (1995) is another strong advocate for the idea that object drop is driven, first and foremost, by pragmatic factors (*contra* Pethõ and Kardos (2006, p. 29)). Lamenting the scarcity of literature on the interaction between lexicon and pragmatics, García-Velasco and Muñoz (2002, p. 7) praise Groefsema, and also Allerton (1975), Fellbaum and Kegl (1989), and Fillmore (1986) as "notable exceptions".

3.4 A note on frequency

Frequency, intended as the number of times a verb (or a verb-object pair) occurs in a corpus, has often been observed to be somewhat correlated with other factors playing a role in felicitous object drop. For instance, Resnik (1996, pp. 149–151) hypothesizes that semantic selectivity (discussed in Section 3.1.1 and Section 6.1.1) should be positively correlated with the corpus frequency of transitive verbs used without a direct object, based on the idea that verbs selecting for highly recoverable objects should occur more easily without them than verbs selecting for scarcely recoverable objects. His results show that while high omission frequency correlates with high semantic selectivity, some verbs deviate from this trend by failing to participate in the implicit indefinite object construction despite being high-recoverability verbs. Interestingly, the opposite (i.e., low-recoverability verbs omitting their object very frequently) never happens, which Resnik takes to mean that "this pattern reflects an underlying hard requirement, namely that strong selection is a necessary condition for object omission". However, these findings are not replicated

by Ruppenhofer (2004, p. 441), who finds no association between verb frequency and their tendency to allow implicit objects in a study using a 34-verb set only partially overlapping with Resnik's 30-verb set.
Medina (2007, p. 165), using the same verb set and semantic selectivity measure as Resnik (1996), further observes that frequency fails to show a precise correlation with gradient grammaticality judgments provided by human subjects about implicit indefinite objects, since some verbs received intermediate judgments despite never being used intransitively in the Brown corpus. Medina provides two interpretations for this phenomenon, one where this is simply an artifact depending on the small size of the corpus, and another ascribing this mismatch between mid-way judgments and null corpus frequency to the existence of a threshold grammaticality value in the minds of speakers blocking them to utter sentences less grammatical than that ideal value (refer to Kempen and Harbusch (2005) for a similar account of the existence of what they call a "production threshold"). The existence of this threshold, and its actual numerical value, is an open question for future studies comparing native judgments and language production.
Goldberg (2005b) offers another relevant insight as to the role frequency plays in object drop. She suggests that frequent use of some verbs (e.g., *to smoke, to drink, to sing, to write*) in contexts favoring a habitual interpretation (discussed in Section 3.3) may give rise to "the grammaticalization of a lexical option, whereby they can appear intransitively in less constrained contexts" such as generic contexts[29] not implying habituality. Under this view, further discussed by Lorenzetti (2008, p. 65), frequency is not a direct cause of object drop, but it appears to be a rather strong facilitator in a diachronic perspective. Another example of frequency facilitating object drop, both in a synchronic and in a diachronic account, is that of verb pairs where the less frequent near-synonym does not allow implicit indefinite objects (e.g., *to devour*, compared to *to eat*). I devoted some space to the conundrum presented by such verb pairs, discussed by Glass (2020), Goldberg (2005b), and Lorenzetti (2008) among others, in Chapter 2 and in Section 3.1.3.
To conclude, indefinite object drop is never a direct consequence of frequency itself, crucially, and frequency has been shown time and again to be a poor correlate of recoverability or object-droppability. Given this, in my experimental setting (presented in Chapter 7) I will employ strategies to avoid having frequency be a confounding factor in my experiments on human acceptability judgments (refer to Section 7.2.2 and Section 8.1.2).

29: Such as the one provided by the sentence *John ate early today*, where generic *to eat* is understood to refer to the act of eating a meal.

3.5 Final considerations

In this Chapter, I presented a series of factors which the literature on indefinite object drop identifies as facilitators of the intransitive use of transitive verbs. In particular, they are:

- semantic factors: object recoverability (strongly correlated with the semantic selectivity of a verb), Agent affectedness, and manner specification (intended as a feature of "semantically marked" counterparts of neutral verbs, e.g., *to devour* with respect to *to eat*);

- actional/aspectual factors: telicity (lexical aspect), perfectivity (grammatical aspect), and aspectually driven pluractionality (i.e., habituality);
- pragmatic, contextual, and discourse factors: iterativity, contextually driven pluractionality, routine, emphasis, coordination, and constrastive focus.

I also noted that corpus frequency, both of verbs themselves and of indefinite null objects, is a rather unreliable correlate of the likelihood of a transitive verb participating in the implicit indefinite object construction. Moreover, it is never a direct cause of indefinite object drop —rather, it is a consequence of other factors at play.

Crucially, no factor among the ones I discussed here is able to predict indefinite object drop with absolute certainty. Indeed, while they all contribute to the phenomenon, none alone is responsible for it. Moreover, there is considerable difference in the effect of each factor on the omissibility of indefinite objects. For instance, while the literature consistently acknowledges that recoverability is the main driver of indefinite object drop, there are way fewer accounts relative to Agent affectedness, and this has also been noted to be a controversial factor on Page 40. Among semantic factors, manner specification (in the sense I use in this thesis) is considered to be a rather strong factor when comparing pairs of hypernym-troponym verbs where one is the manner-specified counterpart of the other. Among aspectual factors, telicity and perfectivity are reliable predictors of object drop, by and large, with telicity playing a somewhat larger role than perfectivity. Among pragmatic factors, which appear to be less powerful and more constrained than semantic and aspectual factors in facilitating object drop, some are expressed by linguistic means (e.g., iterativity, coordination, contrastive focus), while others are (also, or only) dependent on extra-linguistic context (e.g., routine, emphasis).

I picked my factors of choice based on several considerations:

- I want to expand upon the original model of indefinite object drop by Medina (2007), who used semantic selectivity, telicity, and perfectivity as predictors;
- since the same semantic and aspectual factors are at play in context-poor and context-rich utterances (as I argued in Section 2.5), I want to avoid using context-dependent factors in my experiments in order to avoid context-related confounding effects;
- based on previous observations in this Section and throughout this Chapter, Agent affectedness does not appear to be strong enough of a factor to be included in my models;
- corpus frequency is not going to be featured in the model, for reasons stated in Section 3.4 and before in this Section.

Thus, the predictors I will use in my Stochastic Optimality Theoretic models of indefinite object drop are: semantic selectivity (as a proxy to object recoverability), telicity, perfectivity, manner specification, and iterativity (because, among the context-free pragmatic factors, it is the only one requiring just the one verb in the stimulus sentence). I will discuss the experimental implementation of each in Chapter 6.

Towards a Stochastic Optimality Theoretic account of indefinite object drop

4

The model of the implicit object construction presented in this dissertation was developed under the Optimality Theory framework. In particular, it builds on the version of Stochastic Optimality Theory defined by Medina (2007) in order to model indefinite object drop in English.
In Section 4.1, I will present standard Optimality Theory and discuss both its strengths and its shortcomings. In Section 4.2, I will argue in favor of a gradient model of grammar to account for the gradient grammaticality shown by the indefinite implicit object construction. Section 4.2.1 will be devoted to Harmonic Grammar, the historical precursor of Optimality Theory, which is shown to possess attractive mathematical properties its descendant lacks, but also a number of problems making it a bad fit for modeling object drop. In Section 4.2.2 and Section 4.2.3, I will tackle the problem of moving from modeling the binary judgments traditionally used in syntax theory to modeling complex gradient judgments, which lead to the development of probabilistic grammars. A discussion of these and an in-depth introduction to Stochastic Optimality Theory will be provided in Section 4.2.4.

4.1 Standard Optimality Theory

4.1.1 Introduction

In Chapter 3, the optionality of direct objects has been shown to depend on a large number of semantic, aspectual, and pragmatic factors. The challenge a linguistic model of object drop has to undertake is to account not only for the existence of the relevant factors, but also for their combined effect on the grammaticality of the implicit object construction. My model will be based upon a probabilistic variant of Optimality Theory (Prince and Smolensky 1993 [2008], 1997; Smolensky and Prince 1993), devised by Medina (2007). A full introduction to Optimality Theory and its intricacies goes well beyond the scope of this thesis, so I will only provide a short presentation to familiarize the reader with the framework and clarify its role in my research.
In a nutshell, the grammaticality of a linguistic structure in Optimality Theory is defined in terms of its well-formedness with respect to a set of conflicting, re-rankable constraints. The constraints themselves are universal, while the order in which they are ranked is language-specific and determines the optimal (i.e., grammatical) output. Let us rephrase this in more technical, theory-specific terms. A grammar in Optimality Theory is built upon three components:

- **Gen** is the function that generates candidate outputs based on the input provided. This function is unconstrained in its generative power, so that it can create any number and type of candidates by applying any operation, such as insertion and deletion, to the input (a property called "freedom of analysis"). Typically, software (such as OTSoft by Hayes, Tesar, and Zuraw (2003) or SPOT by Bellik and Kalivoda (2019)) is used to trim the candidate set down to a feasible number of candidates to the optimization, since it would be time-prohibitive to do so by hand.

- **Con** is the set of universal, violable constraints, whose ranking hierarchy makes up the grammar of a language. They can be of two different types. **Markedness** constraints force the output to satisfy some requirements, leading it to differ from the input depending on the specific requirement. **Faithfulness** constraints, on the contrary, require that the output be identical to the input and penalize any deviation. The conflict between markedness and faithfulness constraints is at the heart of Optimality Theory.
- **Eval** is the function that picks a winner among the candidates, based on the constraint ranking. In Standard Optimality Theory, the relation between any two constraints is one of *strict domination*, meaning that the higher ranked constraint always dominates the lower ranked one regardless of how many violations each of them incurred in. As I will illustrate later in this Chapter, it is possible to define alternative versions of **Eval** in non-standard Optimality Theories where lower-ranked constraints can be more relevant than a higher ranked one based on their weights or violations.

Crucially, these three components are universal, i.e., they work the same in every language. The only language-specific aspect of an Optimality Theoretic grammar is the constraint hierarchy.
How does this work in practice? This question is best answered by looking at an example of the basic workings of Optimality Theory by Grimshaw and Samek-Lodovici (1998), regarded as a classic model in the introduction to the use of the framework in syntactic theory by Legendre (2001), and finally revised in Legendre (2019). Optimality Theory itself was developed (Smolensky and Prince 1993) within phonological theory, but in these pages I will only focus on its syntactic derivations. The version of the constraints and linguistic examples discussed in the next paragraph is the latest one by Legendre (2019).

4.1.2 Optimality Theory in practice

A linguistic phenomenon often mentioned in textbooks is that languages allowing for Topic-referring subjects to be dropped also require null expletives with weather verbs (as in (1)), while languages with compulsorily overt subjects require overt expletives with weather verbs (as in (2)).

(1) Italian
  a. (Lui/Lei) ha mangiato tre panini.
  b. (*Esso) piove.

(2) English
  a. *(She) has eaten three burgers.
  b. *(It) rains.

In order to account for the difference between Italian and English with respect to their treatment of expletives, as shown in (1-b) and (2-b)

respectively, linguists attempting to model it within Optimality Theory have to individuate the relevant constraints populating Con and to set a conflict between them, i.e., to define them so that satisfying a constraint entails violating another.
In this case, only the two constraints in (3) are at play. In particular, Subject (3-a) is the re-working of the Extended Projection Principle (Chomsky 1982) as a markedness constraint, while Full-Int (3-b) is a faithfulness constraint capturing the Principle of Full Interpretation (Chomsky 1991).

(3) a. Subject: The subject surfaces in SpecTP.
b. Full-Int(erpretation): Lexical items contribute to the interpretation of a structure.

Subject and Full-Int are in conflict in the case of zero-argument verbs (such as weather verbs) because satisfying the former requires the use of an overt expletive, which is forbidden by the latter, while satisfying Full-Int requires a null expletive, which goes against Subject. No sentence can possibly satisfy both constraints at the same time. Conflicts in Optimality Theory are solved by re-ranking the constraints based on the violations each competing candidate incurs into.
Let us discuss the tableaux for Italian and English weather verbs, in Table 4.1 and Table 4.2 respectively, where the competition between candidates and its outcome are made explicit.

Table 4.1: Optimality Theoretic tableau for weather verbs in Italian.

	***piovere*$_v$[present]**	Full-Int	Subject
	a. EXPL piove	*!	
☞	b. piove		*

Table 4.2: Optimality Theoretic tableau for weather verbs in English.

	***rain*$_v$[present]**	Subject	Full-Int
☞	a. EXPL rains		*
	b. rains	*!	

The input to syntactic optimization in Optimality Theory contains only the relevant semantic information (lexical items, argument structures, and tense specifications), and all competitors share the same semantic content. In our case, the input contains the verb, its tense, and no argument structure since weather verbs have no thematic arguments. Crucially, the expletive subject is not part of the input. Instead, it is supplied to a candidate output by Gen, the Generator function mapping the input to the set of all possible candidates to the optimization, i.e., mapping the propositional content of the input to all possible surface forms in the output. The resulting set comprises two candidates, one with an expletive subject (in a.) and one without (in b.). As explained above, candidates with an expletive subject violate Full-Int and subject-less candidates violate Subject. In the tableaux, an asterisk marks a single violation of a constraint, an exclamation mark after an asterisk marks a *fatal* violation (i.e., a violation excluding the candidate from further evaluation), the pointing finger indicates the optimal candidate, and constraints are ordered so that each constraint dominates the one to its right. Given the typology sketched in (1-b) and (2-b), an Optimality Theoretic analysis of Italian and English leads to the conclusion that the difference between the two languages with respect to zero-argument verbs results from the different ranking of two universal constraints, Full-Int and Subject (as shown in the tableaux Table 4.1 and Table 4.2). An explanation of the behavior of weather verbs in Italian and English, as well as a broader account of the typology in (1) and (2), is not something that can happen only under Optimality Theory. For instance, Principles-and-Parameters theory (Chomsky 1981) accounts for the different behavior of Italian and English with the Pro-Drop parameter, so that Italian is a [+pro-drop] language and English is a [-pro-drop] language. In such a framework, inviolable, fixed parameters determine the grammaticality of a linguistic structure in a given language. Optimality Theory on the contrary, as "a formal theory of constraint interaction in Universal Grammar" (Legendre 2001) relies on violable, re-rankable constraints to determine grammaticality. Most importantly, grammaticality is assigned to a linguistic structure as the outcome of a competition among several candidates, instead of being a property of that linguistic structure taken on its own. This is a crucial aspect of Optimality Theory that makes it a suitable framework for my analysis of the implicit object construction, even though *standard* Optimality Theory suffers from problems that are best solved by other variants of the same framework.

4.1.3 Impossible violation profiles

To explain the nature of these problems, I will start with an example. Let us say that we collected a small set of data from a language, that we wanted to model the distribution of these data using three theory-driven constraints, and that these data presented the constraint violation profiles represented in Table 4.3, Table 4.4, and Table 4.5.

	(a) candidate set:	Constr. 3	Constr. 1	Constr. 2
☞	candidate A		*	*
	candidate B	***		

Table 4.3: Hypothetical Optimality Theoretic tableau for mock data by Kuhn (2002)

	(b) candidate set:	Constr. 1	Constr. 2	Constr. 3
☞	candidate A′			*
	candidate B′		*	

Table 4.4: Hypothetical Optimality Theoretic tableau for mock data by Kuhn (2002)

	(c) candidate set:	Constr. 1	Constr. 2	Constr. 3
☞	candidate A″			*
	candidate B″	*		

Table 4.5: Hypothetical Optimality Theoretic tableau for mock data by Kuhn (2002)

It is evident that this model of our mock data is unfeasible, considered that in any given language the constraints have to be ordered consistently (under the assumption of strict dominance). However, in the model above this does not happen, since Table 4.3 is a ranking argument for Constr. 3 $\gg$ Constr. 1 and Constr. 3 $\gg$ Constr. 2, while Table 4.4 and Table 4.5 are ranking arguments for the opposite.

It is possible that these constraints have to be ditched in favor of a whole new set of constraints, but let us say that they are very well motivated by the literature on the topic and that there is no reason in linguistic theory to doubt their effect on the grammaticality of these data. Another possible solution to the inconsistency would be to introduce an additional constraint, Constr. 4, which would be ranked higher than all the others and violated only by candidate B. Provided we could indeed find such a constraint, this would lead to a situation where all the candidate sets from this hypothetical language are consistent with the ordinal constraint ranking Constr. 4 ≫ Constr. 1 ≫ Constr. 2 ≫ Constr. 3. Yet, Constr. 4 may be weakly grounded in the linguistic theory and thus a poor choice for a constraint in an Optimality Theoretic analysis. Remember that constraints are assumed to be part of Universal Grammar, and they have to be ranked consistently within a given language. What if Constr. 4 is found to be incompatible with newer data from the same hypothetical language considered above? The quest to find an additional suitable constraint could continue, but the same problem would surface again and soon lead to an unmotivated number of constraints. I will provide a possible solution to this problem (and discuss its shortcomings) in Section 4.2.

4.1.4 Why standard OT is a bad fit for a model of object drop

I will now focus on a feature of standard Optimality Theory that, while not being a worrying issue *per se*, may become a problem for those using it to model a phenomenon as complex as the implicit indefinite object construction. Let us see why.

Linguists adopting standard Optimality Theory to perform their analyses of syntax (and other aspects of grammar) rely on binary grammaticality judgments. As illustrated above, given a set of candidates competing for optimality, the optimization yields only one optimal candidate, which is the only candidate deemed grammatical, while all the others are equally ungrammatical regardless of the number or ranking of constraint violations they incurred into. Now, considering the phenomenon examined in this dissertation, a traditional Optimality Theoretic approach to modeling the grammaticality of the implicit object construction (such as the one proposed by Yankes (2021 [2022])) would only account for an oversimplified typology such as the one in (4).

(4) a. John sang.
 b. *John built.

Such an account would perfectly describe a world where the linguistic factors presented in Chapter 3 make some object-less sentences with transitive verbs fully grammatical (as in (4-a)) and other ones fully ungrammatical (as in (4-b)). However, the actual scenario is more complex. For instance, a classic Optimality Theoretic analysis of object drop would not capture the different grammaticality of the sentences in (5), even though the literature suggests the imperfective sentence should be at least slightly more acceptable than the sentence in the perfective aspect. Similarly, it would fail to model cases such as (6), where neither candidate is fully grammatical, making it weird, if not altogether wrong, to declare an *optimal* candidate.

(5) a. John had eaten.
 b. John was eating.

(6) a. *John built.
b. ?John was building again.

Example (7) is the one used by Medina (2007, p. 62) to make the point that, crucially, the grammaticality of indefinite object drop varies not only within a group of candidates featuring the same verb, but also across verbs.

(7) a. Jack ate.
b. *Jack found.
c. ?Jack caught.

Standard Optimality Theory fails to capture all of these observations in a viable model of indefinite object drop. I will discuss possible solutions to this problem in Section 4.2.2 and focus on the one I will use for my experiments in Section 4.2.4 and Chapter 5. In particular, I will make it clear that a model is necessary which accounts for nuanced, gradient effects —a goal that is unattainable within the realm of standard Optimality Theory.

Indeed, the standard Optimality Theoretic model of indefinite object drop in English conceived by Yankes (2021 [2022]) suffers from these exact problems. The author acknowledges the concerns relative to the glaringly obvious *gradient* grammaticality of the implicit indefinite object construction, but dismisses them by simply stating that whenever his model chooses the null-object candidate as a winner over the overt-object candidate, this has to be interpreted as a possibility available in the grammar, rather than a strict imposition. On the contrary, he argues that the grammar would never allow for object drop in all the cases where the winner is the overt-object candidate. This explanation serves well a model entrenched in standard Optimality Theory, but, to a closer analysis, it appears to be an attempt to have it both ways. Given the existence of well-known strategies to model linguistic phenomena showing unmistakable gradient grammaticality (both within and without the realm of Optimality Theory, as discussed in Section 4.2), it makes little sense to resort to a theory of grammar where only one winner candidate is possible, only to admit unquantified gradience in some cases. The author claims to be "seeking to apply the concepts of Optimality Theory to a new space", perhaps laying "the groundwork for a more rigorous overhaul in the future", but in doing so he overlooks the gradient, probabilistic model of indefinite null objects in English devised by Medina (2007) within the framework of (stochastic) Optimality Theory. The choice to dismiss probabilistic approaches to grammar in favor of standard Optimality Theory is all the more puzzling if one considers that Yankes (2021 [2022], p. 80) ranks the constraints based on the effect size of the corresponding linguistic features found in the preliminary statistical analysis he performed on the results of his behavioral experiment, then manually fiddling with the constraint ranking to account for several mistaken predictions of the model (Yankes 2021 [2022], pp. 84–90).

I will come back to Medina's solution to the problems of probabilistic constraint re-ranking and the gradient grammaticality of indefinite object drop in Chapter 5.

4.2 Weighted approaches to constraint ranking

4.2.1 Harmonic Grammar

Let us forgo standard Optimality Theory and try replacing ranked ordinal constraints with numerically weighted ones, to avoid resorting to additional constraints and solve the problem of having impossible violation profiles. This was actually the strategy employed by the historical precursor of Optimality Theory, Harmonic Grammar, sketched out by Legendre, Miyata, and Smolensky (1990, 1991) and Smolensky, Legendre, and Miyata (1993), and later discussed in the light of Optimality Theory by Legendre, Sorace, and Smolensky (2006) and Pater (2009). Harmonic Grammar was created as an attempt to bridge the apparently insurmountable gap between generative grammar, which is a formal model of existing languages, and linguistic models based on connectionism, which are instead models of cognition translating the patterns of neural activity into mathematical functions. Most importantly, there is no notion of constraint *ranking* in Harmonic Grammar, so that their ordered depiction in the tableaux has no bearing on the model itself. In practice, Harmonic Grammar models assign a non-negative numerical weight to each constraint, each candidate gets a negative score for each violation of a constraint or a positive score for each satisfaction (unlike in Optimality Theory), and the harmony of each candidate is computed as in Equation 4.1.

$$H = \sum_{k=1}^{K} s_\mathrm{k} w_\mathrm{k} \tag{4.1}$$

Given a set of k constraints, $\mathrm{s_k}$ is the violation score of a candidate and $\mathrm{w_k}$ is the weight of a constraint. The only candidate deemed grammatical, i.e., the winner of the competition for optimality, is the most harmonic in the candidate set. Unlike in Optimality Theory, where strict domination determines that the violation of a higher-ranked constraint is always worse than the violation of a lower-ranked constraint, the weighting of constraints in Harmonic Grammar makes it possible to model the cumulative effect of multiple violations of a constraint (see Table 4.6).
Let us look at some made-up examples in Table 4.6, adapted from Example 7 in Kuhn (2002) to conform to Harmonic Grammar symbolism. Unlike tableaux in Optimality Theory, Harmonic Grammar tableaux have an additional row for the constraint weights and an additional column for the harmony scores of candidates. The starting weights assigned to the three constraints in the examples below do not determine a reliable model, since they are compatible with the candidate sets (a) and (b), but not with (c), where candidate B results more harmonic than candidate A. Based on Equation 4.1, the harmony of each candidate is computed by multiplying each violation score by the weight of the corresponding violated constraint (e.g., -1*4 = -4 and -1*1 = -1 in (a)) and then, by summing up these partial results (in (a), -4 + (-1) = -5). Within a tableau, the grammatical candidate is the most harmonic, i.e., the one with the highest harmony value (in (a), this is candidate A, since -5 is greater than -9). Thus, the starting weights are incompatible with (c), where the

grammatical candidate has a lower harmony score than the other one (since -3 is less than -1).

(a) candidate set:	Constr. 1 4	Constr. 2 3	Constr. 3 1	*H*
☞ candidate A	-1		-1	-5
candidate B		-3		-9

(b) candidate set:	Constr. 1 4	Constr. 2 3	Constr. 3 1	*H*
☞ candidate A′		-1		-3
candidate B′	-1			-4

(c) candidate set:	Constr. 1 4	Constr. 2 3	Constr. 3 1	*H*
??? ☞ candidate A″		-1		-3
candidate B″			-1	-1

Table 4.6: Hypothetical Harmonic Grammar tableaux for mock data, adapted from Kuhn (2002)

The optimization process in this Harmonic Grammar model of our mock data has to continue, by means of updating the weight of each constraint via backpropagation[1] until the model converges on a consistent representation of linguistic data. Interested readers who want to try their hand at this will find that the intended configuration would be $w_1 = 6$, $w_2 = 4$, $w_3 = 5$. At first glance, it would indeed seem that a simple computation saved us the trouble of coming up with unmotivated constraints, and yielded the intended model of our linguistic data. However, as Kuhn (2002) puts it, a weighted-constraint theory has an "undesiderable property" for the linguist trying to model typologically realistic languages with linguistically motivated constraints. Indeed, in Harmonic Grammar, the linguistic motivation of a set of constraints risks being trivial, since it is always possible to bend the weight set until the model accomodates all data, regardless of the constraints used in the model. A severe consequence of this state of affairs is that Harmonic Grammar (or any other version of an Optimality Theoretic-like grammar with weighted constraints) overgenerates data, namely, it predicts both possible and impossible languages (Legendre, Sorace, and Smolensky 2006; Pater 2009).

1: Backpropagation is a resource-intensive training algorithm. Linear Optimality Theory, a later update of Harmonic Grammar by Keller (2000, 2006), determines constraint weights using instead standard Least Square Estimation (and only models violations, not satisfactions, of constraints).

Harmonic Grammar, at least in its original formulation, is a very powerful connectionist tool, but a poor model of the actual typology of human languages. Smolensky and Prince (1993, p. 216) thus conclude that "recourse to the full-blown power of numerical optimization is not required", a concept that has become famous among linguists in its witty form "grammars don't count", and further explored in Smolensky (2006). Optimality Theory was then developed as a more restricted, linguistically motivated spawn of Harmonic Grammar, replacing weighted constraints with constraints ranked according to strict dominance. Optimality Theory is the last step in the path towards symbolic functions and representations that started with fully numerical neural networks and continued with functionally numerical but representationally symbolic Harmonic Grammar. It is important to observe that the weighting-to-ranking shift still has a clear mathematical meaning, which becomes apparent if one were to recast an Optimality Theoretic analysis in an Harmonic Grammar fashion. From this perspective, strict dominance can be seen as the result of exponential weighting of the constraints (so that no violation profile can exist where a lower-ranked constraint outranks a higher-ranked constraint), making Optimality Theory "a very specialized kind of Harmonic Grammar" (Prince and Smolensky 1993 [2008]; Smolensky and Prince 1993). Thus, Optimality Theory is the actual intended link between generative grammar and neural computation, being a formal theory of constraint interaction (i.e., Universal Grammar) grounded in empirical observations about language and connectionist math. However, neither Harmonic Grammar nor standard Optimality Theory are the most suitable framework to model gradient judgments about the implicit object construction.

4.2.2 Going gradient

As argued in Section 4.1.4, a phenomenon as complex as indefinite object drop can't possibly be described by means of binary grammaticality judgments. *Gradient* judgments[2] have to be collected instead, and a suitable linguistic model has to be created to account for *gradient* grammaticality. A first attempt to bridge the gap between the resources of standard Optimality Theory and the need for finer-grained acceptability judgments was made by Keller (1997). In this work, the assumption of standard Optimality Theory that all non-optimal candidates are equally ungrammatical is dropped, in favor of an extended version of the framework where the grammaticality of each candidate is formulated in terms of its relative rank with respect to its competitors. While in standard Optimality Theory grammaticality equates to *global* optimality over the whole candidate set, in Extended Optimality Theory it equates to *local* optimality with respect to a subset of the candidate set (aptly named "suboptimality"). Extended Optimality Theory thus establishes a grammaticality hierarchy among the candidates, and this predicted hierarchy can be evaluated against empirical data, i.e., graded acceptability judgments.

2: For observations on the importance of gradient acceptability judgments for linguistic theory beyond Optimality Theory, as well as some words of caution pertaining their validity, please refer to Bornkessel-Schlesewsky and Schlesewsky (2007), Juzek and Häussler (2019), Lau, A. Clark, and Lappin (2017), Rimmer (2006), Schütze (1996 [2016]), and Sprouse (2007).

Let us look at an example from Keller's experiment on the effect of definiteness and verb class on extraction from verbs (Keller 1997, pp. 10–12). The acceptability judgments he obtained[3] for the sentences in (8) show that extraction from *take*-type verbs is more acceptable than extraction from *destroy*-type verbs, that the extraction of indefinite arguments is more acceptable than the extraction of definite arguments, and that verb class has a stronger effect than definiteness on grammaticality.

3: Mean acceptability ratings obtained via a magnitude estimation experiment (Bard, Robertson, and Sorace 1996) involving nineteen native speakers of English.

(8)	a.	Which man did you take a picture of?	49.39
	b.	Which man did you take the picture of?	43.74
	c.	Which man did you destroy a picture of?	41.01
	d.	Which man did you destroy the picture of?	36.94

The Extended Optimality Theoretic tableaux representing the violation profiles of these candidates is in Table 4.7. Keller (1997) bases his choice of constraints on Diesing (1992), Legendre, C. Wilson, et al. (1995), and Legendre, Smolensky, and C. Wilson (1998), which I won't discuss here in full length due to space constraints. The BAR^n constraints are an implementation of the MINIMALLINK family of constraints, which requires chain links[4] to cross a minimal amount of barriers (see Chomsky (1993) for a full explanation of the Minimal Link Condition, according to which syntactic movement has to target the closest potential landing site). BAR has a counter for each time a barrier gets crossed. SUBCAT requires instead that subcategorization requirements be met. The syntactic representation of each candidate in Table 4.7 follows directly from the theory of extraction by Legendre, C. Wilson, et al. (1995) and Legendre, Smolensky, and C. Wilson (1998), and from the theory of definiteness by Diesing (1992), which requires that [-creation] verbs such as 'to destroy' subcategorize for a definite NP. M is a mapping operator correlating with the definiteness feature, and it turns the projection it adjoins to into a barrier for movement (more on this in Keller (1997, p. 11) and literature referenced hereby).

Let us comment the violation profiles in the tableau. All candidates violate BAR^1 once because the chain $\langle M_{\mathrm{j}}, NP_{\mathrm{j}} \rangle$ crosses the barrier VP. The chain $\langle$which man$_{\mathrm{i}}, t_{\mathrm{i}}\rangle$ crosses only VP in candidate (a), resulting in another violation of BAR^1, while it crosses both VP and IP (made into a barrier by M) in candidate (b), resulting in a violation of BAR^2. Keller (1997) comments that the additional violation of BAR$^{\mathrm{N}}$ candidates (c) and (d) incur into is due to the fact that barrierhood correlates with feature selection, so that a definiteness feature requirement on the verb turns it into an additional barrier for the movement. Candidate (c) also violates SUBCAT because the [-creation] verb subcategorizes for an indefinite NP.

4: When a constituent moves in a syntax tree, chains are the combination of the moved copy and the trace it leaves behind.

	Q$_i$ [NP$_{\mathrm{Subj}}$ V [NP$_{\mathrm{Pict}}$ x_i[+wh]]	BAR 3	SUBCAT	BAR 2	1
☞a.	[$_{\mathrm{CP}}$ which man$_i$ did [$_{\mathrm{IP}}$ you [$_{\mathrm{VP}}$ M$_j$ [$_{\mathrm{VP}}$ take [$_{\mathrm{NP}\ j\ [\text{-def}]}$ a picture of t$_i$[+wh]]]]]]				**
b.	[$_{\mathrm{CP}}$ which man$_i$ did [$_{\mathrm{IP}}$ M$_j$ [$_{\mathrm{VP}}$ you [$_{\mathrm{VP}}$ take [$_{\mathrm{NP}\ j\ [\text{+def}]}$ the picture of t$_i$[+wh]]]]]]			*	*
c.	[$_{\mathrm{CP}}$ which man$_i$ did [$_{\mathrm{IP}}$ you [$_{\mathrm{VP}}$ M$_j$ [$_{\mathrm{VP}}$ destroy$_{[\text{+def}]}$ [$_{\mathrm{NP}\ j\ [\text{-def}]}$ a picture of t$_i$[+wh]]]]]]		*	*	*
d.	[$_{\mathrm{CP}}$ which man$_i$ did [$_{\mathrm{IP}}$ M$_j$ [$_{\mathrm{VP}}$ you [$_{\mathrm{VP}}$ destroy$_{[\text{+def}]}$ [$_{\mathrm{NP}\ j\ [\text{+def}]}$ the picture of t$_i$[+wh]]]]]]	*			*

Table 4.7: Extended Optimality Theory tableau relative to the examples in (8), from Keller (1997, p. 12).

The constraint ranking in Table 4.7 is compatible with the grammaticality judgments in (8). Supported by empirical data, this Optimality Theoretic analysis does more than yield a single optimal candidate. As a matter of fact, it predicts a grammaticality hierarchy by exploiting evidence from the optimal candidate and the suboptimal ones alike. Candidate (a) is the optimal candidate, the one which would stand out as the only winner in a standard Optimality Theoretic analysis. Candidate (b) is suboptimal, in that it violates a higher-ranked constraint than (a), but it still performs better than (c). The same goes for candidates (c) and (d), each showing decreasing degrees of optimality based on their violation profiles.

An important consequence of predicting *sub*optimality in place of plain optimality is that such an analysis makes a stronger case for a given constraint ranking, given that re-ranking a constraint may not affect the optimal candidate, but it may have visible effects on suboptimal

candidates. For instance, re-ranking SubCat over Bar[3] in Table 4.7 would not affect the optimality of candidate (a.), but it would wrongly make candidate (d.) more suboptimal than (c.). Thus, Extended Optimality Theory leverages the same features of the standard framework, but is also able to detect otherwise hidden re-rankings and to include finer-grained grammaticality judgments in the linguistic analysis.
Nevertheless, while this extension of the original framework makes it possible for syntacticians to model non-binary grammaticality judgments, it does not go beyond the assumption of strict domination among constraints. Moreover, even though graded grammaticality judgments are collected from native speakers and averaged to evaluate the (sub)optimality of candidates, these numerical values have only ordinal meaning. In other words, they are used to rank the candidates according to their grammaticality, but they have no other use. As a result, in Extended Optimality Theory the gradience of grammaticality actually stems from discrete, not continuous, values.

4.2.3 Probabilistic grammar

Looking at the bigger picture of what rippled through linguistics at the turn of the century, awareness of the aforementioned issues with Optimality Theory (and its precursor, and its early extension) went hand in hand with a growing discontent with the many limits of traditional linguistic theories based on categorical definitions and binary judgments. The road that took linguists from a categorical view of language to modeling gradient grammaticality was a long, bumpy one. Ever since Chomsky (1957) frowned upon the notion of "probability of a sentence", categoric linguistics was assumed to be linguistics *tout court*, and thwarted was any attempt to bring linguistics closer to information theory and other computation-savvy fields of study. As discussed in Section 4.2.1 and Section 4.2.2, linguistics had to wait several decades for Harmonic Grammar and then Optimality Theory to acknowledge that computation can be beneficial to linguistic theory. Most notably, human cognition entails probability and computation, and modern linguistics is first and foremost a cognitive science.
In the lively debate on probability in linguistics, Bod, Hay, and Jannedy (2003) was a pivotal piece of literature, and its chapter on syntax (Manning 2003) created a compelling narrative of the need for probabilistic grammar, heralding a new season of theoretical (M. Crocker and Keller 2006; Sprouse 2018; Wasow 2007), experimental (Alexopoulou and Keller 2006; Brehm and Goldrick 2017; Bresnan, Featherston, and Sternefeld 2007; Bresnan and Hay 2008; Bresnan and Nikitina 2008; Keller and Sorace 2003; Sorace and Keller 2005; Sprouse 2015), and computational (Turney and Pantel 2010) studies. The main take-home message from Manning (2003) is that categorical accounts of grammar flatten the gradience of grammar by reducing it to an all-or-nothing pattern and underplay complexity as anecdotal observations or "exceptions", whereas probabilistic frameworks can indeed account for the full range of empirically observed gradience[5] while still being consistent with traditional categorical models. Considering the meaning of probability, this is not surprising at all. In fact, the probability of an event is quantified on a continuous numerical scale ranging from 0 (i.e., impossible) to 1 (i.e., certain). Categorical accounts only work with the extreme values of the

5: Sprouse (2007) discusses experimental evidence where human participants to magnitude-estimation experiments actually enact a "spontaneous imposition of a categorical distinction on a continuous rating scale", apparently supporting a categorical approach to grammaticality. However, this result is not found in dozens of other experiments in the same vein, meaning that while it is possible for a gradient scale to be flattened onto a binary opposition, it is never the case that a categoricity-oriented study captures more information than a gradience-oriented one. Indeed, researchers should heed the advice by Bornkessel-Schlesewsky and Schlesewsky (2007) to be wary of overinterpreting "enticing" gradient data, but the point still stands that strictly binary accounts, especially if based on the introspection of a single linguist, just replicate the findings of graded-scale experiments for uncontroversial phenomena (Bader and Häussler 2010; Linzen and Oseki 2018), while limiting the understanding of phenomena involving subtle differences in grammaticality (Keller 1998b).

probability scale, so that under such a model, a linguistic event can only be either impossible (or ungrammatical) or certain (or grammatical). In this case, grammar is conceived as a set of rules to sift out impossible or ungrammatical utterances, and allow grammatical ones. Under a probabilistic model of grammar, on the other hand, the probability of a linguistic event (such as an utterance being produced by a speaker, or a sentence being judged grammatical) can assume any value on the scale, and grammaticality can indeed be modeled as a gradient phenomenon. In this case, there are no rules in the traditional sense, but only soft probabilistic constraints. Moving from categorical to probabilistic models of grammar has been equivalent to a paradigm shift in linguistics, and it has allowed linguists to create far better explanations of what was hastily dubbed "an exception" in the past.

A famous example of a phenomenon whose modeling vastly benefited from this change of perspective is the distinction between arguments and adjuncts, which has been considered binary ever since Tesnière (1959 [2015]), shifted to a 3-way distinction later on (Aldezabal et al. 2003; Dowty 2003; Van Valin and LaPolla 1997; Villavicencio 2002), until linguists, picking up Vater's (1978) intuition, moved to a graded-argumenthood approach (Cennamo and Lenci 2019; Kim, Rawlins, and Smolensky 2018, 2019; Kim, Rawlins, Van Durme, et al. 2019). This thesis aims to be a useful extension to the debate on transitivity, another linguistic phenomenon that has been considered binary (or binary-with-exceptions) for a long time. The utterly positive trade-off between renouncing categorical grammar accounts and adopting probabilistic models has been demonstrated time and again by the experimental literature on the matter, and a mathematical proof of the consistency of Stochastic Optimality Theory with the standard framework has been recently provided by Magri (2018)[6] .

Advocating for probabilistic approaches to syntax, in particular with respect to the shortcomings of Optimality Theory I discussed in Section 4.1.4, several linguists developed their own solutions in the late '90s and early '00s (Alexopoulou and Keller 2006; Boersma 2004; Boersma and Hayes 2001; Davidson and Goldrick 2003; Keller 1998a, 2000, 2006; Sorace and Keller 2005). The common feature underlying all these reformulations of Optimality Theory is that, by dropping the assumption of strict domination made in standard and Extended Optimality Theory, the gradient grammaticality of a candidate is defined as a function of the number and type of constraint re-rankings returning it as optimal. These models all allow for the re-ranking of constraints under an Optimality Theoretic lens, based on different algorithms and mathematical functions. How does this work in practice? For simplicity's sake, in Section 4.2.4 I will focus only on the traditional version of Stochastic Optimality Theory (Boersma et al. 1997; Boersma and Hayes 2001), which is not only "the best motivated and most thoroughly probabilistic extension to Optimality Theory" (Manning 2003, p. 25), but also the one directly inspiring the model of indefinite object drop by Medina (2007), which I will present in Chapter 5.

6: The paper deals with phonology in particular, but the same conclusions can be easily drawn for syntax.

4.2.4 The floating-constraint approach and Stochastic Optimality Theory

In the previous Sections, I demonstrated with examples from the literature the shortcomings of standard Optimality Theory (discussed in

Section 4.1) and its precursor, Harmonic Grammar (discussed in Section 4.2.1). Wrapping up, these theories of grammar are unable to deal with candidate sets where each candidate is assigned a gradient grammaticality score on a continuous scale, yielding instead a single optimal candidate and several equally ungrammatical ones. Even the extended version of Optimality Theory by Keller (1997) (presented in Section 4.2.2), although admitting degrees of suboptimality, still assumes strict domination among constraints, and makes use of grammaticality scores just to give the candidates an ordinal rank.

Let us discuss Stochastic Optimality Theory (Boersma et al. 1997; Boersma and Hayes 2001), which deals with the aforementioned problems in a way that is particularly suitable for modeling phenomena such as the indefinite object construction, i.e., employing floating constraints instead of a fixed ranking as in Standard Optimality Theory. The application of Stochastic Optimality Theory to this particular topic of interest by Medina (2007), upon which I will build my own model of object drop in English and Italian (defined in Chapter 9), will be discussed separately in Chapter 5.

First of all, let us visualize the fixed ranking of three hypothetical constraints in an Optimality Theoretic model, as shown in Figure 4.1.

CONSTRAINT 1 $\gg$ CONSTRAINT 2 $\gg$ CONSTRAINT 3

Figure 4.1: Fixed constraint ranking.

Stochastic Optimality Theory takes the constraint ranking idea to the next level by putting them on a continuous, numerical scale, as shown in Figure 4.2. This first step makes it clear that, in the example I am discussing, CONSTRAINT 1 and CONSTRAINT 2 are closer together than CONSTRAINT 3 is to either of them, so it is possible that the relative ranking of CONSTRAINT 1 and CONSTRAINT 2 is not as strict as the (lower) ranking of CONSTRAINT 3.

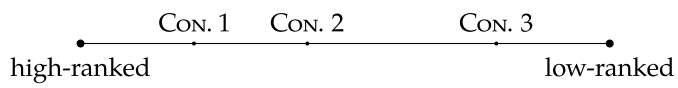

Figure 4.2: Fixed constraint ranking, but on a *continuous* scale.

Now, the model has to implement a way to deal with gradient grammaticality, solving the many issues standard Optimality Theory has with modeling the implicit object construction (debated in Section 4.1.4). Stochastic Optimality Theory achieves this result by means of the Gradual Learning Algorithm, which is intended to be an improvement on the Constraint Demotion algorithm (Tesar and Smolensky 1993) used in standard Optimality Theory. Building a Stochastic Optimality Theoretic grammar is a sensible choice for linguists dealing with any phenomenon where a given input generates candidates with no unique winner, such as language change (with different winners at different times in history), language development (with different winners at different times in the child's life), and complex synchronic phenomena such as indefinite object drop (with different winners at the same time, based on the interaction of the semantic, aspectual, and pragmatic factors listed in Chapter 3).

Stochastic Optimality Theory allows more than one candidate to be optimal at the same time by allowing constraints to "float" on the continuous scale, as if perturbed by numerical noise at the moment of evaluation. In practice, this is possible by assigning to each constraint a full range of values, centered on what was previously a single point on the scale (now called "ranking value"). Given the range of values, the specific value used at evaluation time is called an "evaluation point". Most importantly, the constraint ranking ranges are defined as probability distributions, and distribution overlap determines the probability of two constraints re-ranking with respect to one another. As illustrated in Figure 4.3 and Figure 4.4, probability distributions in Stochastic Optimality Theory are *normal* distributions. The Gaussian curves are defined by their mean value, which is the ranking value, and their standard deviation, which determines how broad the curve is.[7] Since all constraints are assigned the same normal distribution in traditional Stochastic Optimality Theory, the actual value of the standard deviation (the "evaluation noise") has no effect on the constraint re-ranking, and it is arbitrarily set at 2 (a different approach to this matter is provided in Nagy and B. Reynolds (1997) and W. T. Reynolds (1994)). At evaluation time, the selection point will occur most probably in correspondence of the ranking value (given the properties of normal distributions), and its probability steady declines as its value departs from the center of the distribution (i.e., the ranking value).

7: Let μ be the mean value of the distribution and σ the standard deviation. The normal distribution is a function defined by the equation:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

Going back to the simple state of affairs illustrated in Figure 4.2, it is evident from Figure 4.3 that the floating-constraint model makes it now possible for Constraint 1 and Constraint 2 to re-rank. In the picture, Constraint 1 has a ranking value of 6.5 and Constraint 2 a ranking value of 4, ordered from the higher-ranked to the lower-ranked as in the custom of Optimality Theory. In particular, as shown by the overlapping curves, the two constraints re-rank freely when the selection points are comprised between 4 and 6.5, and it is much more probable for Constraint 1 to outrank Constraint 2 than the opposite (graphically, there is only a small area where the curve for Constraint 2 is above the curve for Constraint 1, for x values between 4 and 5.25).

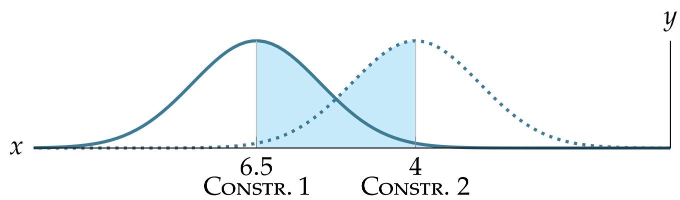

Figure 4.3: Constraint re-ranking distribution in Stochastic Optimality Theory (overlapping constraints).

Instead, Constraint 2 and Constraint 3 never re-rank with respect to one another, since their probability distributions in Figure 4.4 never overlap (ranking values have been randomly assigned in the picture just for the argument's sake). In situations such as this, the constraint ranking on a continuous scale just reproduces the results of the categorical ranking yielded in standard Optimality Theory, i.e., strict domination.

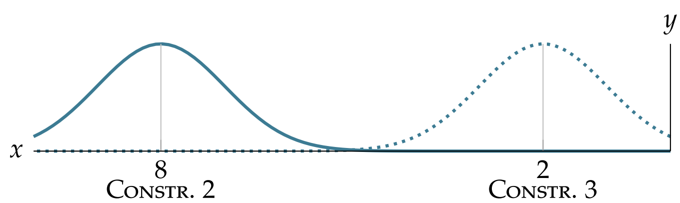

Figure 4.4: Constraint re-ranking distribution in Stochastic Optimality Theory (non-overlapping constraints).

The Gradual Learning Algorithm uses these premises to assign an empirically motivated ranking value to each constraint, modulating its outcome based on an error-driven procedure. The full details of how the algorithm works, which the reader will find in Boersma and Hayes

(2001), are outside the scope of this chapter. For the purposes of my dissertation, it is important to note that Stochastic Optimality Theory provides an optimal environment to create a model of indefinite object drop. Let us consider a simple example such as (9), where the verb *to eat* is used transitively in (9-a) and intransitively in (9-b), and let us ignore the effect of the many factors influencing object drop (discussed in Chapter 3) for the time being. Both (9-a) and (9-b) are grammatical, but the latter is judged slightly less acceptable than the former on average by some hypothetical native speakers of English.

(9) a. John is eating pizza. 7
b. John is eating. 6.5

In our simplified model of object drop, we would posit just two conflicting constraints in the spirit of Optimality Theory (as explained in Section 4.1): *Internal Argument Structure, a markedness constraint penalizing the presence of an overt direct object in the output, and Faithfulness to Argument Structure, a faithfulness constraint requiring all the arguments in the input to be also realized in the output. In a standard Optimality Theoretic analysis of the candidate set in (9), Faithfulness to Argument Structure would be ranked above *Internal Argument Structure and make (9-a) the only winner in the competition, with no reference to the slight acceptability difference. A Stochastic Optimality Theoretic model would solve the problem, so that Faithfulness to Argument Structure would indeed be ranked above *Internal Argument Structure most of the times, but with a large overlap between the two probability distributions (see Figure 4.5). Crucially, one constraint outranks the other only probabilistically, while in the standard Optimality Theoretic model it would do so in an absolute sense due to strict domination.

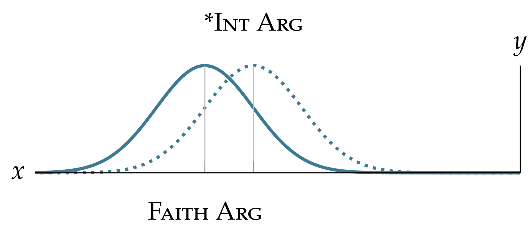

Figure 4.5: Constraint re-ranking distribution in Stochastic Optimality Theory relative to Example (9).

The use of Stochastic Optimality Theory to define a working model of the implicit object construction, aware of the effect of the factors presented in Chapter 3 and of the fact that different transitive verbs are differently prone to be used intransitively, is the topic of Chapter 5.

5 Medina's (2007) model

Medina (2007) created a Stochastic Optimality Theoretic model of the indefinite object drop. An implicit object output, generated by Gen on the basis of the input (Section 5.1), is evaluated against a set of conflicting constraints (Section 5.3). The constraints stem directly from the set of object drop predictors chosen by the author (Section 5.2), and get re-ranked with respect to the verb's semantic selectivity (Section 5.4), computed as the Selectional Preference Strength by Resnik (1993, 1996). The way this model implements a probabilistic ranking of the constraints (detailed in Section 5.5) ensures not only that both implicit and overt objects are allowed in the grammar, but also that an implicit object output has a relative gradient grammaticality across different verbs.

5.1 The input and the output

As I mentioned in Section 4.1, Optimality Theory requires the input to syntactic optimization to contain the relevant lexical and semantic components that will be mapped to syntactically well-formed output forms. Since Medina (2007) defines a model of the indefinite object drop, her input has to provide all the information necessary to generate the two outputs in (1):

(1) a. John was writing.
b. John was writing something.

Hence, at the very least, the input will contain a transitive verb with its complete predicate-argument structure, i.e., a specified subject and an unspecified object. Since the model does not deal with *specified* object drop, the input to the model cannot generate an output with a definite overt object as in (2). To put it better, Medina (2007, pp. 70–71) observes that this output may actually be in the candidate set, but it is always ruled out by a high-ranking faithfulness constraint that keeps output candidates from containing semantically richer information than in the input.

(2) John was writing a book.

In addition to this, the input will also feature all the relevant predictors of object drop used by the author, i.e., semantic selectivity (operationalized as Resnik's Selectional Preference Strength), telicity, and perfectivity. Thus, inputs in Medina (2007) have the form (3):

(3) verb (x,y), x = subject, y = unspecified, SPS = *numerical value*, [+ Past], [± Telic], [± Perfective]

The tense feature is fixed at [+ Past] since the author only modeled past tenses for consistency, but there is no hard theoretical constraint on this. In theory, it would indeed be possible to create similar inputs with other tenses. Looking at a specific case, the input generating the outputs in (1) would look like (4).

(4) write (x,y), x = John, y = unspecified, SPS = 2.54[1] , [+ Past], [-Telic], [- Perfective]

1: This is the value of Selectional Preference Strength that Resnik obtained by performing the computation over the Brown corpus of English (Resnik 1996, p. 150), also reported in Medina (2007, p. 114).

Let us take a closer look at how the three predictors chosen by Medina are implemented in her model.

5.2 Predictors

5.2.1 Semantic selectivity

Medina (2007) uses semantic selectivity as a measure of the recoverability of a transitive verb's direct objects, which the reader will remember being a major determiner of object drop based on the information discussed in Section 3.1.1. As a quick recap, I will just note that the recoverability of a direct object (or, better, of the broad semantic class it belongs to) of a verb correlates with the grammaticality of sentences featuring that same verb used intransitively, as in (5).

(5) a. John ate $\varnothing_{\text{dObj}}$.
$\longrightarrow$ The omitted object belongs to the category of Edibles.
b. *John made $\varnothing_{\text{dObj}}$.
$\longrightarrow$ The omitted object can be virtually anything.

In theory, it would be possible to treat the semantic selectivity of a transitive verb with respect to its direct objects as a binary feature and be done with it. However, this choice would be quite poor both from a methodological point of view, since there is no clear-cut criterion to tell apart recoverable-object and non-recoverable-object verbs, and from a usability point of view, since binary selectivity would be scarcely informative with respect to object drop, semantic recoverability being a gradient notion.
Making use of the experimental literature on the matter, Medina (2007) decided to operationalize semantic selectivity using the Selectional Preference Strength measure developed by Resnik (1993, 1996). Resnik quantifies the selectional preferences of transitive verbs in an information-theoretical model which encodes semantic selectivity as the relative entropy between the distribution of WordNet (Beckwith et al. 1991; G. A. Miller 1995) classes for all the direct objects in a corpus and the distribution of WordNet classes for the direct objects of a specific verb. I will discuss the mathematical meaning and the computational details of Resnik's Selectional Preference Strength in Section 6.1, where I will also present an update of his measure I contributed to create (Cappelli and Lenci 2020), powered by distributional semantics and word embeddings. Here, I will only point out how suitable Resnik's measure is for the purposes of Medina's model of indefinite object drop. First of all, the

Selectional Preference Strength score assigned to a verb is a numerical value (a non-null positive real number) on a continuous scale, which it makes it perfect to capture the fact that semantic selectivity is not a binary variable. In particular, the narrower the selectional preferences of a verb are (i.e., the more recoverable its direct objects are), the greater the Selectional Preference Strength of that verb is. Most conveniently for Medina's thesis in acquisitional linguistics, whose scope goes far beyond the coding of computational technicalities, Resnik (1996, p. 150) provides the Selectional Preference Strength scores computed for 34 English verbs over the Brown corpus (Kučera and Francis 1967), the CHILDES corpus (MacWhinney 2000), and human subject norms. While being a limited set of data, this collection of scores has all the data Medina needed to build her model of the implicit object construction, and it is still a valuable resource for anyone looking into models of semantic selectivity.

5.2.2 Telicity

Telicity has been proven to be an important predictor of the omissibility of direct objects in Section 3.2.1. Medina encodes it as a binary variable, so that the transitive verbs she tested (Resnik's 34 ones) are tagged as either telic or atelic. Doing so, however, is less straightforward than it may seem at first glance.

Telicity, in its typical interpretation, is a property of predicates, in conjunction with their grid of arguments and the way they are filled by linguistic material (Dowty 1979 [2012]; Vendler 1957). This means that the (a)telicity feature may only be assigned to complete verb phrases, complete with arguments and adjuncts, and not to bare verb heads (see the simplified tree in Figure 5.1). Under this view, predicates are telic if the events they describe have an endpoint (encoded as a direct object), or they are atelic if they don't. As a logical consequence, the same verb would be considered telic when used transitively and atelic when used intransitively (Mittwoch 1982; Olsen and Resnik 1997).

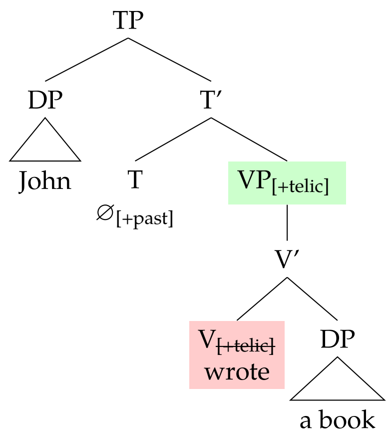

Figure 5.1: Simplified syntax tree illustrating the traditional interpretation of telicity as a property of predicates, not verb heads.

As Medina points out, telicity taking scope over a whole verb phrase poses a severe problem with respect to the use of telicity as a predictor of object drop. How can an object-dropping transitive verb be always considered (a)telic, if the very act of adding or dropping its object alters its telicity? Medina's solution to this conundrum makes use of the analysis of telicity as a privative feature by Olsen (1997 [2014]). Let us discuss this approach in more detail.

According to Olsen (1997 [2014], p. 32), telicity is a semantic feature verbs have if the event they denote can have an endpoint or a result, regardless of whether they actually attain it or not. Based on this, [+telic] verbs cannot lose the feature even if the endpoint is not realized as a syntactic constituent. On the contrary, atelicity is a cancelable conversational implicature, so [-telic] verbs can be interpreted as either having or not having an inherent endpoint or result. To prove this point, Olsen provides two series of examples (adapted from Olsen (1997 [2014], p. 33)), reported in (6) and (7). In (6), telicity is shown to be an inherent semantic property with a non-cancelable marked interpretation, since the addition of the durative adverbial 'for years' turns the accomplishments into *iterative* accomplishments, not into activities (which are [-telic], [+durative] predicates).

(6) a. Eli won for years.
 b. Eli ran to the store for years.

Instead, examples in (7) demonstrate the unmarked nature of atelicity, based on the assumption that progressive forms of atelic verbs entail the corresponding perfect form (as in (7-b)), while the same is not true of telic verbs (as in (7-a) and (7-c)). Thus, it is possible to add material and disrupt the atelic interpretation (compare (7-b) with (7-c)), but it is not possible to cancel telicity (as shown in (6-a) and (6-b)).

(7) a. Eli is winning. $\not>$ Eli has won.
 b. Eli is running. $>$ Eli has run.
 c. Eli is running to the store. $\not>$ Eli has run to the store.

Olsen (followed by Medina) actually uses the notation [0 Telic] to indicate the feature of verbs which lack telic denotation, but in this thesis I will use a more transparent opposition between [+telic] and [-telic] verbs. Now informed about the underlying theory (discussed in more detail in Section 3.2.1), the reader will be able to interpret the [-telic] feature as Olsen intended.

By virtue of this particular interpretation of telicity, Medina considers verbs such as *to make* to be [+telic] and verbs such as *to eat* to be [-telic]. The author assigns telicity (or lack thereof) to each verb on the basis of three tests[2] .

2: All the examples in the list are taken *verbatim* from Medina (2007, pp. 302–303).

- **The *in/for* test** (Dowty 1979 [2012]; Vendler 1957). Predicates featuring [+telic] verbs, as in (8-a), allow more easily *in-X-time* temporal adverbs and less easily *for-X-time* temporal adverbs, while predicates with [-telic] verbs, as in (8-b), have the reverse preference pattern.

 (8) a. Michelle made some stuff in/*for five minutes.
  b. Michelle read *in/for five minutes.

- **The *almost* test** (Dowty 1979 [2012]). The adverb *almost* allows [+telic] verbs, as in (9-a), to have two interpretations (the event has begun but has not finished yet; the event has not begun yet), while [-telic] verbs, as in (9-b), can have only one (the event has not begun yet).

(9) a. Tony almost packed.
⟶ Tony started packing, but hasn't finished yet.
⟶ Tony was about to pack but hadn't yet started.
b. Tony almost ate.
⟶ ~~Tony started eating, but hasn't finished yet.~~
⟶ Tony was about to eat but hadn't yet started.

- **The *counting* test** (Bach 1986). It is more acceptable to count [+telic] predicates, as in (10-a), than [-telic] ones, as in (10-b). In other words, counted [+telic] predicates are interpreted as a single event with repeated acts, while counted [-telic] predicates are interpreted as repeated, distinct events. This is the most controversial test out of the three Medina chose to use.

(10) a. Edgar opened some stuff three times.
b. Edgar watched three times.

Medina tested her target verbs in combination with an indefinite object or in an intransitive sentence. A transitive verb tested this way is [+telic] if at least two tests out of three yield a telic interpretation or [-telic] otherwise, based on the assumption that the verb lacks telic denotation (i.e., is atelic) if it can elicit both a telic and an atelic interpretation. As I will discuss in more detail in Section 6.2, while the *in/for* test is a largely reliable diagnostic for telicity, the other two tests are somewhat problematic.

5.2.3 Perfectivity

While telicity (i.e., lexical aspect) is a semantic property of individual verbs, as illustrated in the previous paragraph, perfectivity (i.e., grammatical aspect) in English is morphologically marked. A more detailed discussion of the effects of perfectivity on the grammaticality of indefinite object drop is provided in Section 3.2.2.
As noted in Section 5.1, Medina (2007) only modeled inputs in the past tense. With respect to the morphological markers of (im)perfectivity, the author followed Olsen (1997 [2014]) in realizing [+perfective] inputs with perfect morphology, in the form "*have* + past participle of the verb", as in (11-a), and [-perfective] inputs with progressive morphology, in the form "*be* + verb + *-ing*", as in (11-b).

(11) a. John had written a book.
b. John was writing a book.

In my own models of object drop (presented in Chapter 9 (*Predicting the grammaticality of implicit objects*)), I marked (im)perfective aspect on English verbs using the same morphology as in Medina (2007), and I devised a similar strategy for my Italian stimuli (Section 6.3).

5.3 Constraints and their ranking

For each input to the optimization in Medina's model, the two candidate outputs (one with an overt object, one with an implicit object, as detailed in Section 5.1) are evaluated against the four constraints in (12). Their labels and definitions are taken *verbatim* from Medina (2007, p. 72).

(12) a. *Int Arg (*Internal Argument Structure)
The output must not contain an overt internal argument (that is to say, a direct object).
b. Faith Arg (Faithfulness to Argument Structure)
All arguments in the input must be present in the output.
c. Telic End (Telic Endpoint)
The endpoint of a [+ Telic] event must be bounded by the presence of an overt argument in the output.
d. Perf Coda (Perfective Coda)
The coda of a perfective event [+ Perfective] must be identified by the presence of an overt argument in the output.

Let us discuss each constraint more extensively. *Int Arg[3] is a markedness constraint belonging to a broader class of economy-of-structure constraints, *Struc (Buchwald et al. 2002; Hartkemeyer 2000), operating in syntax and every other domain of grammar. By penalizing candidates with an overt object, which have a greater degree of syntactic structure than candidates with an implicit object, *Int Arg is the only constraint in Medina's set to favor implicit-object candidates. As I will explain in Section 5.4, this unique behavior of *Int Arg is not just intended, but downright necessary for the probabilistic re-ranking of the constraint to happen.

3: In Optimality Theory, constraints are named after features or linguistic elements required in the grammar. An asterisk is added at the beginning of a constraint's name if that constraint is violated by the presence, not the absence, of a particular feature.

Faith Arg is a faithfulness constraint requiring that all arguments in the input be overtly realized in the output, thus penalizing candidates with an implicit object. As discussed in Section 4.1, faithfulness constraints are a unique feature of Optimality Theory, created in order to conflict with economy-of-structure markedness constraints like the framework requires. Without faithfulness constraints, the optimal candidate would always be the least marked one, i.e., the one violating the lowest-ranked constraints (Legendre 2001, p. 3). In the specific case of Medina's model, Faith Arg conflicts directly with *Int Arg, determining a state of affairs that closely resembles the situation previously described in Figure 4.5. The picture is made more complex by the interaction of two additional constraints and the inclusion of semantic selectivity in the model.

Finally, the markedness constraints Telic End and Perf Coda are the Optimality Theoretic implementations of telicity (Section 3.2.1) and perfectivity (Section 3.2.2) as predictors of object drop, both penalizing object-dropping output candidates. Notably, while *Int Arg and Faith Arg are always active regardless of the input features, Telic End and Perf Coda are only actively used by the Eval component if the input is, respectively, [+telic] and [+perfective]. Let us consider the examples in Table 5.1 to Table 5.4, adapted from Medina (2007). The absence of the pointing hand indicates that no winner has been chosen among the candidates, since these examples are just here to illustrate the possible violation profiles determined by different inputs. In the same spirit, the dotted lines show that no constraint ranking (neither in standard nor in stochastic Optimality Theory) has been determined.

A transitive verb which is both telic and perfective, such as *to catch* in Table 5.1, has a full constraint violation profile involving all the four constraints at play in Medina's model. Instead, the output candidates for a telic but imperfective input verb (such as *to catch* in Table 5.2) only violate *Int Arg, Faith Arg, and Telic End. In this situation, Perf Coda is vacuously satisfied, i.e., there is no candidate in the candidate set with the ability to violate the constraint (given that it is violated by object-dropping perfective candidates, and here there is none). Similarly, Telic End is vacuously satisfied in Table 5.3, and both Telic End and Perf Coda are vacuously satisfied in Table 5.4. In these four tableaux, vacuously satisfied constraints are marked in light gray for the sake of clarity. Typically, authors leave them out of their tableaux in Optimality Theoretic literature.

	catch (x,y), x = Jack, y = unspecified, SPS = n/a, [+Past], [+Telic], [+Perfective]	*Int Arg	Faith Arg	Telic End	Perf Coda
a.	Jack had caught.		*	*	*
b.	Jack had caught something.	*			

Table 5.1: Optimality Theory tableau illustrating the constraint violation profile in the model of object drop by Medina (2007), relative to a telic perfective verb.

	catch (x,y), x = Jack, y = unspecified, SPS = n/a, [+Past], [+Telic], [-Perfective]	*Int Arg	Faith Arg	Telic End	Perf Coda
a.	Jack was catching.		*	*	
b.	Jack was catching something.	*			

Table 5.2: Optimality Theory tableau illustrating the constraint violation profile in the model of object drop by Medina (2007), relative to a telic imperfective verb.

	eat (x,y), x = Jack, y = unspecified, SPS = n/a, [+Past], [-Telic], [+Perfective]	*Int Arg	Faith Arg	Telic End	Perf Coda
a.	Jack had eaten.		*		*
b.	Jack had eaten something.	*			

Table 5.3: Optimality Theory tableau illustrating the constraint violation profile in the model of object drop by Medina (2007), relative to an atelic perfective verb.

	eat (x,y), x = Jack, y = unspecified, SPS = n/a, [+Past], [-Telic], [-Perfective]	*Int Arg	Faith Arg	Telic End	Perf Coda
a.	Jack was eating.		*		
b.	Jack was eating something.	*			

Table 5.4: Optimality Theory tableau illustrating the constraint violation profile in the model of object drop by Medina (2007), relative to an atelic imperfective verb.

So far, so good. However, the attentive reader will have noticed that no mention has been made of object recoverability (quantified via semantic selectivity) among the constraints at play, even though it has been said to be a crucial predictor of object drop in Section 3.1.1 and in Section 5.2. As Medina (2007, p. 76) observes, it would indeed be easy to define an *Overt Recoverable Object constraint penalizing overt objects occurring with high-selectivity verbs, or a *Non-recoverable Implicit Object constraint penalizing implicit objects occurring with low-selectivity verbs. This works perfectly for binary predictors such as telicity and perfectivity, but semantic selectivity is not a binary predictor (as noted in Section 5.2.1). Not only is it quantified by means of a continuous numerical variable, but it is structurally impossible to define a threshold value separating high- and low-selectivity transitive verbs.

Yankes (2021 [2022], p. 75) tried to account for the effect of object recoverability[4] by introducing the faithfulness constraint (penalizing object-dropping candidates) in (13). However, Yankes merely states that a candidate violates Info if the verb in the input lacks recoverability "in sufficient degree", without attempting a more rigorous definition or quantification of such a degree beyond personal intuition. The blurry definition of this binary constraint is the only way to make his standard Optimality Theoretic model of indefinite null objects work, but this phenomenon has to be accounted for within a gradience-compatible framework (as argued in Section 4.1.4).

(13) Info(rm): Important, noninferable speech content may not be omitted.

Let us go back to Medina's model, and her implementation of recoverability via Resnik's measure of semantic selectivity. In the original works about a computational model of semantic selectivity as a proxy to argument recoverability (Resnik 1993, 1996), the author himself observed that albeit his Selectional Preference Strength measure correlates well with the acceptability of implicit objects, there are indeed some cases where a high-SPS verb does not allow for its object to be dropped (e.g., *to hang, to wear*). For this reason, as I will detail in Section 5.4, SPS will play a crucial role in Medina's probabilistic constraint ranking, despite not being defined as a constraint *per se*.

The gradient nature of object recoverability, be it computed via SPS (Resnik 1993, 1996) or measures such as PISA (Cappelli and Lenci 2020), makes it a bad candidate for constraint-hood. Moreover, even if it were possible to binarize it into a viable constraint, it would still suffer from a problem afflicting all the binary constraints at play, i.e., out-of-range output generation both within-constraint and across constraints. In other words, each constraint in the model under- or over-generates outputs across verbs, since experimental data show that some telic verbs occur with implicit objects while others do not, that perfective aspect favors implicit objects but does not force them, and so on. This is normal, expected behavior for classic Optimality Theoretic constraints, easily solved by means of re-ranking in order to select the winner candidate. However, in this case the state of affairs is more complicated.

As shown in Table 5.1 to Table 5.4, forcing the implicit object construction into a standard Optimality Theoretic model (as attempted by Yankes (2021 [2022])) has two major drawbacks. One is that the same constraint violations apply to any verb with a given aspectual profile, meaning that a tableau like Table 5.3 would apply to *to eat* and to any other atelic perfective input (since, in standard Optimality Theory, the ranking of the constraints has to be consistent within the same language). As just observed, this does not match the actual linguistic data on indefinite object drop. The second issue with such a model is that it only allows for a single winner out of a candidate set, so that (also considering the previous problem) for a given aspectual profile, regardless of the specific verb in the input, the direct object would be either obligatory, or obligatorily dropped. Looking again at tableaux Table 5.1 to Table 5.4, the model would determine which of the two candidates would win the competition solely based on the relative ranking of *Int Arg with respect to any of the other constraints (since the relative ranking of Faith Arg, Telic End, and Perf Coda is irrelevant for the purposes of electing

a winner). Given that the the felicity of the implicit object construction depends on the interaction of several factors, and given that no factor or combination of factors actively forces or prohibits object drop, the logical conclusion is that a standard Optimality Theoretic model of object drop fails to meet the speaker's intuitions.
Taking all of this into consideration, a non-standard Optimality Theoretic comprehensive model of the implicit object construction based on the linguistic factors hereby described has to account for two aspects, i.e.,

1. incorporating semantic selectivity (a continuous, non-binary factor) into the model, and
2. allowing the model to yield outputs with varying degrees of grammaticality, instead of having a winner and a loser.

5.4 Defining a probabilistic constraint ranking

5.4.1 Introduction

Medina (2007) built such a model under the premises of Stochastic Optimality Theory, whose tenets were introduced in Section 4.2.4. This framework accounts for gradient grammaticality out-of-the-box, since it assigns each candidate a probability of it being the winner, to be interpreted as a degree of grammaticality. As for the implementation of semantic selectivity as a continuous factor, Medina came up with a personal variation on Stochastic Optimality Theory.
As noted in tableaux Table 5.1 to Table 5.4, the winner in a Standard Optimality Theoretic model is determined based on the ranking of *Int Arg relative to the other three constraints at play. The same would hold true for a traditional Stochastic Optimality Theoretic version of the same model, where all constraints would be assigned a given normal distribution with a fixed evaluation noise (fully described by its standard deviation). Medina's stochastic model obtains the probability of *Int Arg dominating the other three constraints as a function of the verb's semantic selectivity, computed using Resnik's Selectional Preference Strength (Resnik 1993, 1996). Compared with traditional Stochastic Optimality Theory, Medina's proposal achieves the same goal of having candidates with *gradient* grammaticality (instead of a single absolute winner), while having two advantages:

- implementing semantic selectivity in the model properly, and
- defining the relative ranking of *Int Arg with respect to the other three constraints as independent computations, yielding fine-grained interactions of semantic selectivity with telicity, perfectivity, and faithfulness to the input.

In such a model, the grammaticality of candidates is assessed across all possible constraint re-rankings, each of which is assigned a probability by the model. Medina's implementation of this model stems from a three-step logic:

1. the probability of *Int Arg dominating each of the other constraints is expressed as a function of the input verb's semantic selectivity, computed via Resnik's Selectional Preference Strength;

2. the values of the function are used to compute the relative probabilities of each of the four possible re-rankings of *Int Arg with respect to the three other constraints at play;
3. these relative probabilities determine the relative probability (and thus grammaticality) of the implicit object output for a given input, depending on semantic selectivity and the input's aspectual type.

In the next Sections, I will explain each step in depth and unveil the underlying mathematical processes. Before doing this, a brief introduction to the mathematical premises of Medina's model is in order.
Let us consider a hypothetical model with n constraints. These constraints could be re-ranked in $n!$[5] ways. However, since in the stochastic model of indefinite object drop the only relevant ranking is that of *Int Arg relative to the other constraints, the mathemathics involved in the hypothetical n-constraint model can be simplified considerably. With this restriction, the constraints in the model can only be re-ranked in n ways, because the focus is on whether one constraint is in first, second... nth position, while the order of all the other $n-1$ constraints is irrelevant. Mathematically, this translates into a trivial case of partial permutation of k elements out of a set of n elements total[6] , where k equals one. As stated at the end of the previous paragraph, the model assigns each constraint re-ranking a value on the 0-1 probability scale. If all ranking orders in the model are equally probable, each of them would have a probability of $\frac{1}{n}$.
Medina's stochastic model of object drop employs the four constraints introduced in Section 5.3, i.e., *Int Arg, Faith Arg, Telic End, and Perf Coda. Applying the math discussed right before, it follows that these four constraints can be re-ranked in 24 different ways. Each ranking selects either an implicit object or an overt object as a winner, based on the aspectual features of the input. To illustrate this point, Table 5.5 reproduces the summary table from Medina (2007, p. 89), rearranging the lines in three different groups to make the reasoning more transparent.

5: This notation stands for "n factorial" and it is equivalent to the number of possible permutations of the set $1, 2, \ldots, n$, i.e., a list of n unique elements. It is computed as $n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 1$.

6: Expressed as "n permute k", and written in the form

$$\frac{n!}{(n-k)!}$$

The summary in Table 5.5 follows directly from the definition of the four constraints in (12) and from the above analysis of tableaux Table 5.1 to Table 5.4, where it was made evident that *Int Arg favors an implicit object output, while the three other constraints are violated by such a candidate. In particular,

1. 6 constraint re-rankings always yield an implicit object output regardless of the aspectual features of the input, because *Int Arg dominates Faith Arg (first group in Table 5.5);
2. 12 constraint re-rankings always yield an overt object output regardless of the aspectual features of the input, because Faith Arg dominates *Int Arg (second group in Table 5.5);
3. 6 constraint re-rankings yield either an implicit or an overt object output based on the aspectual features of the input (last group in Table 5.5), i.e., based on the position of *Int Arg relative to the relevant, non-vacuously satisfied constraints. The relative ranking of the three other constraints is irrelevant, while the position of *Int Arg is crucial since it is the only constraint to favor object drop.

In a floating-constraint approach to the issue of creating a probabilistic model of a linguistic phenomenon, such as Medina's variant of Stochastic Optimality Theory, the ratio of re-ranking orders returning a specific output (in this case, the implicit object output) out of the total number of

	Telic Perf	Telic Imperf	Atelic Perf	Atelic Imperf
$*I \gg F \gg T \gg P$	implicit	implicit	implicit	implicit
$*I \gg F \gg P \gg T$	implicit	implicit	implicit	implicit
$*I \gg T \gg F \gg P$	implicit	implicit	implicit	implicit
$*I \gg P \gg F \gg T$	implicit	implicit	implicit	implicit
$*I \gg T \gg P \gg F$	implicit	implicit	implicit	implicit
$*I \gg P \gg T \gg F$	implicit	implicit	implicit	implicit
$F \gg *I \gg T \gg P$	overt	overt	overt	overt
$F \gg *I \gg P \gg T$	overt	overt	overt	overt
$F \gg T \gg *I \gg P$	overt	overt	overt	overt
$F \gg P \gg *I \gg T$	overt	overt	overt	overt
$F \gg T \gg P \gg *I$	overt	overt	overt	overt
$F \gg P \gg T \gg *I$	overt	overt	overt	overt
$T \gg F \gg *I \gg P$	overt	overt	overt	overt
$T \gg F \gg P \gg *I$	overt	overt	overt	overt
$T \gg P \gg F \gg *I$	overt	overt	overt	overt
$P \gg F \gg *I \gg T$	overt	overt	overt	overt
$P \gg F \gg T \gg *I$	overt	overt	overt	overt
$P \gg T \gg F \gg *I$	overt	overt	overt	overt
$T \gg *I \gg F \gg P$	overt	overt	implicit	implicit
$T \gg *I \gg P \gg F$	overt	overt	implicit	implicit
$T \gg P \gg *I \gg F$	overt	overt	overt	implicit
$P \gg *I \gg F \gg T$	overt	implicit	overt	implicit
$P \gg *I \gg T \gg F$	overt	implicit	overt	implicit
$P \gg T \gg *I \gg F$	overt	overt	overt	implicit

Table 5.5: Set of 24 possible re-rankings of the four constraints *Int Arg, Faith Arg, Telic End, and Perf Coda, with implicit/overt object output based on aspectual features of the input (Medina 2007, p. 89).

rankings can be used as a proxy to the relative frequency of that same output in a corpus, or to its gradient grammaticality as judged by native speakers. Going back to the math discussed above, each ranking would have a 1/24 probability if all of them were equiprobable, making it very straightforward to compute the probability (and hence, grammaticality) of an implicit object output for each of the four aspectual types under examination. Based on Table 5.5, an implicit object output would then be expected:

1. 25% of the time for telic perfective inputs, since 6 rankings out of 24 favor the implicit object construction,
2. 33% of the time both for telic imperfective inputs and for atelic perfective inputs (8/24 rankings),
3. 50% of the time for atelic imperfective inputs (12/24 rankings).

However, the model (and the underlying math) has yet to account for semantic selectivity.

5.4.2 Logical step 1: Re-ranking probability as a function of semantic selectivity

Earlier in this Section, I observed that Medina's model of the gradient grammaticality of the implicit object construction takes full account of semantic selectivity as a continuous factor, and it also defines the ranking of *Int Arg with respect to the other constraints independently. The author achieved this outcome by renouncing the equiprobability tenet assumed so far, defining instead the probability of *Int Arg being ranked above each constraint as a function of Resnik's Selectional Preference

Strength (Resnik 1993, 1996) for each verb. The linguistic reason to do so lies in the role of semantic selectivity as a predictor of object recoverability (and, thus, of object omissibility), discussed extensively in Section 3.1.1, Section 5.2, and Section 6.1. Numerically speaking, the model has to assign a higher probability (hence, grammaticality) to implicit objects occurring with verbs having a higher semantic selectivity, all the while modulating the computation based on the other factors at play.

Stochastic Optimality Theory defines the re-ranking range of each constraint in a model as a probability distribution (Section 4.2.4). However, all constraints re-rank one with respect to another based on the same function and standard deviation, making it impossible to register the effect of varying degrees of semantic selectivity. The solution Medina devised (Medina 2007, p. 94) is to define the probability of *INT ARG being ranked above any other constraint as a linear function (instead of a fixed normal curve) whose value is directly proportional to the Selectional Preference Strength of the verb featured in the input. This way, some constraints among the 24 in Table 5.5 are more probable than others, depending on the semantic selectivity of the verb.

Let us look at the underlying mathematics in more detail. A function $f(x) = y$ is a relation that maps each element x from the domain set X to one and only one value y from the co-domain set Y. A *linear* function, which is the relevant kind of function for the linguistic model under discussion, has the form $y = mx + q$ and is represented as a straight, non-vertical[7] line on a plane. In particular, m and q are numerical constants, the former representing the slope of the function (i.e., its "steepness" with respect to the x axis) and the latter indicating the intercept (i.e., the point where the curve meets the y axis). This is rendered graphically in Figure 5.2, where it is possible to gauge the effects of varying the values of m and q on the shape of the curves.

7: A vertical line (described by the equation $x = k$, where k is a numerical constant) is not a function, because the unique value k from the domain is mapped to more than one value in the co-domain.

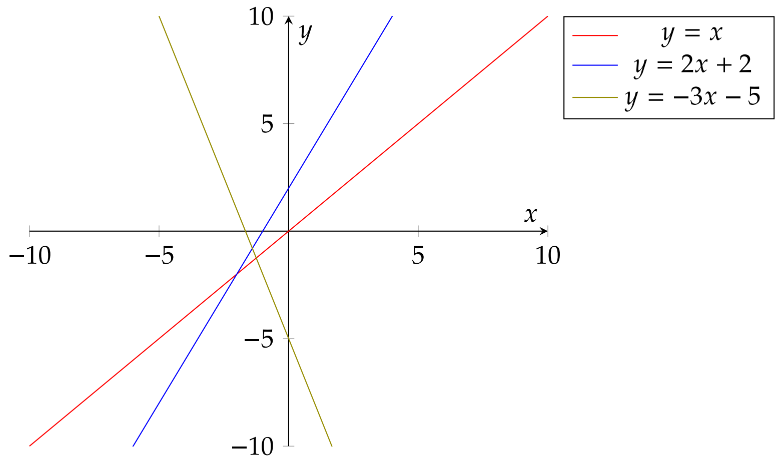

Figure 5.2: Graphical representation of three linear functions as lines on the Euclidean plane.

Here, $y = x$ meets the y axis in the origin because q is null, and it bisects perfectly the first and third quadrants of the cartesian plane because m is equal to 1. The curves described by $y = 2x + 2$ and $y = -3x - 5$ are steeper (i.e., y grows/decreases faster than x), because the absolute value of m is greater than 1, and they meet the y axis at 2 and -5 respectively. Moreover, $y = x$ and $y = 2x + 2$ are *positive* linear functions because their slope is positive, i.e., y values increase when x values do, while the opposite holds for $y = -3x - 5$.

Since any flavor of Optimality Theory requires a conflict between constraints, and since *INT ARG is the one constraint to conflict with all the others in this account of object drop, Medina's model defines a linear function for the pairwise ordering of *INT ARG with respect to FAITH ARG (Equation 5.2), TELIC END (Equation 5.3), and PERF CODA (Equation 5.4), for a total of three different linear functions. In general, each function takes the form in Equation 5.1, where the probability of *INT ARG being ranked above another constraint is y, SPS_i is x, $\frac{\delta_k - \gamma_k}{SPS_{max} - SPS_{min}}$ is m, $\gamma_k - SPS_{min}(\frac{\delta_k - \gamma_k}{SPS_{max} - SPS_{min}})$ is q, and δ_k and γ_k are, respectively, the values the function assumes at SPS_{max} and SPS_{min}.

$$p(\text{*Int Arg} \gg \text{con}) = \frac{\delta_k - \gamma_k}{SPS_{max} - SPS_{min}} \cdot (SPS_i - SPS_{min}) + \gamma_k \quad (5.1)$$

In particular, the three linear functions at play are the ones in Equation 5.2, Equation 5.3, and Equation 5.4.

$$p(\text{*Int Arg} \gg \text{Faith Arg}) = \frac{\delta_1 - \gamma_1}{SPS_{max} - SPS_{min}} \cdot (SPS_i - SPS_{min}) + \gamma_1 \quad (5.2)$$

$$p(\text{*Int Arg} \gg \text{Telic End}) = \frac{\delta_2 - \gamma_2}{SPS_{max} - SPS_{min}} \cdot (SPS_i - SPS_{min}) + \gamma_2 \quad (5.3)$$

$$p(\text{*Int Arg} \gg \text{Perf Coda}) = \frac{\delta_3 - \gamma_3}{SPS_{max} - SPS_{min}} \cdot (SPS_i - SPS_{min}) + \gamma_3 \quad (5.4)$$

These functions take positive values in a range of possible values depending on the verbs' semantic selectivity.

5.4.3 Logical step 2: relative probabilities of the 4 constraint re-rankings

At the beginning of this Section, it was observed that four constraints result in 24 possible re-ranking orders (listed in Table 5.5), and that this large set of permutations actually reduces to just four possible re-rankings if one only cares for the relative position of *INT ARG. In fact, the model will favor an implicit object output whenever *INT ARG is ranked above all the relevant constraints, regardless of their order with respect to one another. A telic perfective input only results in an implicit object output when *INT ARG is ranked above the three other constraints, a telic imperfective input when *INT ARG is ranked above FAITH ARG and TELIC END, an atelic perfective input when *INT ARG is ranked above FAITH ARG and PERF CODA, and an atelic imperfective input when *INT ARG is ranked above FAITH ARG alone. This results in the four possible orderings in Table 5.6.

Table 5.6: Set of the four possible re-rankings of *Int Arg with respect to Faith Arg, Telic End, and Perf Coda, these being unordered with respect one to another.

	Telic Perf	**Telic Imperf**	**Atelic Perf**	**Atelic Imperf**
*I ≫ {F, T, P}	implicit	implicit	implicit	implicit
P ≫ *I ≫ {F, T}	overt	implicit	overt	implicit
T ≫ *I ≫ {F, P}	overt	overt	implicit	implicit
{T, P} ≫ *I ≫ F	overt	overt	overt	implicit

The probability of each individual re-ranking ordering in Table 5.6 is equal to the joint probabilities of the independent pairwise orderings that comprise it. For instance, the probability of *Int Arg being ranked above all the other constraints is equal to the intersection (represented with ∩ in set theory) between the probability spaces of *Int Arg being ranked above Faith Arg, *Int Arg being ranked above Telic End, and *Int Arg being ranked above Perf Coda. This is rendered graphically in the Venn diagram in Figure 5.3, where the probability of *Int Arg being ranked above all the other constraints is colored in blue.

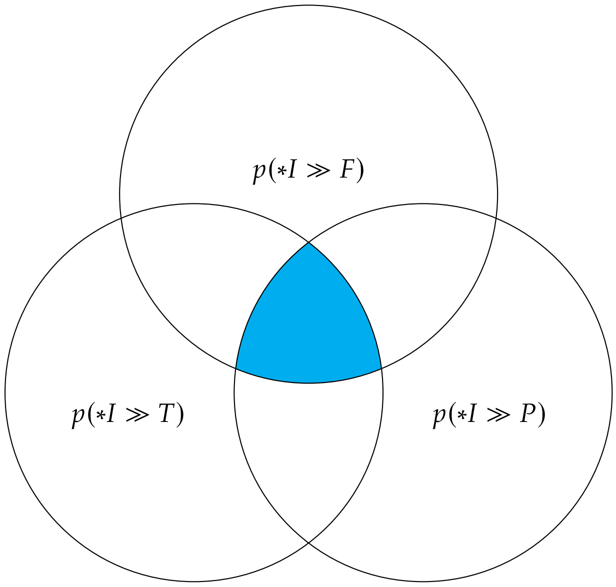

Figure 5.3: Graphical representation of the probability of *Int Arg being ranked above the three other constraints, in cyan in the Venn diagram.

In summary, the probabilities of the four rankings in Table 5.6 are computed as in Equation 5.5 to Equation 5.8. Whenever an equation features the subtraction of the probability of an event from 1, it means that the computation is taking into account the cases when that event is *not* happening, given that probabilities can vary between 0 (impossible event) and 1 (certain event). For instance, since in Equation 5.6 Faith Arg ranks above *Int Arg, the partial probability included in the computation is [1 - p(*Int Arg ≫ Faith Arg)], namely the probability of *Int Arg *not* being ranked above Faith Arg.

$$p(*I \gg F, T, P) = p(*I \gg F) \cdot p(*I \gg T) \cdot p(*I \gg P) \tag{5.5}$$

$$p(P \gg *I \gg F, T) = p(*I \gg F) \cdot p(*I \gg T) \cdot [1 - p(*I \gg P)] \tag{5.6}$$

$$p(T \gg *I \gg F, P) = p(*I \gg F) \cdot [1 - p(*I \gg T)] \cdot p(*I \gg P) \tag{5.7}$$

$$p(T, P \gg *I \gg F) = p(*I \gg F) \cdot [1 - p(*I \gg T)] \cdot [1 - p(*I \gg P)] \tag{5.8}$$

5.4.4 Logical step 3: grammaticality of implicit object output as sum of probabilities for each aspectual type

Returning to Table 5.6, the next step consists in determining the probability of an implicit object output for each of the four aspectual features in the input (the columns of the table). It is possible to achieve this result by summing the probabilities of the individual partial orderings where *Int Arg is ranked above the relevant constraints (the rows of the table). Thus, the likelihood of the object being dropped for each aspectual type is computed as in Equation 5.9 to Equation 5.12.

$$p(\text{implicit})_{\text{Telic Perfective}} = p(*I \gg F, T, P) \tag{5.9}$$

$$p(\text{implicit})_{\text{Telic Imperfective}} = p(*I \gg F, T, P) + p(P \gg *I \gg F, T) \tag{5.10}$$

$$p(\text{implicit})_{\text{Atelic Perfective}} = p(*I \gg F, T, P) + p(T \gg *I \gg F, P) \tag{5.11}$$

$$p(\text{implicit})_{\text{Atelic Imperfective}} = p(*I \gg F, T, P) + p(T \gg *I \gg F, P) + \\ + p(P \gg *I \gg F, T) + p(T, P \gg *I \gg F) \tag{5.12}$$

As stated at the beginning of this Section, in the stochastic model of the implicit object construction by Medina (2007) these probabilities indicate the gradient grammaticality of indefinite object drop for each aspectual type of input.
Atelic imperfective inputs violate a proper subset of the constraints violated by other aspectual types of input. Hence, also considering Table 5.6 and the probabilities Equation 5.9 to Equation 5.12, it is evident that the implicit object output is most likely to be grammatical with atelic imperfective inputs. Following this line of reasoning, it results that telic perfective inputs have the lowest probability to yield an implicit object output, and that the likelihood of atelic perfective and telic imperfective inputs to allow for the object to be dropped is intermediate. The relative object-dropping probability of atelic perfective inputs with respect to telic imperfective inputs depends on the parameters of the functions in Equation 5.2, Equation 5.3, and Equation 5.4. This state of affairs is depicted in Figure 5.4, an adaptation of the graph provided by Medina (2007, p. 108).

The curves in Figure 5.4 are cubic functions because the computations in Equation 5.5 to Equation 5.8 involve the multiplication of the three linear functions of the type illustrated in Equation 5.1. This means that the unknown SPS_i (the x of the linear function) gets multiplied three times by itself, resulting in ${SPS_i}^3$, which yields a cubic curve. Of course, all the parameters in the complex function in Equation 5.1 are part of the newly created cubic function, but they do not result in a higher grade (i.e., more than cubic) polynomial function since they are not unknowns of the function.

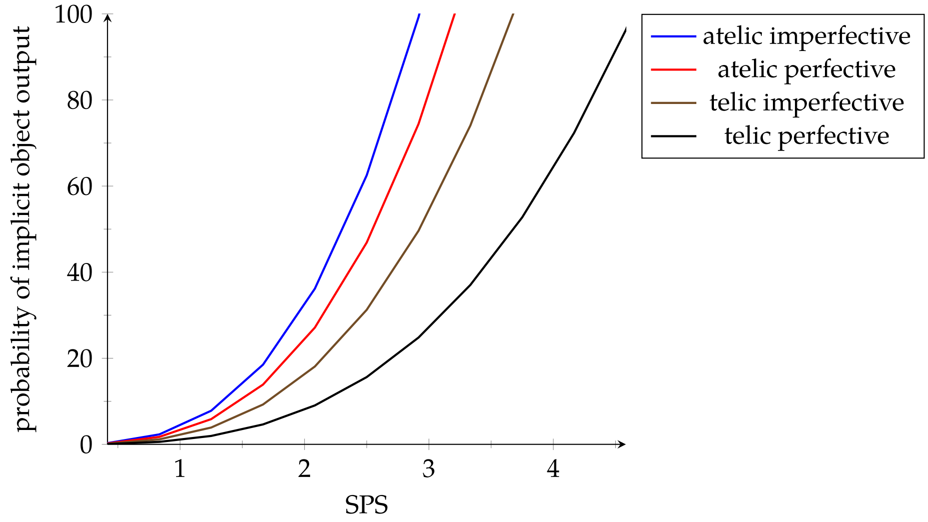

Figure 5.4: Hypothetical representation of the relationship between semantic selectivity and the probability (as a proxy to grammaticality) of an implicit object output in the stochastic model by Medina (2007).

5.5 Implementing a probabilistic constraint ranking

5.5.1 Introduction

The three-step logic illustrated in Section 5.4 is mirrored by the three-step procedure used to computationally implement the probabilistic constraint ranking. The computational steps devised by Medina are as follows:

1. the grammaticality of the indefinite object drop is quantified via an acceptability judgment survey, the results thereof are equated to the probability of an implicit object output for a given input;
2. the probability of each of the four possible constraint orderings can be estimated via the probability of an implicit object output;
3. knowing the probability of each constraint ordering, it is possible to estimate the probability of *Int Arg dominating each constraint.

As is evident from comparing the logical and the computational steps (see Table 5.7), the technical implementation of the model goes backwards with respect to the underlying logic.

Table 5.7: Three-step design of Medina's model, where the computational steps mirror the logical steps.

Logic ↓	
step 1	step 3
step 2	step 2
step 3	step 1
	↑ **Computation**

Let us discuss each computational step in more detail.

5.5.2 Computational step 1: Collecting acceptability judgments

The collection of acceptability judgments is not a part of the computational procedure *per se*, unlike the estimation of the parameters of the linear functions. However, since acceptability judgments are directly equated to the probability of an implicit object output, the fine points of their collection belong to this Section nonetheless. I will now present the main aspects of Medina's experimental design, which the interested reader can integrate with additional information by Medina (2007, pp. 110–134). The linguistic predictors of object drop under consideration are semantic selectivity, telicity, and perfectivity. Among these, semantic selectivity and telicity[8] are inherent properties of each verb, while perfectivity is a sentence-level property. For this reason, the verbs in Medina's experiment are annotated with respect to their semantic selectivity and telicity, while sentences are manipulated for perfectivity. This results in a 2x2 factorial design[9] where each verb, having its own selectivity and telicity profile, appears both with and without a direct object, both having and lacking the perfectivity feature. The examples in (14) are adapted from Medina (2007, p. 113).

8: Based on Olsen (1997 [2014]) and pertaining observations throughout this chapter. Refer to Section 3.2.1 for more details.

9: An experimental design with two independent variables having two levels each, resulting in four experimental conditions.

(14) a. Michael had brought.
b. Michael was bringing.
c. Sarah had brought a gift.
d. Sarah was bringing a gift.

The telicity of each verb was deemed to be [+Telic] if two out of three tests yielded a telic interpretation, [-Telic] otherwise. As discussed in Section 6.2 in regards to the telicity tests of choice in my own model, there is something to be said about the feasibility of this particular set of tests. However, they yielded very consistent results, and henceforth they cannot be considered as a crack in the otherwise solid foundation of Medina's design. The three tests used by Medina (2007, pp. 302–303), first introduced in Section 5.2.2, are as follows.

- **The *almost* test.** Predicates marked as [+telic] and appearing with the adverb *almost* (e.g., *Tony almost packed*) can be interpreted as describing either an event that has begun but has not finished, or an event that has not yet begun. On the contrary, [-telic] predicates with *almost* (e.g., *Tony almost ate*) can only get the second interpretation.
- **The *in/for* test.** [+telic] predicates are more natural with *in X time* as an adjunct (e.g., *Michelle made some stuff in/*for five minutes*), while [-telic] ones prefer *for X time* (e.g., *Michelle read *in/for five minutes*).
- **The *counting* test.** Counting a [+telic] predicate results in a natural interpretation where it denotes multiple, separate events (e.g., *Edgar opened some stuff three times*), while counted [-telic] predicates appear as if denoting multiple instances of the same event (e.g., *Edgar watched some stuff three times*).

Using a set of 30 transitive verbs of interest and 10 intransitive verbs resulted in a set of 160 sentences to be used as experimental stimuli (because of Medina's 2x2 design). The 30 transitive verbs were the same used by Resnik (1993, 1996) to test his Selectional Preference Strength, since Medina (2007) employs the same measure to quantify semantic selectivity. Of the 160 stimuli, the 40 intransitive sentences resulting from the 10 intransitive non-target verbs were used as fillers to distract the participants from the real focus of the experiment, the 60 transitive sentences with an overt object were used as controls (since they had to be grammatical by default), and the remaining 60 sentences featuring a transitive verb used *without* a direct object were the actual target of the experiment.

A total of 15 native speakers of English were recruited as participants among the undergraduate students of Johns Hopkins University and rewarded class credit for their effort. Each of them saw all the stimuli in every experimental condition in randomized order, i.e., the experiment followed a within-subject crossed design. Participants partook in a short training session with 3 mock stimuli before accessing the experiment proper, and received immediate feedback. Both in the training session and the experimental session, participants had to score each stimulus on a 5-point Likert scale ranging from 1 (ungrammatical) to 5 (fully grammatical).

5.5.3 Computational step 2: Judgments and probabilities

The acceptability judgments thus obtained for each verb in each experimental condition were considered equal to the probability of the implicit object output to be returned by the stochastic model based on all the possible re-rankings of the constraints at play for that given input. The judgment scores were then used to estimate the probabilities of the four possible constraint rankings in Table 5.6, which in turn were used to estimate the probabilities of *Int Arg dominating each of the other three constraints. Let us retrace the computational steps that lead to this result.

As shown in Section 5.4.4, the probability of an implicit object output for each aspectual type of input is equal to the sum of the probabilities of the individual partial orderings where *Int Arg is ranked above the relevant constraints. This is shown in Equation 5.9 to Equation 5.12, reported here again for ease of consultation.

$$p(\text{implicit})_{\text{Telic Perfective}} = p(*I \gg F, T, P) \qquad \text{(Equation 5.9)}$$

$$p(\text{implicit})_{\text{Telic Imperfective}} = p(*I \gg F, T, P) + p(P \gg *I \gg F, T) \qquad \text{(Equation 5.10)}$$

$$p(\text{implicit})_{\text{Atelic Perfective}} = p(*I \gg F, T, P) + p(T \gg *I \gg F, P) \qquad \text{(Equation 5.11)}$$

$$\begin{aligned} p(\text{implicit})_{\text{Atelic Imperfective}} &= p(*I \gg F, T, P) + p(T \gg *I \gg F, P) + \\ &+ p(P \gg *I \gg F, T) + p(T, P \gg *I \gg F) \end{aligned} \qquad \text{(Equation 5.12)}$$

Based on the calculations in Equation 5.5 to Equation 5.8, the aforementioned sums of probabilities can be computed as follows in Equation 5.13 to Equation 5.16, where the probability of *Int Arg outranking *all* the relevant constraints is equal to the joint probability of the independent pairwise rankings of *each* constraint with respect to *Int Arg (as explained in Section 5.4.3).

$$p(\text{implicit})_{\text{Telic Perfective}} = p(*I \gg F) \cdot p(*I \gg T) \cdot p(*I \gg P) \quad (5.13)$$

$$p(\text{implicit})_{\text{Telic Imperfective}} = p(*I \gg F) \cdot p(*I \gg T) \cdot p(*I \gg P) + \\ + p(*I \gg F) \cdot p(*I \gg T) \cdot [1 - p(*I \gg P)] \quad (5.14)$$

$$p(\text{implicit})_{\text{Atelic Perfective}} = p(*I \gg F) \cdot p(*I \gg T) \cdot p(*I \gg P) + \\ + p(*I \gg F) \cdot [1 - p(*I \gg T)] \cdot p(*I \gg P) \quad (5.15)$$

$$p(\text{implicit})_{\text{Atelic Imperfective}} = p(*I \gg F) \cdot p(*I \gg T) \cdot p(*I \gg P) + \\ + p(*I \gg F) \cdot [1 - p(*I \gg T)] \cdot p(*I \gg P) + \\ + p(*I \gg F) \cdot p(*I \gg T) \cdot [1 - p(*I \gg P)] + \\ + p(*I \gg F) \cdot [1 - p(*I \gg T)] \cdot [1 - p(*I \gg P)] \quad (5.16)$$

5.5.4 Computational step 3: Parameter estimation

The last step involves Equation 5.2, Equation 5.3, and Equation 5.4, where the ranking of *Int Arg with respect to the other three constraints was defined as a (linear) function of the input verb's semantic selectivity. Plugging them in Equation 5.13 results in the (cubic) function in Equation 5.17, describing the probability of an implicit object output with a telic perfective input depending on the specific verb's semantic selectivity.

$$p(\text{implicit})_{\text{Telic Perfective}} = [\frac{\delta_1 - \gamma_1}{SPS_{max} - SPS_{min}} \cdot (SPS_i - SPS_{min}) + \gamma_1] \cdot \\ \cdot [\frac{\delta_2 - \gamma_2}{SPS_{max} - SPS_{min}} \cdot (SPS_i - SPS_{min}) + \gamma_2] \cdot \\ \cdot [\frac{\delta_3 - \gamma_3}{SPS_{max} - SPS_{min}} \cdot (SPS_i - SPS_{min}) + \gamma_3] \quad (5.17)$$

The value of the function, i.e., $p(\text{implicit})_{\text{Telic Perfective}}$, is the average acceptability judgment for a telic perfective input with a known SPS, i.e., SPS_i in the equation. The values of SPS_{max} and SPS_{min} are known, since they are the maximum and minimum Selectional Preference Strength values in Resnik's list, respectively. Henceforth, the equation has only to be solved for δ_i and γ_i, which are the values the function takes at SPS_{max} and SPS_{min} respectively. A similar reasoning applies to telic imperfective, atelic perfective, and atelic imperfective inputs, once again plugging Equation 5.2, Equation 5.3, and Equation 5.4 into Equation 5.14, Equation 5.15, and Equation 5.16.
At the end of this process, the experimenter is left with a set of n equations, each having six unknowns to be solved for, where n is the total number of target sentences among the stimuli. In order to calculate δ_i and γ_i, the judgments (i.e., the probabilities of object drop which serve as values of the functions) underwent two preprocessing steps in Medina's pipeline, namely:

1. for each target sentence, the 15 judgments provided by 15 participants were averaged into a single numerical value, and
2. this numerical value was converted linearly to fall between 0 and 1, since this is the proper range for probabilities.

Medina estimated the values of the unknown parameters using Excel Solver[10] , based on the following two constraints:

10: An add-in component of Microsoft Excel used to find the optimal value of a function based on several constraints.

- δ_i and γ_i have to fall between 0 and 1;
- the sum-squared error between the predictions of the model and the actual grammaticality judgment data have to be minimized.

Thanks to these constraints, the model outputs predicted grammaticality values in the 0-1 probability range.

5.5.5 Summary of results

The values of δ_i and γ_i Medina found as a result of this optimization indicate that, in English, *Int Arg is more likely to dominate each of the other three constraints when the verb is highly semantically selective, and this is most evident for Telic End. The actual values computed by the model are in Table 5.8.

Table 5.8: Values of unknown parameters δ_i and γ_i as computed in Medina's stochastic model.

	γ	δ
p(*Int Arg ≫ Faith Arg)	0.70	0.82
p(*Int Arg ≫ Telic End)	0.36	1.00
p(*Int Arg ≫ Perf Coda)	0.65	0.80

These results are visualized in Figure 5.5, adapted from the figures in Medina (2007, pp. 143–144).

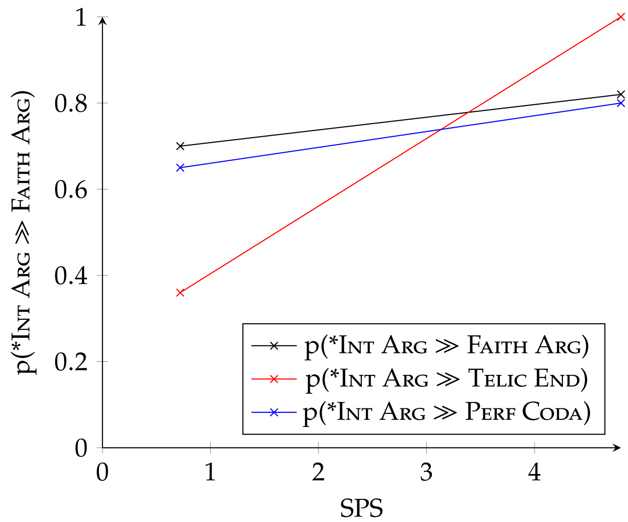

Figure 5.5: Graphical representation of the values of γ and δ in Medina's model.

Knowing the actual values of these parameters, it is possible to plug them into Equation 5.17 (and the equivalent functions for the three other aspectual types of input) to let the model predict the grammaticality of an implicit object output in terms of its probability as a function of semantic selectivity. Graphically, the state of affairs that was hypothesized in Figure 5.4 looks like in Figure 5.6 (adapted from Medina (2007, p. 145)) in Medina's model of indefinite object drop in English. The figure shows the probability of an implicit object output for all possible aspectual types depending on SPS values ranging from 0 to 5, based on the γ and δ values estimated in the model (shown in Table 5.8).

Finally, it results that:

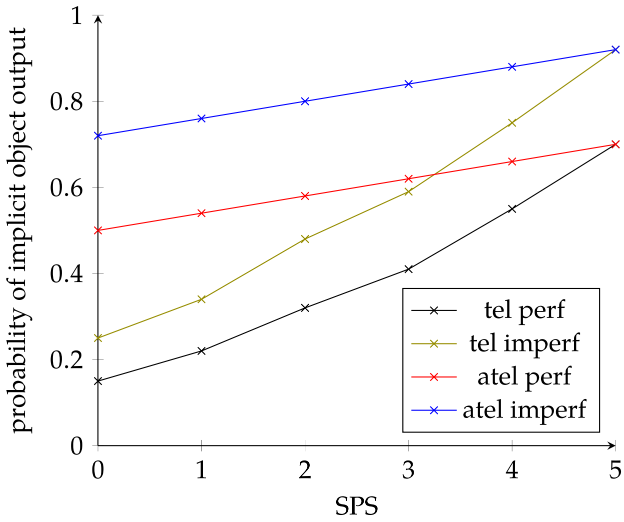

Figure 5.6: Realistic representation of the relationship between semantic selectivity and the probability (as a proxy to grammaticality) of an implicit object output in Medina's model, based on computed γ and δ values.

- the probability of an implicit object output is directly proportional to the verb's SPS for all aspectual types of input;
- the relative probabilities hypothesized in Figure 5.4 are also shown in the actual results, with the implicit object construction being most likely with atelic imperfective inputs and least likely with telic perfective inputs;
- due to semantic selectivity, there is an interaction between aspectual features so that telic imperfective inputs are less likely to accept object drop than atelic perfective inputs when the verb's SPS is low (approximately lower than 3 in Resnik's and Medina's case), while the opposite happens for verbs with a higher SPS;
- in Figure 5.6, the curves for inputs having the same telicity feature are more-or-less parallel because of the strong telicity effect shown in the steep p(*Int Arg ≫ Telic End) curve in Figure 5.5, where it is also easy to spot the interaction between telicity and perfectivity that was mentioned in the previous bullet point.

In Chapter 6, I will open the experimental part of this thesis by presenting the five predictors I included in my own Stochastic Optimality Theoretic model of the implicit indefinite object construction in English and Italian.

Experiments and results

6 Linguistic factors used as predictors

This Chapter presents the five linguistic factors I will use as predictors in my Stochastic OT model of implicit indefinite objects, defined in Chapter 9. The theory linking each of these factors to the omissibility of direct objects was discussed in Chapter 3. In particular, I explained my reasoning behind these choices in Section 3.5.

6.1 Recoverability

As discussed throughout Chapter 2 and in Section 3.1.1, direct objects are likely omitted if they are "sufficiently recoverable" (Glass 2013) from the meaning of the verb. Following Medina (2007), object recoverability (as a predictor in the Stochastic Optimality Theoretic model proposed in this thesis) is a continuous variable, for two main reasons stated in Chapter 5. First, it would not be possible to set a fixed threshold value to separate recoverable and non-recoverable object verbs. Second, a continuous predictor works especially well within a model of *gradient* grammaticality, such as the one hereby provided.
While Medina only employs the taxonomy-based Selectional Preference Strength measure by Resnik (1993, 1996), I will model object recoverability using three different measures of a verb's semantic selectivity as a proxy to object recoverability:

- SPS, the taxonomy-based Selectional Preference Strength measure by Resnik (1993, 1996);
- Computational PISA, a novel similarity-based measure of Preference In Selection of Arguments by Cappelli and Lenci (2020) leveraging distributional semantics;
- Behavioral PISA, a behavioral measure inspired by Computational PISA and computed as the Object Similarity measure by Medina (2007), based on human similarity judgments.

These three measures are detailed in this Section.

6.1.1 Resnik's SPS

Description Resnik (1993, 1996) was the first to link the recoverability of direct objects to the selectional preferences of transitive verbs in a computational model, substantiating this claim by showing that his measure of selectional preference correlates well with plausibility and typicality judgments provided by human subjects. Resnik observed that the distribution of the classes of entities used as direct objects in a corpus regardless of the predicate (called "prior distribution") is different from the distribution of the same classes used as direct objects of a given verb (called "posterior distribution"). A graphical representation of this observation, relative to the verb *to grow*, is provided in Figure 6.1.

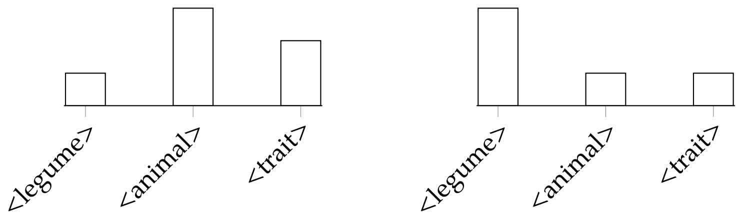

Figure 6.1: Hypothetical representation of the prior (on the left) and posterior (on the right) distribution of direct object classes with respect to the verb *to grow*, adapted from Resnik (1993, p. 54).

Under this view, selectional preferences are fully encoded by the change between the prior and the posterior distribution. Resnik implemented this intuition with his Selectional Preference Strength (SPS) measure, defined as the Kullback-Leibler divergence (relative entropy) between the two distributions, as in Equation 6.1.

$$SPS_{v,r} = \sum_{c \in classes} p(c|v,r) \log_2 \frac{p(c|v,r)}{p(c|r)} \tag{6.1}$$

In Equation 6.1, $p(c|v,r)$ is the posterior distribution of the argument classes (each called c) occurring with a given verb v in a given relation r with the verb, and $p(c|r)$ is the prior distribution of the argument classes participating in the r relation with any verb. The only relevant relation for the purposes of this thesis is the verb-object relation, but it can also be any other grammatical function (such as verb-subject) or semantic role (such as the verb-Instrument relation, as modeled in Cappelli and Lenci (2020)).
The probabilities in Equation 6.1 are estimated from corpus frequencies. Thus, assuming that only nouns participating in the verb-object relation are considered, $p(c|v,r)$ is computed as in Equation 6.2, and $p(c|r)$ is computed as in Equation 6.3.

$$p(c_i|v,r) = \frac{f(v,r,c_i)}{\sum_{c \in classes} f(v,r,c)} \tag{6.2}$$

$$p(c_i|r) = \frac{f(r,c_i)}{\sum_{c \in classes} f(r,c)} \tag{6.3}$$

The frequencies associated to each class are equated to the sum of the frequencies of the nouns (each called n) belonging to that class, since they cannot be extracted directly from the corpus. Thus, $f(v,r,c)$ is computed as in Equation 6.4 and $f(r,c)$ is computed as in Equation 6.5.

$$f(v,r,c) = \sum_{n|c \in hyp(n)} f(v,r,n) \tag{6.4}$$

$$f(r,c) = \sum_{n|c \in hyp(n)} f(r,n) \tag{6.5}$$

In order to overcome the problem of word sense disambiguation posed by polysemous nouns, Resnik (1993, p. 28) proposes to "distribute the credit" for a noun uniformly over its possible classes, namely, to average the frequency of the noun over all the k classes subsuming that noun, as in Equation 6.6.

$$f(r, n) = \frac{\sum_{i=1}^{k} f(r, n, c_i)}{k} \tag{6.6}$$

Given a set of transitive verbs used in a corpus, SPS scores will be higher for transitive verbs having a narrow selectional range (as in (1-a)) and lower for those having a broader selectional range (as in (1-b)).

(1) a. John ate $\varnothing_{\text{object}}$.
    b. *John made $\varnothing_{\text{object}}$.

Crucially, selectional preferences as measured by SPS provide insight about the recoverability of direct objects, and their recoverability affects the grammaticality of the sentences in (1). In particular, *to eat* selects mostly for edible items, so that when reading (1-a) we can reasonably assume John ate some kind of food, while *to make* selects for a wide array of objects from semantically different classes, so it is impossible to know what is that John made in (1-b).
The classes considered when computing the SPS scores have to belong to a lexical taxonomy, so that each sense of each noun in the lexicon is mapped to a concept. Using Resnik's example (Resnik 1993, p. 59), the noun *baseball* can be mapped to two concepts, i.e., a hyponym of the concept <ball> and a hyponym of the concept <field game>. Moving from concepts to taxonomy classes, a noun belongs to any class having one of its concepts as a hyponym, even indirectly. Thus, as illustrated in Figure 6.2, *baseball* belongs not only to the classes <ball> and <field game>, but also to <game equipment>, <artifact>, <inanimate object>, <outdoor game>, <sport>, <human activity>, and <entity> (the root node of the whole hierarchy).

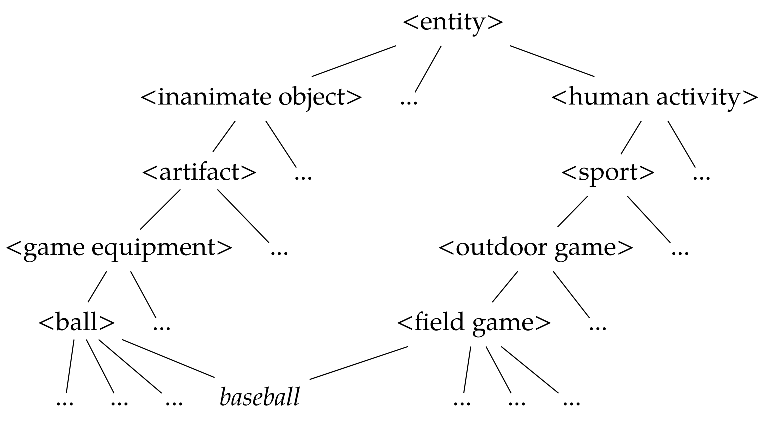

Figure 6.2: Simplified representation of the noun *baseball* in the lexical taxonomy.

In order to compute SPS scores, Resnik adopts WordNet (Beckwith et al. 1991; G. A. Miller 1995) as a computational model of the lexical taxonomy. Its appeal to the author (Resnik 1993, p. 32) lies in that the WordNet taxonomy encodes knowledge in an explicit, hierarchical fashion, which is "intuitively reasonable" and "widely accepted".
Resnik (1993, 1996) also defines the Selectional Association (SA) of a

verb-relation-*class* triple as the ratio of the SPS for that class and the overall SPS of the verb-relation pair (see Equation 6.7).

$$SA_{v,r,c} = \frac{p(c|v,r)\, log \frac{p(c|v,r)}{p(c|r)}}{SPS_{v,r}} \tag{6.7}$$

Then, the SA of a verb-relation-*argument* triple is defined as the highest verb-relation-*class* SA among those computed for each WordNet class the argument belongs to, as in Equation 6.8.

$$SA_{v,r,a} = \max_{c_i \in hyp(a)} \frac{p(c_i|v,r)\, log \frac{p(c_i|v,r)}{p(c_i|r)}}{SPS_{v,r}} \tag{6.8}$$

Resnik (1996, p. 142) computed SA scores for a set of transitive verbs in English, finding that verb-object pairs where the object is plausible for the verb (such as *read-article, write-letter*) received a significantly higher SA score than pairs featuring an implausible object (such as *read-fashion, write-market*).

Resnik's work inspired new models of SA over the years (Abe and Li 1996; Alishahi and Stevenson 2007; Bergsma, Lin, and Goebel 2008; Brockmann and Mirella Lapata 2003; Ciaramita and Johnson 2000; S. Clark and D. Weir 2001; Grishman and Sterling 1992; Nadejde, Birch, and Koehn 2016; U. Padó, M. W. Crocker, and Keller 2009; Shutova, Tandon, and Melo 2015; Van de Cruys 2014), used in a variety of linguistic tasks from semantic role classification to metaphor detecting (Haagsma and Bjerva 2016; Schulte im Walde et al. 2008; Zapirain et al. 2013), but no further refinements of the SPS itself. As I will show in Section 6.1.2 relative to my novel distributional measure of semantic selectivity (Cappelli and Lenci 2020), computing SA scores is a crucial step in the calculation of Computational PISA scores, despite it being intended as an improvement on Resnik's SPS rather than on SA itself.

While powerful in many respects, Resnik's model of selectional preferences has a crucial drawback, which is the need for a manually-built lexicon. This requirement makes it difficult to compute SPS scores for verbs in languages without a WordNet, for neologisms, and for special registers not yet encoded in WordNet. Given this severe limitation, I decided to improve on Resnik's SPS to model the recoverability of direct objects, creating a model based on distributional semantics (Lenci 2008, 2018) that can be applied in all the cases where the SPS measure cannot.

Results The SPS scores for the English and Italian transitive verbs under consideration are fully listed in Appendix C.1.

6.1.2 Computational PISA

Distributional semantics is based on the Distributional Hypothesis, which states that words occurring in the same contexts tend to have similar meanings, or, quoting the popular formula from Firth (1957), that "you shall know a word by the company it keeps". Such a framework puts the meaning of words in the ever-changing use speakers make of language, instead of securing it in the nodes of a static lexical taxonomy.

Most importantly, distributional semantics posits a correlation between distributional similarity and semantic similarity, and uses the former to model the latter. Roughly, the semantic space where words "keep company" can be modeled as a vector space where each word is a vector whose dimensions are context words. An example is provided in Figure 6.3, where the two words *hamburger* and *dragon* populate a semantic space defined by the three context words *big*, *tasty*, and *mythological*. The coordinates of *hamburger* are (2,2,0) because in this hypothetical situation it occurs twice with *big*, twice with *tasty*, and never with *mythological*, while the coordinates of *dragon* are (2,0,1) because it occurs twice with *big*, never with *tasty*, and once with *mythological* in the hypothetical corpus.

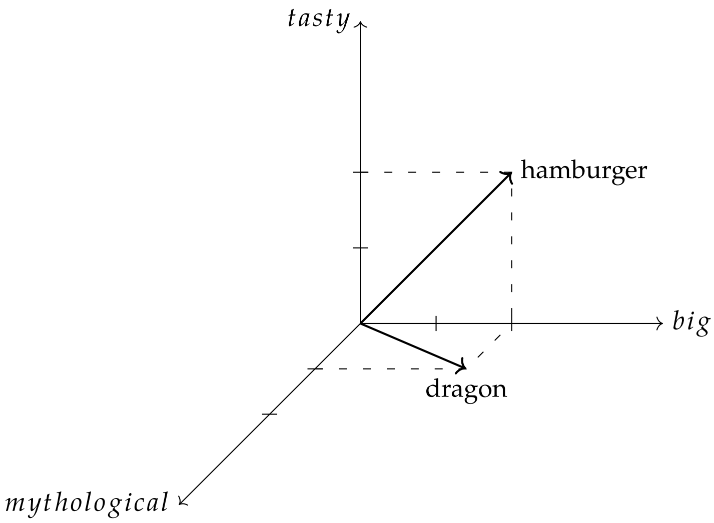

Figure 6.3: Simplified representation of the words *hamburger* and *dragon* in a vector space.

How does all of this translate into a solution to the problem of having a taxonomy-free model of argument recoverability? Let us go back to the example in (1) and consider the distribution of the arguments of the two verbs *to eat* and *to make* in a corpus. Ideally, collapsing on a 2-dimensional grid the n-dimensional space populated by these arguments, the distribution of the arguments of *to eat* would resemble Figure 6.4 and the distribution of the arguments of *to make* would resemble Figure 6.5. The (recoverable) arguments of *to eat* are close together in the vector space, while the (non-recoverable) arguments of *to make* are very sparse in the same space.

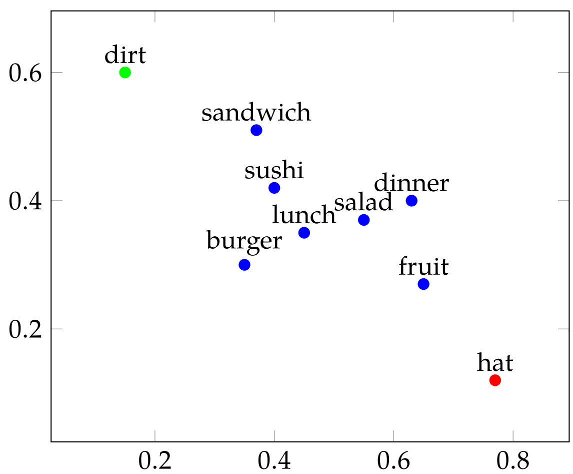

Figure 6.4: Made-up representation of the possible distribution of the direct objects of the verb *to eat* in a vector space.

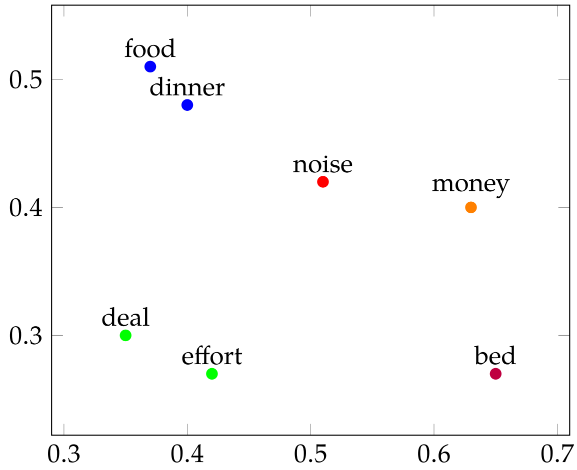

Figure 6.5: Made-up representation of the possible distribution of the direct objects of the verb *to make* in a vector space.

Based on these consideration, the conclusion is that the closer the direct objects of a verb are in a vector space, the more recoverable they should be. This is the main intuition behind Computational PISA, whose computational implementation I will discuss in the next paragraph.

Computational implementation Computational PISA is defined as the semantic density of a verb-relation pair, i.e., the mean value of the pairwise cosine similarity between the arguments of the pair. In this thesis, this means calculating the mean pairwise cosine similarity between all the direct objects of a given transitive verb. The full script used to obtain Computational PISA scores using a corpus and a list of verbs as input is available on GitHub[1] .

1: https://github.com/ellepannitto/PISA

This is done in two steps, for each transitive verb under consideration. The first one is to compute the Selectional Association between the verb and each one of its direct objects as defined by Erk (2007) and Erk, S. Padó, and U. Padó (2010) in Equation 6.9. This is a measure of the strength of the selectional preference SA of a verb for a possible argument a_0, modeled as the weighted sum of the similarities between the candidate argument a_0 and the actual arguments found in the corpus (each called a in the formula). In Cappelli and Lenci (2020) and this dissertation, I measure argument similarity with the cosine similarity.

$$SA_{v,r}(a_0) = \sum_{a \in args(v,r)} wt_{v,r}(a)\, sim(a_0, a) \tag{6.9}$$

Then, Computational PISA scores are computed as in Equation 6.10, i.e., by averaging Equation 6.9 over the n direct objects of a given transitive verb.

$$PISA_{v,r} = \frac{1}{n} \sum_{i=1}^{n} SA_{v,r}(a_i) \tag{6.10}$$

In Cappelli and Lenci (2020) five weighting functions are used to compute Equation 6.9 and then Equation 6.10 (Equation 6.11, Equation 6.12, and Equation 6.13 are taken from Erk, S. Padó, and U. Padó (2010)). In detail, the functions are as follows.

- **UNI** assumes a uniform distribution. This actually yields an *un*weighted model, because in Equation 6.9 the argument similarity is always multiplied by 1.

$$wt_{v,r}(a) = 1 \tag{6.11}$$

- **FRQ** is the co-occurrence frequency of a given direct object with the transitive verb under consideration.

$$wt_{v,r}(a) = \text{freq}(a, v, r) \tag{6.12}$$

- **IDF** is inspired to the well-known Inverse Document Frequency weighting scheme by Spärck Jones (1973), which assigns higher scores to direct objects occurring with fewer transitive verbs (the minimum would be 0, for an argument that occurs with every verb in the corpus). This is done in order to mitigate the frequency effect of arguments occurring with too many verbs to be considered relevant for the specific target verb under examination. In Equation 6.13, $|v, r|$ is the number of transitive verbs in the corpus, and $|v, r : a \in v, r|$ is the number of transitive verbs having a as a direct object.

$$wt_{v,r}(a) = \log_2 \frac{|v, r|}{|v, r : a \in v, r|} \tag{6.13}$$

- **LMI** is the Local Mutual Information (Evert 2005, p. 89) of a direct object and a given transitive verb, computed as in Equation 6.14. The LMI compares the probability of a noun occurring in a corpus as the direct object of a verb with the probability of the noun and the transitive verb occurring in a corpus without any relation to one another. In other words, given a noun and a transitive verb, their LMI is computed as the logarithmic ratio of their joint probability and the product of their individual probabilities in the corpus (i.e., their joint probability if they were statistically independent events).

$$wt_{v,r}(a) = f(a, v, r) \log_2 \frac{p(a, v, r)}{p(a)p(v, r)} \tag{6.14}$$

- **ENT** is the entropy (Shannon 1948) of the direct objects of a given transitive verb. Information theory uses entropy to quantify the informativity of a given event, inheriting its mathematical definition from thermodynamics. The entropy of an event (which in Equation 6.15 is the direct object itself) is a function whose value decreases as the probability of the event increases.

$$wt_{v,r}(a) = - \sum_{a \in args(v,r)} p(a) \log_2 p(a) \tag{6.15}$$

  In Equation 6.15, $p(a) = \frac{f(a)}{\sum_{a_0 \in A} f(a_0)}$, where A is the complete set of the direct objects of the target verbs, extracted from the corpus.

The original experiment In Cappelli and Lenci (2020), I computed both weighted models of argument recoverability (as explained in the previous paragraph) and unweighted models, only taking into account 300 direct objects for each transitive verb. For each verb, the relevant 300 objects were selected after sorting the entire list of direct objects based on the FRQ, IDF, LMI and ENT functions. The reason to include unweighted models in the experiment stems from the observation that the computation of Equation 6.10 for verbs with a large number of direct objects can get cumbersome, and it may be possible to achieve a comparable degree of informativity by only considering the most relevant objects occurring with each verb.

I tested the models on a 99-verb set of transitive verbs, extracting their direct objects from ukWaC, a 2-billion token part-of-speech tagged and lemmatized corpus of English (Ferraresi et al. 2008). Direct objects were modeled as bare head nouns, excluding any determiner and modifier present in the DP (e.g., *sword* instead of *a big rusty sword*). I obtained the vector representation of direct objects by using 12 different 300-dimensional embeddings trained on a concatenation of ukWaC and a 2018-dump of English Wikipedia, including both SVD-reduced count-based DSMs and neural embeddings.

Since Computational PISA was intended to model argument recoverability, I tested the results of each model by means of a Mann-Whitney U test comparing the mean Computational PISA score of the recoverable-object transitive verbs with the mean Computational PISA score of the non-recoverable-object transitive verbs. Summing up the discussion of the results carried out in the original paper (here in Table 6.1), it was found that the weighted versions of Computational PISA yield highly significant results, while the unweighted versions yield results with a comparable degree of significance just with the FRQ sorting function and, in particular, running the model on word2vec (Mikolov et al. 2013) distributional spaces.

Operative choices Drawing from the results discussed in Cappelli and Lenci (2020) and summarized in the previous paragraph, I computed Computational PISA scores for my verbs of interest in English and Italian. For English, I based the calculations on ukWaC as in the original study, and for Italian, I based them on itWaC, a 2-billion token part-of-speech tagged and lemmatized corpus of Italian (Baroni et al. 2009). Given that web-scraped, automatically tagged corpora this large inevitably suffer from significant noise in the data, which then results in possibly unreliable results, I pre-processed the data extracted from both corpora to minimize the impact of noise and tagging mishaps. First of all, I filtered

Table 6.1: Mann-Whitney U tests comparing recoverable- and non-recoverable-argument verbs (significance levels).

		weighted	**top** k	**bot** k
UNI	SVD	***	-	-
	w2v	***	-	-
	w2vf	**	-	-
FRQ	SVD	***	**	ns
	w2v	***	***	ns
	w2vf	***	**	ns
IDF	SVD	***	**	ns
	w2v	***	***	***
	w2vf	**	ns	ns
LMI	SVD	***	**	ns
	w2v	***	*	*
	w2vf	***	*	*
ENT	SVD	***	ns	ns
	w2v	***	***	ns
	w2vf	***	*	*

the 65,000 direct objects of my 30+30 target verbs so that they were:

- not hapaxes (considering the verb-noun frequencies);
- having an absolute frequency in the corpus greater than 100;
- present in WordNet (to eliminate misspelled words).

Then, I manually filtered the remaining nouns so that each verb only takes direct objects which do not belong to any of these categories:

- idiomatic senses (e.g., 'ice' as an object of 'to break');
- metaphoric senses (e.g., 'mind' as an object of 'to poison');
- direct objects of an unintended meaning of the verb (e.g., 'salmon' as an object of 'to smoke', since the intended sense of 'to smoke' in my study is only the one related to inhaling the byproduct of the combustion of tobacco and other plants);
- unintended direct objects (e.g., Recipients in double-object constructions such as 'pupils' in 'to teach pupils Linguistics', since in my study I am only interested in Theme/Patient direct objects);
- mistagged direct objects (e.g., 'disorder' as an object of 'to eat', which is clearly the result of the automatic corpus tagger interpreting 'eating' in 'eating disorder' as a verb rather than as an adjective);
- 'thing' (and 'stuff'), which appears with every transitive verb and is thus irrelevant.

The files containing the complete list of direct objects of each verb, both raw and cleaned, both in English and in Italian, are available for consultation on my GitHub profile[2] .
I based my operative choices regarding the vector spaces and weighting functions on the results I obtained in the original Computational PISA study. In particular, I computed word2vec neural embeddings with the Python library Gensim with the same parameters for English and Italian, i.e., Skipgram with Negative Sampling (SGNS) with 5 noise words, window of 10 words, 300-dimension vectors, ignoring all words with an absolute frequency lower than 10. Among the five functions defined in Cappelli and Lenci (2020), I employed the FRQ weighting function because it is the best-performing among all five, both in weighted and in sorted models of Computational PISA.

2: https://github.com/giuliacappelli/dissertationData

Results The Computational PISA scores for the English and Italian transitive verbs under consideration are fully listed in Appendix C.2.

6.1.3 Behavioral PISA

Introduction Despite representing a clear step forward compared to Resnik's taxonomy-based measure of selectional preferences, a corpus-based measure such as Computational PISA still has notable downsides. First of all, the corpus on which Computational PISA calculations are based has to be large enough to feature all the verbs under consideration, and for each verb to present a sufficient number of direct objects to allow for the computation of meaningful Computational PISA scores. It may well be the case that using a small corpus forces an experimenter to draw from a different set of verbs than intended for their study, if not all of them are featured in the corpus. Moreover, low-frequency verbs may

appear with a small set of direct objects even in large corpora, making it potentially difficult to obtain reliable Computational PISA scores. Another order of problems (which larger, automatically-tagged corpora are more prone to suffer from than smaller, manually-tagged ones) depends on noisy or otherwise inaccurate corpus data. For instance, large corpora constructed by crawling the Web (such as ukWaC and itWaC) typically present a substantial number of misspelled words and mis-tagged parts of speech, making it necessary to clean them manually beforehand. In addition to this, even clean corpora often lack fine-grained semantic information about thematic roles (e.g., distinguishing the Theme and the Recipient in *John gave Mary a book*), polysemy (e.g., *to graduate* may mean 'to become a doctor', but also 'to arrange something in gradations'), and idiomatic uses of verbs (e.g., there is usually no actual bucket being kicked when someone *kicks the bucket*). It is also important to point out that the static word embeddings used to obtain Computational PISA scores are unable to discriminate between different senses of a word, unlike dynamic, contextual embeddings such as BERT (Devlin et al. 2018), where a word gets represented by different vectors in different contexts.
In order to overcome these problems and obtain equally reliable scores for each verb in my experiment, I decided to implement a behavioral variant of the original Computational PISA measure. In the case of Behavioral PISA, the semantic similarity of a given verb's direct objects is not approximated via their distributional similarity in a corpus, but instead via their psychological similarity as judged by native speakers of the languages under study. Such a measure is intended to provide robust data for each verb regardless of its corpus frequency or scarcity of direct objects with respect to other verbs, at the cost of having to perform a behavioral experiment with human subjects.

Experimental protocol The experimental procedure to obtain the relevant data and compute the Behavioral PISA scores follows closely the method Medina used in her thesis to compute a comparable measure, which she calls "Object Similarity" (Medina 2007, pp. 173–178), and whose purpose is to overcome the shortcomings of Resnik's SPS. In a sense, Object Similarity may be viewed as a behavioral precursor of Computational PISA, both being based on a broad notion of selectivity-as-semantic-closeness.
In order to build the stimuli, I picked 6 pairs of direct objects for each verb of interest in my two 30-verb sets (one for English and one for Italian, as detailed in Chapter 7). For each verb, the direct objects comprising the 6 pairs were randomly selected from the manually cleaned (see Section 6.1.2) verb-object lists, so that each pair contained two different direct objects. This operation resulted in 180 stimuli (30 verbs x 6 direct objects) for each language, which the interested reader will find in Appendix B. Each pair of direct objects works as a stimulus, without explicit mention of the verb which subcategorizes the objects in the pair.
The two lists of 180 stimuli were used to create two Google Form surveys (with randomized stimuli), and 25 unpaid raters were recruited online for each language among native speakers holding at least a Bachelor's degree. Each rater had to judge the similarity of the two objects in each pair on a 7-point Likert scale in a single experimental session. Crucially, the participants to the experiment were not provided with a strict definition of similarity, and they did not know that the stimuli were direct objects of

a given set of verbs. Instead, they were told that the pair 'love - upholstery' should get a rating of 1 an the pair 'cat - dog' should get a rating of 7, to familiarize them with the Likert scale, and they were encouraged to make use of the whole 7-point scale whenever necessary.

Computational implementation Behavioral PISA is defined as the mean pairwise similarity between a subset of direct objects of a transitive verb. The pairwise similarity was obtained via human similarity judgments on a 7-point Likert scale, as described right above. In order to account for individual differences in the use of the scale, I computed the within-subject z-scores of these results, and then I averaged the normalized scores to obtain a single value for each target verb v (Kim, Rawlins, and Smolensky 2018, 2019; Kim, Rawlins, Van Durme, et al. 2019) as in Equation 6.16, where r is a (normalized) rating and i is the total number of ratings.

$$PISA_v = \frac{\sum_i r_v}{i} \tag{6.16}$$

The Behavioral PISA scores thus obtained were then normalized to fall between 0 and 1. The Python script I coded to generate the stimuli and compute the Behavioral PISA scores for each verb is freely available on my GitHub profile[3] .

3: https://github.com/giuliacappelli/behavioralPISA

Results The Behavioral PISA scores for the English and Italian transitive verbs under consideration are fully listed in Appendix C.3.

6.1.4 Model evaluation

The three models of semantic selectivity as a proxy to the recoverability of the direct objects of a verb discussed in this section (i.e., Resnik's SPS, Computational PISA, and Behavioral PISA) all capture different aspects of the same phenomenon and leverage different computational implementations. Resnik's SPS is a taxonomy-based measure which looks at the WordNet classes direct objects belong to, in order to compute the selectional strength of a transitive verb. Computational PISA is a corpus-based measure whose implementation is rooted in distributional semantics. Behavioral PISA models the recoverability of the direct objects of a given transitive verb via acceptability judgments elicited from native speakers relative to the similarity of several pairs of those direct objects. Using human judgments on object similarity as a baseline to evaluate the performance of computational models, in this Section I will consider Behavioral PISA the benchmark against which Resnik's SPS and Computational PISA are to be compared. Based on the correlation matrix in Table 6.2, relative to the three models of semantic selectivity in English, it appears that Computational PISA is a significantly reliable model (Pearson r 0.685, $p < 0.001$), while Resnik's SPS does not make the cut (Pearson r 0.286, non-significant).

Similarly, based on the correlation matrix in Table 6.3, relative to the three models of semantic selectivity in Italian, it results that Computational PISA in Italian is just as good a model of semantic selectivity as it

Table 6.2: Correlation matrix for the three models of semantic selectivity in English. Significance levels are star-marked in the table as follows:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

	SPS	**Comp. PISA**	**Behav. PISA**
SPS		0.330 .	0.286
Comp. PISA	0.330 .		0.685 ***
Behav. PISA	0.286	0.685 ***	

is in English, with a Pearson r of 0.687, significant at the 0.001 level. Interestingly, Resnik's SPS is also somewhat a good model if compared against human judgments (Pearson r 0.522, $p < 0.05$), but way less so than Computational PISA.

Table 6.3: Correlation matrix for the three models of semantic selectivity in Italian. Significance levels are star-marked in the table as follows:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

	SPS	**Comp. PISA**	**Behav. PISA**
SPS		0.746 ***	0.522 *
Comp. PISA	0.746 ***		0.687 ***
Behav. PISA	0.522 *	0.687 ***	

Thus, Computational PISA compares favorably to Behavioral PISA both in English and in Italian, making it a sensible choice for a distributional model of a verb's semantic selectivity, especially considering that both PISAs are based on the pairwise similarity of direct objects.
Incidentally, Table 6.2 and Table 6.3 also show that there is a much higher correlation between Computational PISA and Resnik's SPS in Italian (Pearson r 0.746, $p < 0.001$) than in English (Pearson r 0.330, not quite significant). These results suggest that, while being methodologically closer to Behavioral PISA as a pairwise similarity-based model, Computational PISA in Italian shares with Resnik's SPS something even more relevant. Given that Computational PISA, unlike Resnik's SPS, does not make use of WordNet or other taxonomies, it stands to reason that the common factor between these two measures of semantic selectivity in Italian has to be itWaC, i.e., the corpus used in both their computational implementations. More light on the difference between ukWaC and itWaC on the linguistic models of object drop they yield will be shed in Chapter 8 and Chapter 9.

6.2 Telicity

6.2.1 Telicity tests

Section 3.2.1 introduced telicity as a major predictor of object omissibility. Since aspectual interpretation is compositional (Olsen 1997 [2014], p. 14), a transitive verb which can be used both transitively and intransitively tends to get a telic interpretation in the first case and an atelic interpretation in the second case (refer to Section 5.2.2 for more details).
In my behavioral experiments and subsequent linguistic models of object drop, I operationalize telicity as a binary feature, based on theoretical claims discussed in Chapter 3 and Chapter 5. A verb is considered [+telic] if two tests yield a telic interpretation, [-telic] otherwise. Out of Medina's set of tests, I only used the staple *in/for* test for telicity, rejecting the *almost* test since it is notoriously quite problematic with achievements (please refer to Bertinetto and Delfitto (2000) for an extensive discussion on the issue), and the *counting* test because speakers seldom agree on its results. Verbs were tested without a direct object to avoid involuntarily eliciting a telic interpretation, or with a generic object (e.g., "something") if the

intransitive use yields a grammatically unacceptable interpretation. The two tests used for the experiments of this thesis are as follows (the reader is referred to Borik (2006) and L. Liu (2014) for a detailed review of possible telicity tests)[4] .

4: I also considered the *conjunction* test for telicity, but then decided to discard it because it is weaker than the other two. This diagnostic requires that the predicate be tested in a sentence where it is modified by two conjoined temporal adjuncts denoting consecutive time slots, as in (2). Telic verbs in this construction, as in (2-a), imply two separate events (i.e., John built something on Saturday and something else on Sunday). On the contrary, sentences with atelic verbs, as (2-b), can be interpreted either as two separate events (i.e., John sang on Saturday, stopped, and resumed his singing on Sunday) or as a continuous event (John sang on both days without interruption).

(2) a. John built something on Saturday and on Sunday.
b. John sang on Saturday and on Sunday.

***In*/*for* test** Based on this largely agreed-upon diagnostic for telicity, telic predicates (e.g., *to build* in (3)) are only grammatical if used with time-frame adverbials such as "in X time", while atelic predicates (e.g., *to sing* in (4)) are grammatical if used with time-span adverbials such as "for X time".

(3) a. John built something in an hour.
b. %John built something for an hour.

(4) a. *John sang in an hour.
b. John sang for an hour.

***Progressive* entailment test** The progressive form of a predicate bears different pragmatic implicatures based on whether it is telic or atelic. In particular, saying that the subject "was *verb*ing" implies that they *verb*ed with atelic verbs as in (5-b), while this is not true for telic verbs as in (5-a).

(5) a. John was building something.
b. John was singing.

6.2.2 Results

The telicity features of each English and Italian transitive verb are listed in Appendix C.4.

6.3 Perfectivity

Perfectivity was argued to be as a possible determinant of object drop in Section 3.2.2. The main point is that transitive verbs resist the omission of their direct object when used in the perfective aspect (as in (6-a)), while verbs in the imperfective aspect are much more likely to allow for their object to be dropped (as in (6-b)).

(6) a. ?John painted.
b. John was painting.

Like telicity, perfectivity appears as a binary feature in my experiments and models. Unlike telicity, perfectivity is not an inherent feature of verbs, so it cannot be tested beforehand: instead, I will include it in the behavioral experiments by having both perfective and imperfective uses of the same verb (please refer to Chapter 7 for a complete overview of the experimental design). The reader will find the results of these operative choices in Appendix D.

6.3.1 A note on tense

To choose the correct tenses to encode (im)perfectivity in English and in Italian, a critical discussion on the relation between tense and aspect[5] is in order. A common perspective on grammatical aspect in English, as summarized by C. S. Smith (1991, p. 106) and Wagner (2001, p. 663), is that perfective aspect corresponds to the simple form of the verb, while imperfective aspect is derived via the progressive construction (i.e., by adding the auxiliary *be* and *-ing* to the main verb). This perspective has the undeniable merit of keeping grammatical aspect and lexical aspect apart, acknowledging the orthogonality of (a)telicity and (im)perfectivity (refer to Bertinetto (2001) for an extensive discussion on the issue). However, the view that the simple form of a verb is necessarily interpreted as perfective has long since been abandoned. Indeed, based on Bertinetto (2001) and Olsen (1997 [2014]), the simple past in English is not marked for the perspective aspect. On the contrary, simple verb forms are aspectually unmarked, and they get different aspectual interpretations based on the interaction of tense, lexical aspect, and the different context they appear in. For this reason, for the experimental stimuli I will choose tenses that are aspectually marked, as to avoid misinterpretations.

5: Please refer to Section 3.2.3 for broader considerations on the way telicity, perfectivity, and tense interact with one another.

However, this observation about the absence of a one-to-one correspondence between past tense and perfective aspect should not be taken to mean that there is absolutely no link between the two. Even though simple past does not equate to perfective aspect, it is nevertheless true that past tense at least *hints* towards perfectivity, as noted by Comrie (1976) and later on by Medina (2007), Olsen (1997 [2014]), and Wagner (2001), among others. So, a morphologically unmarked past form is more likely to be interpreted as perfective than a non-past form. Moreover, there are differences internal to the group of past forms. In fact, as we saw before, the simple past is aspectually unmarked, but the past perfect (e.g., *John had written a book.)* is marked as perfective.

6.3.2 Operative choices

Considering what has been noted in Section 6.3.1, and following Medina (2007), in the experimental stimuli perfective aspect will be encoded with perfect morphology, and imperfective aspect will be encoded with progressive morphology.

For English, perfective stimuli will be in the past perfect tense (as in (7-a)) and imperfective stimuli in the past continuous tense (as in (7-b)). For Italian, perfective stimuli will be in the *trapassato prossimo* tense (as in (8-a)) and imperfective stimuli in the continuous past form created with the auxiliary 'to stay' in the *imperfetto* tense and the simple gerund of the main verb (as in (8-b)).

(7) a. John had played.
    b. John was playing.

(8) a. Gianni aveva cantato.
    b. Gianni stava cantando.

On a side note, Italian also has perfective-only and imperfective-only simple past tenses, namely *passato remoto* and *imperfetto*, but I chose to use compound tenses since *passato remoto* lost some ground to *passato prossimo* over the last few decades, and also considering the variety of regional uses of *passato remoto* (Bertinetto and Squartini 1996). Most importantly, the *trapassato prossimo* tense was chosen to encode perfective aspect in Italian because it's the aspectually closest Italian tense to the English past perfect, despite the glaring differences between the aspectual systems in these two languages (Bertinetto 1992).

6.4 Iterativity

As argued in Section 3.3.2, iterativity and other types of pluractionality favor the omission of direct objects (compare (9-a) and (9-b)). Although well-known in the theoretical literature (Glass 2013, 2020; Ruda 2017), to my knowledge this predictor of object drop appears here for the first time in a comprehensive linguistic model based on acceptability judgments.

(9) a. #The Joker killed.
b. The Joker killed again.

Iterativity is yet another binary predictor and it will be encoded in the behavioral experiments in the same way as perfectivity, that is to say, by creating both iterative and non-iterative stimuli for the same verb as in (9) (please refer to Appendix D for the complete set).

6.5 Manner specification

The effects of manner specification on indefinite object drop were presented in Section 3.1.3. As argued by Ruda (2017) a.o., if a transitive verb allows for its direct object to be dropped, then its troponyms or near-synonyms with a specific manner component block it. This is evident in (10): the base verb *to eat* allows for object drop in (10-a), but its manner-specified counterparts *to devour/nibble/chew* in (10-b) do not.

(10) a. John ate.
b. *John devoured/nibbled/chewed.

In the following experiments, manner specification is treated as a binary predictor. The verbs used in the stimuli are marked in Appendix C.5 as either manner-specified or manner-unspecified. Just as in (10), verbs tagged as manner-specified are marked counterparts of manner unspecified verbs present in the list.
On Page 35 I observed, in passing, that manner-specified verbs tend to imply that the action described by the verb reached its natural endpoint, e.g., *to devour*, unlike *to eat*, strongly implies that the Patient is utterly consumed by the devourer. Based on the way I will implement telicity in my experiment (taken from Olsen (1997 [2014]), following Medina (2007)), then, one could wonder as to the relation between telicity and manner specification. In other words, if manner-specified verbs imply that the action got to its natural endpoint, is then manner specification

collinear or highly correlated with telicity? If so, one or the other should be excluded from the set of predictors used in the experiment. However, this is not the case, as shown in Appendix C. Indeed, while there is noticeable overlap between telicity and manner specification (with 12 atelic, manner-unspecified verbs and 8 telic, manner-specified verbs), there is also a number of cases where the relation does not hold (i.e., 6 atelic, manner-specified verbs and 4 telic, manner-unspecified verbs). Moreover, a sanity check I performed before running the linear mixed-effects models in Chapter 8 showed that telicity and manner specification are indeed not collinear with one another.

7 Collecting acceptability judgments: materials and methods

7.1 Operative choices

The development of both experiments is organized in three steps, i.e., building, running, and recruiting, with a different platform employed for each step. The merits of the PsychoPy-Pavlovia-Prolific pipeline, described below in full detail, are making it a growingly popular choice among behavioral experimenters. Let us examine each step separately.

7.1.1 Building the experiment locally

I built the experiment using the graphical interface (called "Builder") of PsychoPy v2020.2.10 (Peirce et al. 2019). PsychoPy is an open-source, cross-platform software package allowing experimenters to build any kind of experiment in psychology, neuroscience, psychophysics, linguistics, and other behavioral sciences. It makes it possible to code an experiment from scratch using the Coder interface (or any Python-friendly Integrated Development Environment), or to build one using the graphical interface provided by the Builder. To a skilled Python programmer, the main appeal of the Builder lies in that it has a feature to translate the built-in Python functions into JavaScript code that can run online, while Coder-created experiments can only be run locally on the experimenter's computer.

7.1.2 Running the experiment online

In order to be run online on the participants' devices, a PsychoPy Builder experiment first has to be uploaded on Pavlovia, a hosting platform for behavioral experiments coded using PsychoPy, lab.js, or jsPsych. Pavlovia can be fine-tuned to launch experiments online with a variety of options (e.g., launch pilots or full-fledged experiment, save or discard incomplete submissions), and it integrates seamlessly with popular participant recruiting platforms. Moreover, it doubles up as a source code repository thanks to integration with GitLab.
The source code for both my experiments, as well as the stimuli used in them, is available here for English[1] and Italian[2] . The files can be read directly online, but the code requires PsychoPy to be edited.

1: https://github.com/giuliacappelli/psychopy_exps/tree/main/eng

2: https://github.com/giuliacappelli/psychopy_exps/tree/main/ita

7.1.3 Recruiting participants

Both surveys were run on Prolific (formerly known as "Prolific Academic"), a large crowdsourcing platform that was specifically developed to cater to the needs of researchers. For a review of the merits of running behavioral experiments online and recruiting participants via crowdsourcing, please refer to Erlewine and Kotek (2016), Gibson, Piantadosi, and Fedorenko

(2011), and Grootswagers (2020).
The experimental tasks were carried out by 30 people for the study on Italian and 30 people for the study on English. All 60 people are native speakers of Italian and English, untrained in Linguistics, holding at least a Bachelor's degree (in order to minimize the effect of education on their judgments), and lacking any knowledge of the goals of this dissertation. The estimated duration for the experiments was about 30 minutes. Participants who completed the experiment were given £3.00 each as compensation for their effort, in compliance with Prolific's policy calling for ethical rewards.

7.2 Target verbs

Both for English and for Italian, the verb dataset used to build the stimuli comprises 30 target transitive verbs and 10 filler intransitive verbs. The reader will find both sets, together with the relevant frequency information, in Appendix A.
The optimal verb set used in the creation of the stimuli must be balanced in every relevant aspect. This means that:

- it has to include the same number of English and Italian verbs (leading to the same number of stimuli in both Likert[3] experiments);
- the verbs have to span over different frequency ranges within the same language (i.e., they cannot all be high-frequency or low-frequency verbs);
- each English verb has to have a corresponding Italian verb with roughly the same meaning and comparable (relative) frequency in the corpus.

3: Please refer to Langsford et al. (2018) and Weskott and Fanselow (2011) for an analysis of the reasons why judgments on a 7-point Likert scale are a better alternative to both binary judgments and judgments collected via a Magnitude Estimation (Bard, Robertson, and Sorace 1996) task.

This last requirement on verb frequency is crucial, since word frequency is typically confounded with the variables under consideration in (psycho)linguistic experiments[4] .
The creation of a verb set for the two Likert experiments was accomplished in two steps, as detailed below.

4: See, for instance, Arunachalam (2013) and Brysbaert, Mandera, and Keuleers (2018) on the word frequency effect and other confounding variables.

7.2.1 Creation of verb lists

First of all, a list was made of several transitive English verbs (3 from Resnik (1993, p. 138), 8 from Levin (1993, p. 33), 4 from both papers, and 5 commonly used verbs of my choosing). I further added to the list 10 verbs specified with respect to the manner component (discussed in Section 6.5), based on my personal opinion as a linguist and person in charge of this study. This resulted in a list containing 30 transitive verbs, reported in Table 7.1.
The English verbs in the list have unambiguous meanings (e.g., *to slice*, *to wash*). I did not include all of Resnik's original set, which featured highly polysemous verbs such as *to have* and *to do*, precisely to comply with this requirement. In order to verify that the 30 verbs are indeed monosemous, quantifying my intuition in a more rigorous way, I used WordNet (G. A. Miller 1995) to check that each verb belongs to just one synset (i.e., has just one sense) or, if it belongs to more than one as it often happens, that these synsets are very close together semantically.

In order to achieve this result, I computed the Wu-Palmer similarity[5] between all the WordNet synsets each verb belongs to. I defined strict criteria for monosemy, namely, a verb is only considered monosemous if no more than 20% of its senses have a Wu-Palmer similarity score lower than a set threshold. For lack of a standardized threshold in the literature, I set it at 0.15, since it is close to the scores of very different word senses of the notoriously non-monosemous English noun *bank*, a now classic textbook example (0.1428 for 'bank$_1$: sloping land' versus 'bank$_2$: financial institution', 0.1538 for 'bank$_1$: sloping land' versus 'bank$_6$: gambling house funds'). All the 30 verbs in the set are acceptable, based on these requirements. It is possible to reproduce the results (or apply the test to novel data) using my Python script, available here on GitHub[6] .

5: The Wu-Palmer similarity (WP) is a similarity measure based on the depth of the two synsets (s1 and s2) in the taxonomy and the depth of their closest ancestor, i.e., the Least Common Subsumer (LCS).

$$WP = 2 * \frac{depth(LCS(s1, s2))}{depth(s1) + depth(s2)}$$

It can vary between 0 and 1 ($0 < WP \leq 1$), with higher scores corresponding to more similar senses.

6: https://github.com/giuliacappelli/checkPolysemy

Keeping the semantic requirement for monosemy in mind, I created the set of 30 Italian verbs (also listed in Table 7.1) by translating each English verb from the list into Italian, and checked that they are not polysemous using the same criteria applied to the English verb set. For each verb in the English set, the corresponding Italian verb is the first translation found in the WordReference English-Italian Dictionary ©2020. I had to choose the second translation by WordReference just for *to chop* (because the first translation, *tagliare*, suits best the verb *to cut*) and for *to swig* (because the first translation, *tracannare*, features in the itWaC corpus only 32 times, while *trangugiare* is found 647 times).
Based on the information discussed in Chapter 6, I computed the semantic selectivity scores for each verb in both lists, and I annotated each verb pair with their telicity and manner specification features (the full details are collected in Appendix C).

7.2.2 Frequency check

The second step in the creation of the verb set is a "sanity check" of the verb frequencies, both within-language and between-language, as detailed in this paragraph. The (absolute) frequency of each verb was extracted from ukWaC for English and itWaC for Italian (Baroni et al. 2009).
Absolute frequencies have to be transformed in order to be compared, since they are corpus-dependent, and also because words occur in a corpus according to the power law known as "Zipf's law".[7] Computing the relative frequency, or the frequency per million words, would solve the first problem but not the second. Log-transforming either of these would solve both problems, but low-frequency words would yield negative logarithms, so that the scale would not be easily human-readable and it would be quite difficult to use for the purposes of this experiment.
In order to overcome these problems, I used the "Zipf scale" by Van Heuven et al. (2014), i.e., a logarithmic scale going from 1 (very-low-frequency words) to 7 (very-high-frequency words), much like a Likert scale. The human (or automatic) interpretation of values on the Zipf scale is straightforward, and it does not vary across corpora. The middle of the scale, i.e., 4, is the tipping point between low- and high-frequency words, and words with a Zipf score higher than 6 are very likely to be function words (i.e., non-content words, such as determiners and auxiliaries, often called "stop words" in computational literature).

7: According to Zipf's law, The rth most frequent word has a frequency $f(r)$ that scales according to:

$$f(r) \propto \frac{1}{r^{\alpha}}$$

for $\alpha \approx 1$. Based on this law, the distribution of word frequencies w.r.t. the word rank r is a logarithmic distribution (crucially, not a linear one).

Table 7.1: The sets of English and Italian verbs of interest.

English verbs	Italian verbs	source
behead	decapitare	new (manner of killing)
break	rompere	new
build	costruire	new
chop	spaccare	new (manner of cutting)
clean	pulire	Levin (1993)
cook	cucinare	Levin (1993)
cut	tagliare	new
devour	divorare	new (manner of eating)
doodle	scarabocchiare	new (manner of writing)
drink	bere	Levin (1993) and Resnik (1993)
eat	mangiare	Levin (1993) and Resnik (1993)
embroider	ricamare	Levin (1993)
hum	canticchiare	Levin (1993)
kill	uccidere	new
knife	accoltellare	new (manner of cutting)
poison	avvelenare	new (manner of killing)
polish	lucidare	Levin (1993)
pour	versare	Resnik (1993)
sew	cucire	Levin (1993)
sign	firmare	new (manner of writing)
sing	cantare	Levin (1993) and Resnik (1993)
sip	sorseggiare	new (manner of drinking)
slice	affettare	new (manner of cutting)
smoke	fumare	new
steal	rubare	Resnik (1993)
swig	trangugiare	new (manner of drinking)
teach	insegnare	Levin (1993)
wash	lavare	Levin (1993)
watch	guardare	Resnik (1993)
write	scrivere	Levin (1993) and Resnik (1993)

Knowing the absolute frequency of a verb in a corpus (f) and the corpus size in tokens (c), Zipf scores (Z) are easy to compute as in Equation 7.1 or, equivalently, as in Equation 7.2 (both are the base 10 logarithm of the frequency-per-billion-words).

$$Z = \log_{10} \frac{f * 1,000,000,000}{c} \tag{7.1}$$

$$Z = \log_{10} \frac{f}{c} + 9 \tag{7.2}$$

Having chosen a suitable scale to compare my verb frequencies, I performed the within- and between-language tests. As for the within-language check, I verified that the Zipf scores of the verbs for both languages were compatible with a distribution spanning from low- to high-frequency verbs, avoiding extremes (English verbs 2.078 ÷ 5.520, Italian verbs 1.681 ÷ 5.763). Lastly, I made sure that each English verb belonged in the same Zipf score tier as its Italian translation, by computing the difference between the English and the Italian Zipf scores. Since the Zipf scale is designed so that words in each of the 7 tiers have significantly different corpus frequencies, I take an English-Italian Zipf difference to be acceptable if smaller than 1, while I would reject any English-Italian verb pair with a difference in Zipf scores equal to or greater than 1. The within-and between-language frequency tests showed that the two verb sets comply with the above requirements. The reader will find the frequencies and Zipf scores in Appendix A.

7.3 Design

Both experiments follow a within-subject fully crossed design, where every participant sees all the stimuli in random order and provides judgments for each on a 7-point Likert scale.
The stimuli are organized in a 2x2x2 factorial design, summarized in Table 7.2, with 3 independent variables (presence of an overt direct object, perfectivity, iterativity) having 2 levels each (presence or absence of the feature). This experimental design is more complex than the 2x2 design by Medina (2007), since I am adding iterativity as an independent variable.

overt dObj	perfectivity	iterativity
+	+	+
+	+	-
+	-	+
+	-	-
-	+	+
-	+	-
-	-	+
-	-	-

Table 7.2: The 2x2x2 factorial design used in both Likert experiments.

Each verb of interest (fully listed in Appendix A) participates in each of the 8 experimental conditions. Since telicity, semantic recoverability, and manner specification are inherent properties of each verb, they are not part of the experimental design.

7.4 Stimuli

All the 30 target verbs plus 10 intransitive filler verbs participate in all the experimental conditions, leading to a total of 320 sentence stimuli (twice as in Medina (2007)) for each language in the study.
The list of English and Italian intransitive filler verbs is provided in Table 7.3. Since these verbs are part of the experimental design but irrelevant for the subsequent analysis, they were not controlled for frequency, nor were they annotated with semantic or aspectual information.

English fillers	Italian fillers
clap	applaudire
fast	digiunare
knock	bussare
laugh	ridere
limp	zoppicare
rest	riposarsi
scream	urlare
sleep	dormire
smile	sorridere
stagger	barcollare

Table 7.3: The sets of English and Italian filler verbs.

As mentioned in Chapter 6, semantic recoverability, telicity, and manner specification vary *across verbs*, while the presence of an overt direct object, perfectivity, and iterativity vary *across sentences*. This means that the two groups of object drop predictors are treated in different ways. Recoverability, telicity, and manner specification are intrinsic characteristics of each verb, respectively continuous the first and binary the last two (refer to Appendix C for the complete set of verbs with their features). On the contrary, the presence of a direct object, perfectivity, and iterativity are binary features that need to be encoded by creating a pair of minimally different sentences for each.
Let us consider the eight[8] example stimuli for the verb *to eat* in (1):

8: Based on the design in Section 7.3.

(1)	a.	John had eaten pizza again.	[dObj+, perf+, iter+]
	b.	John had eaten pizza.	[dObj+, perf+, iter-]
	c.	John was eating pizza again.	[dObj+, perf-, iter+]
	d.	John was eating pizza.	[dObj+, perf-, iter-]
	e.	John had eaten again.	[dObj-, perf+, iter+]
	f.	John had eaten.	[dObj-, perf+, iter-]
	g.	John was eating again.	[dObj-, perf-, iter+]
	h.	John was eating.	[dObj-, perf-, iter-]

All the 30 verbs of interest and the 10 filler verbs, both for English and Italian, will feature in stimuli like the ones listed in (1). In transitive sentences, regardless of the verb being transitive or intransitive in nature, the direct object is semantically compatible with the meaning of the verb so that the violation of selectional preferences does not act as a confound in the experiment. Moreover, in order to minimize the possible confounding effect of other factors which Hopper and Thompson (1980) identified as relevant in determining prototypical transitivity (as discussed in Section 2.1), all the stimuli are in the indicative (*realis*) mood, they are affirmative rather than negative, and they feature a human subject (which is the most volitional, high-in-potency Agent possible). As anticipated in Section 2.5, no context is provided in the stimuli, since it is known to enhance the recoverability of objects and, thus, to influence the grammaticality of indefinite object drop.
The reader will find the full list of stimuli and verb-object pairings in Appendix D, and detailed information on the manipulation of perfectivity and iterativity in Section 6.3.

7.5 Setting

7.5.1 Informed consent

First of all, the participants had to accept the privacy policy in order to proceed to the survey, or else decline the terms and exit the experiment. The English and the Italian versions of the privacy policy are reported below.

Privacy policy for the English survey

During this survey, you will not be asked personal information. In no way will it be possible for anyone to find out your identity by having access to your answers to the survey. The data you will provide will be used for the purposes of this linguistic experiment and they may be shared with third parties anonymously. By completing the survey, you accept these terms. If you leave the survey early, your data will not be used in the study and you will not be compensated.

Privacy policy for the Italian survey

Nel questionario non ti saranno chieste informazioni personali. Non sarà possibile per nessuno risalire alla tua identità a partire dalle risposte che darai nel questionario. I dati che fornirai saranno usati ai fini di questo esperimento linguistico e potranno essere condivisi con terzi in forma anonima. Completando il questionario, dichiari di accettare questi termini. Se abbandoni il questionario in anticipo, i tuoi dati non saranno usati nello studio e non riceverai alcun compenso.

7.5.2 Instructions

Then, participants were given instructions in their native language (i.e., the target language of the survey). The full text is available right below for both English and Italian.

Instructions for the English survey

This survey takes about 30 minutes to complete, and you will be rewarded £ 3.00 as compensation if you complete the survey.
You will see a series of sentences, one by one. For each, you are asked to judge how acceptable it is to you on a graded scale. You should rate a sentence 1 if it sounds utterly bad, 7 if it sounds perfectly fine, or choose any in-between score if you think it applies. Let us consider some examples:

- John laughs stories.

This sentence should score 1, because you can't "laugh something".

- Mario walked on the path.

This sentence should score 7, because it's perfectly acceptable.
This is not an exam! By virtue of being a native speaker of English, you will provide the right answers. Beware: to avoid cheating and random clicking, the survey is interspersed with hidden control questions. If you fail them, you will be kicked out of the survey and receive no compensation.

Instructions for the Italian survey

Il questionario richiede - minuti in media per essere completato e riceverai 3.00 £ come compenso se lo completerai tutto.
Vedrai una serie di frasi, una alla volta. Per ciascuna, devi giudicare quanto ti sembra accettabile in una scala di valori. Dovresti dare a

una frase il punteggio 1 se ti sembra del tutto sbagliata, 7 se ti suona del tutto corretta, o scegliere punteggi intermedi se ti sembra il caso. Guardiamo alcuni esempi:

- Gianni ride storie.

A questa frase dovresti dare 1, perché non si può "ridere qualcosa".

- Mario camminava sul sentiero.

A questa frase dovresti dare 7, perché è perfettamente accettabile. Questo non è un esame! Le tue risposte sono tutte giuste, perché sei un parlante nativo di italiano. Attenzione, però: per impedire che vengano date risposte a caso, il questionario contiene domande di controllo nascoste. Se le sbaglierai, ti sarà impedito di continuare a rispondere e non riceverai alcun compenso.

7.5.3 Screening survey

Finally, the participants were asked to complete a short screening survey before entering the actual linguistic judgment survey. The screening questions are presented in (2) for English and in (3) for Italian:

(2) a. Are you a native speaker of English?
b. Have you got a Bachelor's (or higher) degree?
c. Have you understood the instructions above?

(3) a. Sei un parlante nativo di italiano?
b. Hai una laurea triennale (o titolo superiore)?
c. Hai capito le istruzioni presentate sopra?

Participants could click on either "Yes" or "No" buttons to answer. Answering "No" to any screening questions meant being automatically kicked out of the survey.

7.5.4 Training session

Before entering the actual experimental session, participants had the chance to accustom themselves to the task in a short training session. They were asked to judge the acceptability of each sentence on a 7-point Likert scale, as in the full experimental session. Likert scales are a reliable method to test grammaticality (Weskott and Fanselow 2011), and the 7-point variant (unlike the 5-point scale used by Medina (2007)) is the most common in experimental linguistics (Juzek 2016).
In order to keep the training session as short as possible while maximizing its usefulness, subjects only judged a fully-grammatical sentence and a fully-ungrammatical sentence (in (4) for the English survey and (5) for the Italian survey).

(4) a. Jack had opened a bar.
b. *Ann went a sandwich.

(5) a. Sergio ha aperto un bar.
b. *Anna è andata un panino.

Unlike the real experimental session, the training session allowed participants to keep rating a sentence as many times as they wanted, instead of kicking them out in the case of a mistake. After each judgment on the Likert scale, participants received immediate feedback in the same interface window. They were either prompted to provide a different judgment on the same sentence, if the score they chose was off-scale, or they were asked to continue to the next task, if they rated the sentence correctly. Expected scores in the training session were quite strict, i.e., no less than 6 for (4-a) and (5-a), and no more than 2 for (4-b) and (5-b).

7.5.5 Experimental session

The 320 stimuli were presented in randomized order, since order is well-known to have an effect on acceptability judgments (Juzek 2016; Myers 2009). In general, randomizing stimuli or counterbalancing conditions are good experimental practice to counteract the effects of carryover, fatigue, and practice. Moreover, the stimuli were presented one by one, in order to prevent participants to compare the stimuli one to another instead of judging them individually. The concern for similar task-specific strategies and the need to avoid eliciting them is raised, among others, by Myers (2009).

Participants were instructed to click on their chosen score and then press their spacebar to proceed to the next stimulus, so they could change their mind before submitting their judgment for good. The experiment was coded as to prevent participants from skipping stimuli.

7.5.6 Reliability of judgments

The reliability of the collected judgments was ensured in two different ways. On a general note, paid compensation is standard practice in behavioral studies where attentiveness is key, as well as being a popular way to make up for the time participants invested[9] in the experiment.

A more task-specific method to elicit reliable judgments involves using the clearly ungrammatical and the clearly grammatical control sentences as monitoring stimuli. This means that ungrammatical control sentences like (6), where an intransitive filler verb appears with a direct object, should get a very low score on the 7-point Likert scale.

(6) * John had slept pillows.

Likewise, grammatical control sentences like (7-a), where a transitive verb appears with a semantically compatible direct object, or one like (7-b), where an intransitive verb is used intransitively, should get a very high score on the 7-point Likert scale.

(7) a. John was eating pizza.
    b. John was limping.

Summing up, based on the experimental design previously depicted in Table 7.2, the stimuli are divided into target stimuli and (un)grammatical control stimuli as in Table 7.4.

9: Please refer to Permuth-Wey and Borenstein (2009) for an extensive debate on the ethical and practical implications of financial remuneration in behavioral research.

Table 7.4: Summary of which stimuli are targets and which ones are (un)grammatical controls in both experiments.

overt dObj	perfectivity	iterativity	trans. verbs	intrans. verbs
+	+	+	/	*control
+	+	-	control	*control
+	-	+	/	*control
+	-	-	control	*control
-	+	+	target	/
-	+	-	target	control
-	-	+	target	/
-	-	-	target	control

Since getting just one judgment wrong out of 320 would cost a participant his reward (and, incidentally, cost this study valuable data), the requirements for a judgment to be deemed correct were softened with respect to those used in the training session. Thus, participants who provided a score higher than 3 (i.e., not in the lower half of the scale) to filler sentences or lower than 5 (i.e., not in the higher half of the scale) to control sentences were kicked out of the experiment and did not receive any compensation, since out-of-range scores would mean that they were not paying enough attention to the task or were downright clicking at random.

Notably, iterative sentences with transitive verbs used transitively and intransitive verbs used intransitively were not included among the control stimuli, because they are not prototypical examples of perfectly grammatical sentences. It would have been frustrating for participants to be excluded from the experiment (and the reward) for a single mistake on such a sentence, especially considering that half the stimuli already were control sentences.

The results of the English and Italian experiments are presented in full detail in Chapter 8. The full English and Italian stimuli, together with the raw scores provided by the participants on a 7-point Likert scale for the target sentences, are also available on my GitHub profile in a dedicated repository[10] .

10: https://github.com/giuliacappelli/dissertationData

Exploring the acceptability judgments | 8

8.1 Making sense of the results: computational implementation

8.1.1 Operative pipeline

This paragraph outlines the technical steps needed to replicate my results by running my scripts on the raw judgment data I collected. As pointed out before, all the raw input data necessary to run my scripts, which in this case are the Likert judgments provided by human participants to the experiment in Chapter 7, are available in a dedicated GitHub repository*. The first step in this pipeline is the cleansing and reshaping of the output generated in the PsychoPy-Pavlovia-Prolific process of judgment gathering detailed in Section 7.1. This result can be achieved using my dedicated script† on GitHub, which takes care of taking the full Pavlovia-generated file as input and yielding a tabular output with the minimal information necessary to run my Stochastic Optimality Theoretic analysis. The script also anonymizes the participants' names to make the data shareable, and it can (optionally, depending on the experimenter's needs) filter out any participant providing polar either-1-or-7 judgments when prompted to make full use of the 7-point Likert scale.
The judgments are now ready to be processed with the main Python program‡, which preprocesses the judgments as described in Section 8.1.2, generates the data used in the analysis provided in this Chapter, and models the judgments according to the Stochastic Optimality Theory requirements described in Chapter 4 (final results in Chapter 9).

8.1.2 Data preprocessing

Before moving forward to the actual data analysis and modeling, the main script carries out three preprocessing steps:

1. computing the min-max normalized semantic selectivity values for Resnik's SPS, Computational PISA, and Behavioral PISA input files (first introduced in Section 6.1), to make the results comparable across models;
2. multiplying the semantic selectivity score of each verb by its Zipf value (first introduced in Section 7.2), i.e., the base 10 logarithm of the frequency-per-billion-words of the verb in a given corpus, to avoid having the verb's frequency confound the information provided by the semantic selectivity models;

* https://github.com/giuliacappelli/dissertationData
† https://github.com/giuliacappelli/PsychopyToMedina
‡ https://github.com/giuliacappelli/MedinaStochasticOptimalityTheory

3. computing the within-subject z-scores for the judgments, then averaging these scores to obtain the mean judgment for each sentence in the stimuli list, then normalizing the mean judgments between 0 and 1 (following the technique by Kim, Rawlins, and Smolensky (2018, 2019) and Kim, Rawlins, Van Durme, et al. (2019)), to account for inevitable differences in the way each participant makes use of the Likert scale.

This kind of preprocessing also improves on Medina's (2007) setting, where both semantic selectivity and judgment data were analysed by considering their raw, original values, because it minimizes the potentially disruptive influence of external factors such as corpus frequencies and individual differences in humans on the final Stochastic Optimality Theoretic model of object drop.
In the next sections, following Medina (2007), I will provide a thorough description of the way the acceptability judgments about the implicit object construction are influenced by each factor separately and by all the factors together, both in English and in Italian. This analysis will show that:

- no factor alone has a main effect so strong as to fully predict the grammaticality of the implicit object construction;
- a comprehensive Stochastic Optimality Theoretic model of object drop based on all five predictors in Chapter 6 is indeed feasible.

The model itself is presented and discussed in Chapter 9.

8.2 English results

8.2.1 Semantic selectivity

The effect of semantic selectivity on the acceptability of the implicit object in English is quantified by means of a Pearson correlation between them. The results of this computation are visualized in Figure 8.1 for Resnik's SPS, in Figure 8.2 for Computational PISA, and in Figure 8.3 for Behavioral PISA.
The first thing to strike the eye of the observer is that the three models of semantic selectivity correlate with varying degrees of accuracy with the human judgments. The Selectional Preference Strength, a now-classic measure by Resnik (1993, 1996), yields unsatisfactory results which fall quite short of statistical significance. Computational PISA performs much better, with significant ($p = 0.038$) results, even though the correlation between it and the judgments is very modest (Pearson's $r = 0.381$). Finally, Behavioral PISA appears to be by far the best-performing model of semantic selectivity, with a Pearson's r of 0.494 against human judgments and a p value of 0.006.
Keeping in mind what I concluded about the three models of semantic selectivity back in Section 6.1.4 (see Table 6.2 in particular), these results should not come as a surprise. Indeed, being based on human judgments, Behavioral PISA is expected to yield the best results among the three models used here. Computational PISA correlated quite well with the Behavioral PISA benchmark, and we can see that it also correlates nicely with the acceptability judgments regarding the implicit object

construction. Resnik's SPS, on the contrary, was found to be a poor model of semantic selectivity if compared to Behavioral PISA, and it is also a poor fit if compared to human ratings about object drop. The good performance of Computational PISA and Behavioral PISA against Resnik's SPS can be explained by referring to the way these models were created, since both PISA models are based on pairwise similarity scores between pairs of direct objects for a given verb, while Resnik's SPS is taxonomy-based.

All in all, semantic selectivity (especially the PISA models) is not a bad predictor of object drop in English, but it's far from being a reliable one when considered in isolation from all the other possible factors.

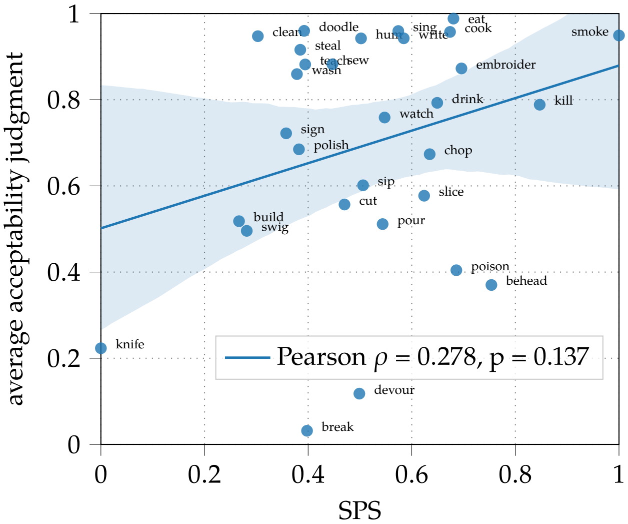

Figure 8.1: Correlation between semantic selectivity (Resnik's SPS) and normalized acceptability judgments on object drop in English.

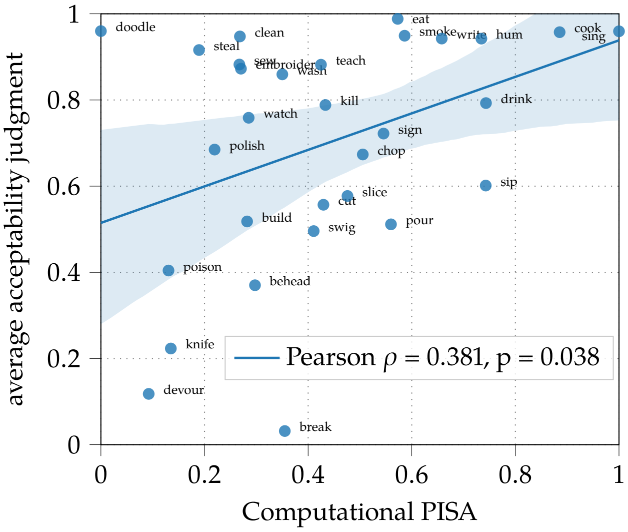

Figure 8.2: Correlation between semantic selectivity (Computational PISA) and normalized acceptability judgments on object drop in English.

8.2.2 Binary predictors

Telicity The boxplots in Figure 8.4 illustrate the main effect of telicity on the acceptability judgments on the implicit object construction in English. A Mann-Whitney U test reveals that telic verbs were judged as significantly ($p < 0.0001$) less grammatical than atelic verbs, consistently with expectations (refer back to Section 3.2.1 and Section 6.2). In particular,

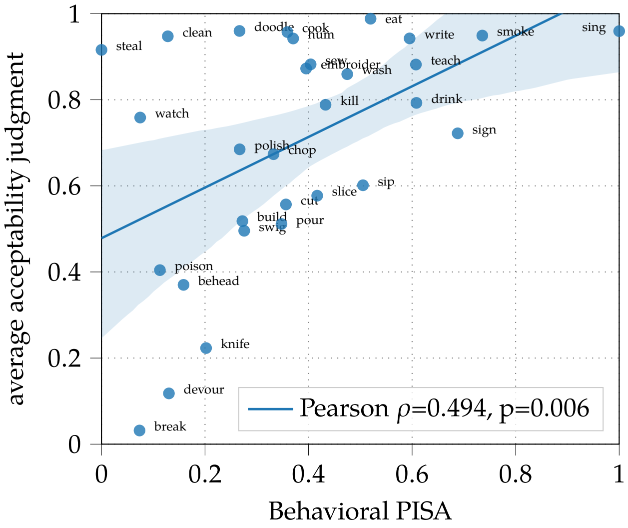

Figure 8.3: Correlation between semantic selectivity (Behavioral PISA) and normalized acceptability judgments on object drop in English.

the median rating for telic verbs is 0.501 and the median rating for atelic verbs is 0.906.

Despite the statistical significance of the difference between the ratings of telic and atelic verbs, it is not the case that all telic verbs receive ratings below a given threshold and all atelic verbs receive ratings above it. On the contrary, judgments for telic verbs span almost all the way from 0 to 1, and while judgments for atelic verbs have a much tighter distribution (with their interquartile range[1] being fully above the interquartile range for telic verbs), they still overlap in a non-negligible way. Figure 8.4 only shows a single outlier among the atelic verbs, corresponding to the atelic, manner-unspecified verb *to cut* in the perfective, non-iterative sentence stimulus *Sean had cut* (normalized acceptability rating of 0.286). This may depend on the fact that not only this verb is fairly resistant to object drop despite its atelicity, with all its ratings being within the lower 18 positions among the 72 atelic target stimuli, but this stimulus in particular also has two features which tend to favor the use of overt objects in sentences (i.e., perfectivity and lack of iterativity).

1: The interquartile range is the difference between the first quartile and the third quartile, which are the medians of the lower and the upper half of the dataset, respectively. Graphically, it is rendered as the so-called "box" in the boxplot. The other parts of a boxplot are the median (second quartile), cutting the interquartile range, and the so-called "whiskers", i.e., the minimum and maximum values in the dataset. Outliers are shown in these boxplots as little diamonds outside of the boundaries traced by the whiskers.

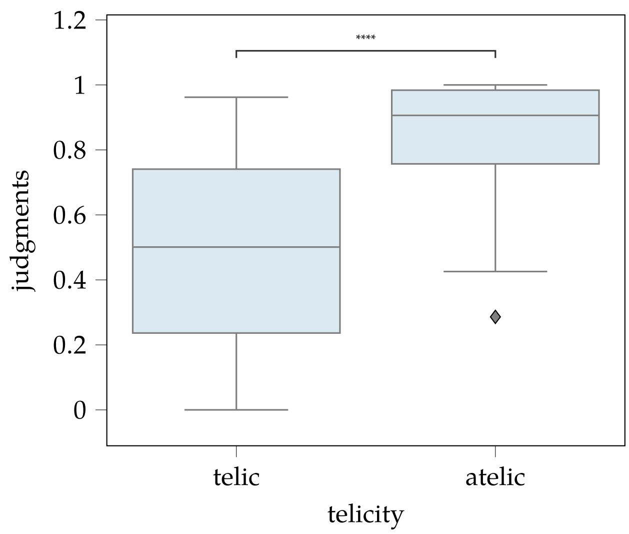

Figure 8.4: Effect of telicity on normalized acceptability judgments about object drop in English.

Perfectivity The boxplots in Figure 8.5 illustrate the main effect of perfectivity on the acceptability judgments on the implicit object construction in English. The median rating for imperfective stimuli is 0.854 while the median rating for perfective stimuli is 0.703, and a Mann-Whitney U test shows that these medians are significantly different ($p < 0.01$). This result is compatible with the hypothesis that the imperfective aspect favors the omission of direct objects and perfective aspect resists it (refer back to Section 3.2.2 and Section 6.3).
However, the distribution of judgments both for imperfective and for perfective stimuli is very sparse, given that both span almost all the way from 0 to 1, and there is significant overlap between both interquartile ranges. Neither imperfective nor perfective stimuli received outlier ratings.

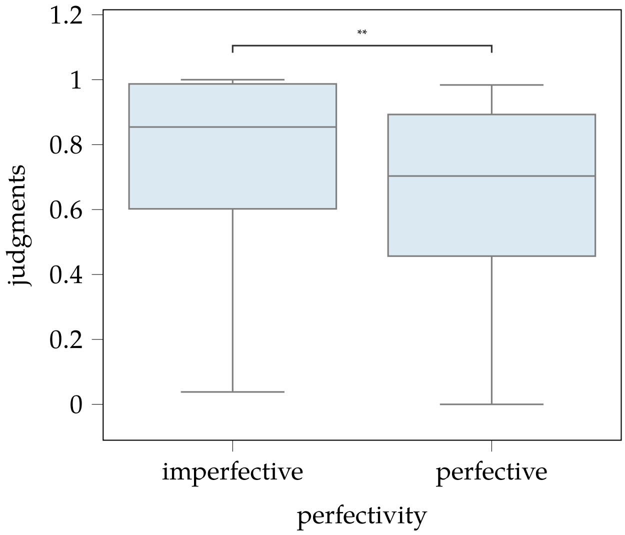

Figure 8.5: Effect of perfectivity on normalized acceptability judgments about object drop in English.

Iterativity The boxplots in Figure 8.6 illustrate the main effect of iterativity on the acceptability judgments on the implicit object construction in English. The median rating for iterative stimuli (0.826) is higher than the median rating for non-iterative stimuli (0.753), consistently with the literature on the matter (refer back to Section 3.3.2 and Section 6.4). However, the difference is not stark enough to be statistically significant according to a Mann-Whitney U test, which may depend on native speakers being less sensitive to iterativity if compared to other linguistic factors (such as telicity and perfectivity) when it comes to judging the grammaticality of the implicit object construction.
Once again, the distribution of judgments for both types of stimuli covers almost all the possible 0-1 range, and there are no outlier ratings.

Manner specification The boxplots in Figure 8.7 illustrate the main effect of manner specification on the acceptability judgments on the implicit object construction in English. A Mann-Whitney U test shows that the median rating for manner-unspecified verbs (0.898) is significantly higher ($p < 0.001$) than the median rating for manner-specified verbs (0.645), consistently with the literature and the hypothesis (refer back to Section 3.1.3 and Section 6.5).
The distribution of judgments for manner-specified verbs ($0.060 \div 0.987$) is more sparse than the distribution of judgments for manner-unspecified

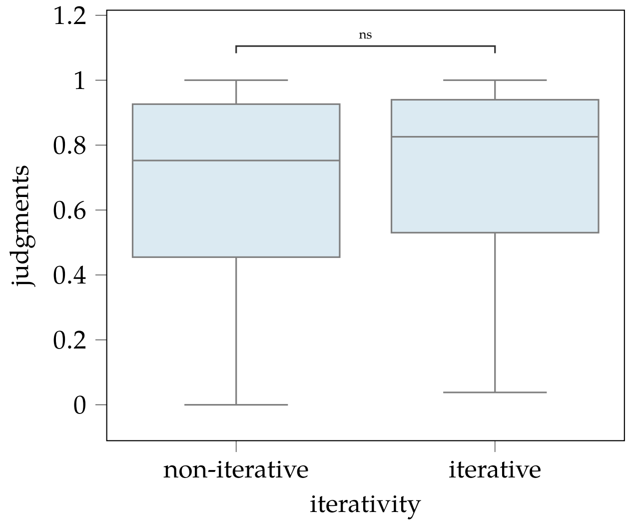

Figure 8.6: Effect of iterativity on normalized acceptability judgments about object drop in English.

verbs (0.286 ÷ 1), if one does not consider the five outliers among the latter. These outliers are the ratings for:

- the four target stimuli for the verb *to break* (perfective non-iterative[2] 0, imperfective iterative 0.038, perfective iterative 0.039, imperfective non-iterative 0.050);
- the perfective, non-iterative stimulus for the verb *to build* (0.124), i.e., *Paul had built*.

2: As detailed in Section 7.3, semantic selectivity, telicity, and manner specification are properties of each target verb, which participate in a 2x2x2 experimental design where the presence of a direct object, perfectivity, and iterativity are manipulated at the sentence level.

As was the case with the atelic outlier in Figure 8.4, the outlier stimulus for the verb *to build* is both perfective and non-iterative, making it a very unlikely candidate for felicitous object drops. The verb *to break* appears to be quite resistant to object drop regardless of the experimental conditions, given that all the target stimuli featuring it are outliers (below the lower whisker of the boxplot) in the distribution of judgments for manner-unspecified verbs.

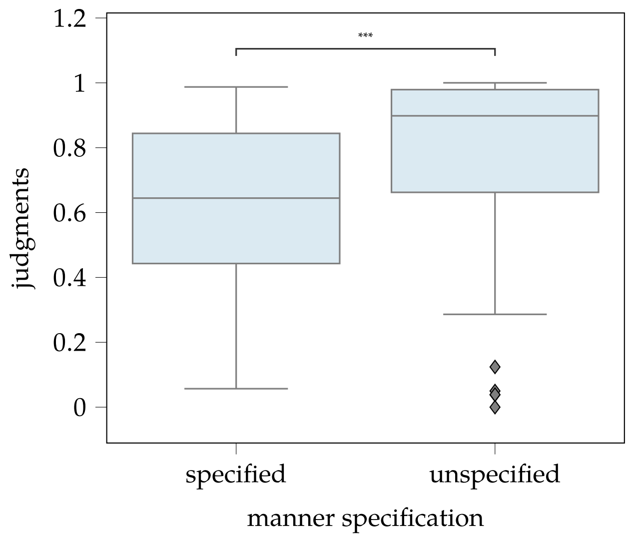

Figure 8.7: Effect of manner specification on normalized acceptability judgments about object drop in English.

8.2.3 Joint effect of predictors

In this section I will consider the joint effect of all five predictors of object drop on the grammaticality ratings in a statistical model, in order to gauge the feasibility of a linguistically-motivated probabilistic model of the implicit object construction. I will compute a linear mixed-effects model[3] for each measure of semantic selectivity I employed, accounting both for the fixed effects determined by the five predictors and for the random effects determined by my choice of verbs and participants, whereas Medina (2007, p. 131) computed a multiple linear regression (which is not endowed to account for random effects in addition to the fixed ones). I carried out my analysis using the Python package `statsmodels` and completed it with the R function `report()` of the package `easystats` (Makowski et al. 2021), which takes as its input the linear mixed model created with the R function `lmer()` of the package `lme4` (Bates et al. 2015), to compute the conditional R^2 and marginal R^2 of the model[4] .
The three linear mixed-effects models I computed are reported in:

- Table 8.1 (conditional $R^2 = 0.53$, marginal $R^2 = 0.20$), where semantic selectivity is measured with Resnik's SPS;
- Table 8.2 (conditional $R^2 = 0.53$, marginal $R^2 = 0.20$), where semantic selectivity is measured with Computational PISA;
- Table 8.3 (conditional $R^2 = 0.53$, marginal $R^2 = 0.22$), where semantic selectivity is measured with Behavioral PISA.

In general, it is already possible to observe that, while the total explanatory power is the same for the three models, Behavioral PISA generates a slightly better model than Resnik's SPS and Computational PISA when considering the fixed effects alone.

3: For observations on the feasibility of linear mixed-effects models applied to rating data, such as Likert-scale judgments, refer to Bross (2019), Cunnings (2012), Endresen and Janda (2017), Gibson, Piantadosi, and Fedorenko (2011), and Kizach (2014).

4: The conditional R^2 quantifies the total explanatory power of the model, i.e., how much both fixed effects and random effects explain the variance in the data. The marginal R^2 instead quantifies the explanatory power of fixed effects alone.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	4.117	0.334	12.333	0.000	3.463	4.771
tel[atelic]	1.504	0.388	3.870	0.000	0.742	2.265
perf[imperf]	0.541	0.045	11.903	0.000	0.452	0.630
iter[iter]	0.221	0.045	4.864	0.000	0.132	0.310
spec[nospec]	0.280	0.383	0.732	0.464	-0.470	1.031
sps	0.199	0.183	1.088	0.277	-0.159	0.557
subject Var	0.351	0.071				
verb Var	0.943	0.196				

Table 8.1: Linear mixed-effects model of the five predictors of object drop in English as fixed effects, with verb and participant subject as random effects, measuring semantic selectivity with Resnik's SPS.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	4.167	0.352	11.854	0.000	3.478	4.856
tel[atelic]	1.437	0.407	3.526	0.000	0.638	2.235
perf[imperf]	0.541	0.045	11.903	0.000	0.452	0.630
iter[iter]	0.221	0.045	4.864	0.000	0.132	0.310
spec[nospec]	0.261	0.390	0.669	0.504	-0.504	1.026
cpisa	0.182	0.198	0.920	0.358	-0.206	0.571
subject Var	0.351	0.071				
verb Var	0.956	0.198				

Table 8.2: Linear mixed-effects model of the five predictors of object drop in English as fixed effects, with verb and participant subject as random effects, measuring semantic selectivity with Computational PISA.

The three models all converge[5] on similar results. In particular, they show that:

- the effect of (im)perfectivity, (a)telicity, and iterativity is statistically significant and positive;

5: Linear mixed-effects model are generated by optimizing a complex function over thousands of steps. The optimizer, which in the case of lme4::lmer is nloptwrap by default, stops as soon as it finds a solution to the optimization problem or after a given number of unsuccessful iterations. In the first case, the model is said to have converged and it is reliable. In the other case, a warning is issued that the model has not converged, meaning that the estimates it yielded may not be reliable.

Table 8.3: Linear mixed-effects model of the five predictors of object drop in English as fixed effects, with verb and participant subject as random effects, measuring semantic selectivity with Behavioral PISA.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	4.238	0.336	12.601	0.000	3.579	4.898
tel[atelic]	1.278	0.405	3.157	0.002	0.485	2.072
perf[imperf]	0.541	0.045	11.903	0.000	0.452	0.630
iter[iter]	0.221	0.045	4.864	0.000	0.132	0.310
spec[nospec]	0.306	0.367	0.832	0.406	-0.415	1.026
bpisa	0.331	0.190	1.740	0.082	-0.042	0.705
subject Var	0.351	0.071				
verb Var	0.883	0.183				

- the effect of manner (non-)specification and semantic selectivity is statistically non-significant and positive.

These results lead to the conclusion that the joint effect of the five predictors of indefinite object drop in English can provide a good explanation of the acceptability judgments relative to this phenomenon. Therefore, this means that it will make sense to compute a linguistically-motivated, probabilistic model of the effect of all the five predictors on the grammaticality of the implicit object construction. The implementation of this model within the framework of Stochastic Optimality Theoretic (in the linear variant defined by Medina (2007)) will be discussed in Chapter 9, where I will also discuss its predictions relative to the grammaticality of indefinite object drop in English and Italian, and perform a comparison between my own results and the ones Medina obtained in her original model.

8.3 Italian results

8.3.1 Semantic selectivity

The effect of semantic selectivity on the acceptability of the implicit object in Italian is quantified by means of a Pearson correlation between them. The results of this computation are visualized in Figure 8.8 for Resnik's SPS, in Figure 8.9 for Computational PISA, and in Figure 8.10 for Behavioral PISA.

What I observed in Section 8.2 about the correlations between the three models of semantic selectivity and human judgments about object drop in English still holds true, *mutatis mutandis*, when considering the Italian data. First of all, it appears that Resnik's SPS is once again the worst-performing model among the three (with a staggeringly low, non-significant Pearson's r of -0.055), Computational PISA makes the situation somewhat better but still not enough to be statistically significant (Pearson's r = 0.223), and Behavioral PISA is quite a good model of semantic selectivity (Pearson's r = 0.481, p value = 0.007).

Once again, this state of affairs mirrors the situation depicted in Section 6.1.4 (see Table 6.3 in particular), where I made the case that Behavioral PISA, the human judgment-based benchmark model of semantic selectivity, correlates better with Computational PISA than with Resnik's SPS. Moreover, the non-significant correlation yielded by both Resnik's SPS and Computational PISA in Italian relative to the acceptability judgments on the implicit object construction mirrors the high correlation shown in Table 6.3 between Resnik's SPS and Computational PISA. It would thus appear that the itWaC corpus has a stronger effect on the semantic similarity measures based on it than ukWaC has on the ones computed for English, as shown earlier in Section 6.1.4.

Concluding, Behavioral PISA is a satisfactory predictor of object drop in Italian, with a correlation against human acceptability judgments on object drop comparable with the one obtained by Behavioral PISA in English (compare Figure 8.3 and Figure 8.10). However, as is the case with English, Behavioral PISA is not able to fully predict the feasibility of object drop for a given transitive verb.

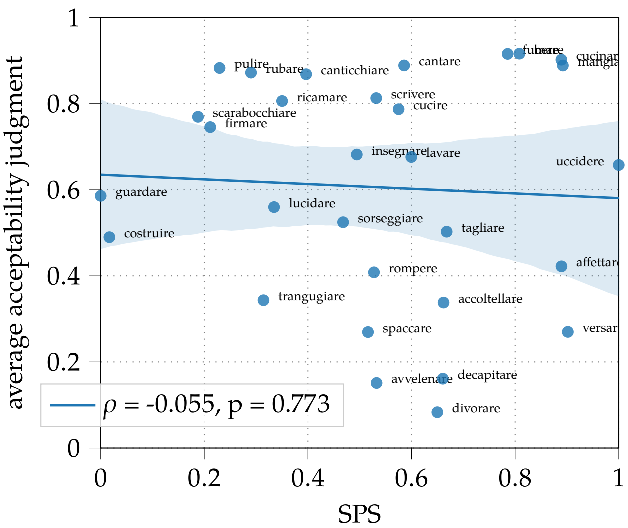

Figure 8.8: Correlation between semantic selectivity (Resnik's SPS) and normalized acceptability judgments on object drop in Italian.

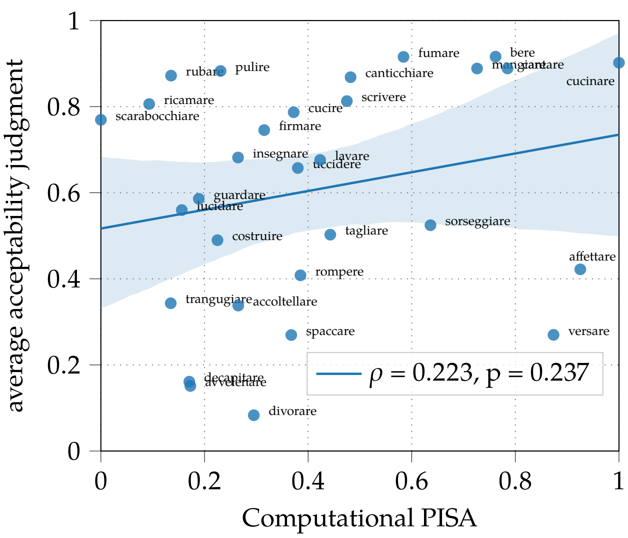

Figure 8.9: Correlation between semantic selectivity (Computational PISA) and normalized acceptability judgments on object drop in Italian.

8.3.2 Binary predictors

Telicity The boxplots in Figure 8.11 illustrate the main effect of telicity on the acceptability judgments on the implicit object construction in Italian. A Mann-Whitney U test reveals that the median judgment for atelic verbs (0.823) is significantly higher ($p < 0.0001$) than the median judgment for telic verbs (0.384), consistently with abundant literature on the effect of telicity on object drop (refer back to Section 3.2.1 and Section 6.2).

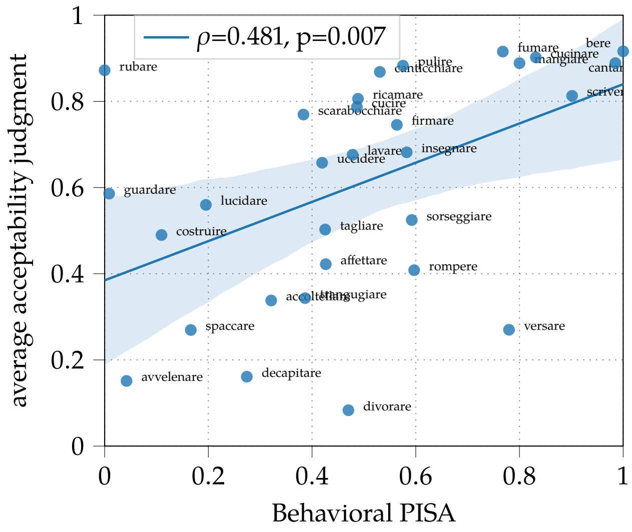

Figure 8.10: Correlation between semantic selectivity (Behavioral PISA) and normalized acceptability judgments on object drop in Italian.

The interquartile ranges of telic and atelic verbs do not overlap, as shown in the boxplots, but the overall distributions of ratings for the two types of verbs do indeed overlap for the most part. This shows that, despite the high statistical significance of the difference in judgments between telic and atelic verbs, telicity alone is not a sufficient predictor of object drop in Italian.

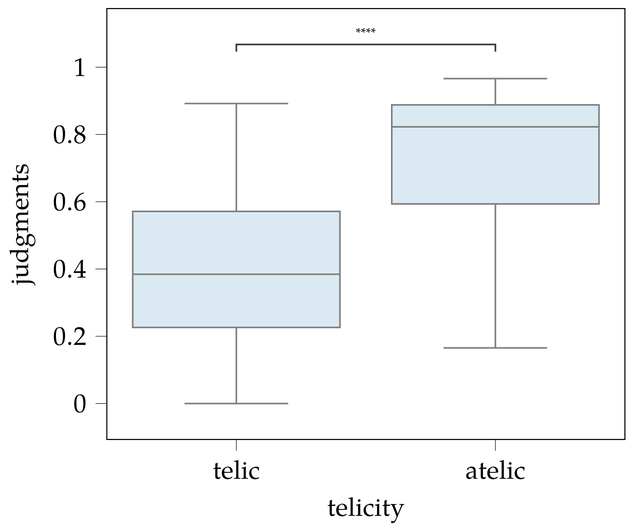

Figure 8.11: Effect of telicity on normalized acceptability judgments about object drop in Italian.

Perfectivity The boxplots in Figure 8.12 illustrate the main effect of perfectivity on the acceptability judgments on the implicit object construction in Italian. The median rating for imperfective stimuli (0.670) is significantly higher ($p < 0.05$) than the median rating for perfective stimuli (0.562), consistently with the hypothesis (refer back to Section 3.2.2 and Section 6.3).

However, the distribution of ratings for both imperfective and perfective stimuli is quite sparse, and thus there is significant overlap between them. The significant main effect of perfectivity on the grammaticality of the implicit object construction cannot be considered reason enough to use it as the sole predictor of object drop.

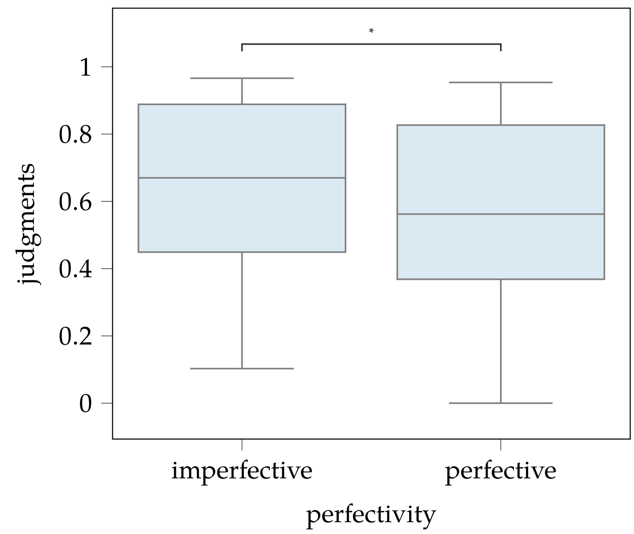

Figure 8.12: Effect of perfectivity on normalized acceptability judgments about object drop in Italian.

Iterativity The boxplots in Figure 8.13 illustrate the main effect of iterativity on the acceptability judgments on the implicit object construction in Italian. Interestingly, both in English (see Section 8.2) and Italian there is no significant main effect of iterativity on the grammaticality of the implicit object construction, and once again the question arises of whether this depends on the weakness of this factor if compared against the other predictors, or whether it will be solved by considering its action in a joint statistical model of all five predictors. The median for iterative stimuli is indeed higher than the median for non-iterative stimuli (0.659 the former, 0.645 the latter), consistently with literature on the matter (refer back to Section 3.3.2 and Section 6.4), but the difference is way too small to even approach statistical significance. Moreover, there is almost complete overlap between the distributions of judgments for both types of stimuli.

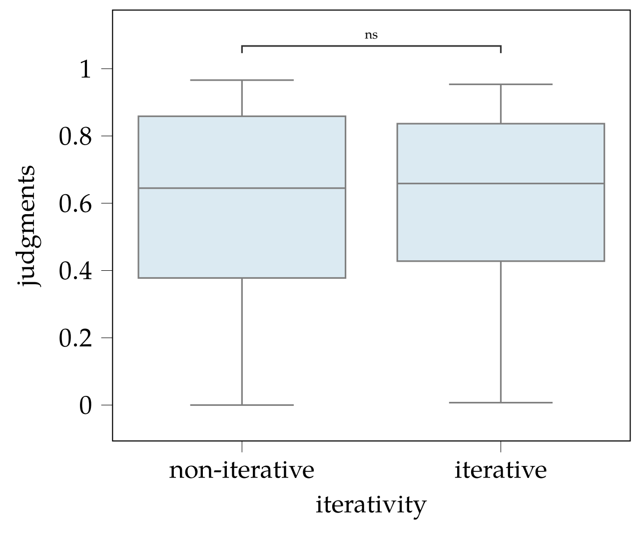

Figure 8.13: Effect of iterativity on normalized acceptability judgments about object drop in Italian.

Manner specification The boxplots in Figure 8.14 illustrate the main effect of manner specification on the acceptability judgments on the implicit object construction in Italian. A Mann-Whitney U test reveals the difference between the medians of judgments for manner-specified (0.469) and manner-unspecified (0.833) verbs to be statistically significant

($p < 0.0001$), consistently with expectations (refer back to Section 3.1.3 and Section 6.5).
The distribution of ratings for manner-unspecified verbs is tighter than the distribution of ratings for manner-specified verbs, but there is still relevant overlap between them despite the high statistical significance of their difference.

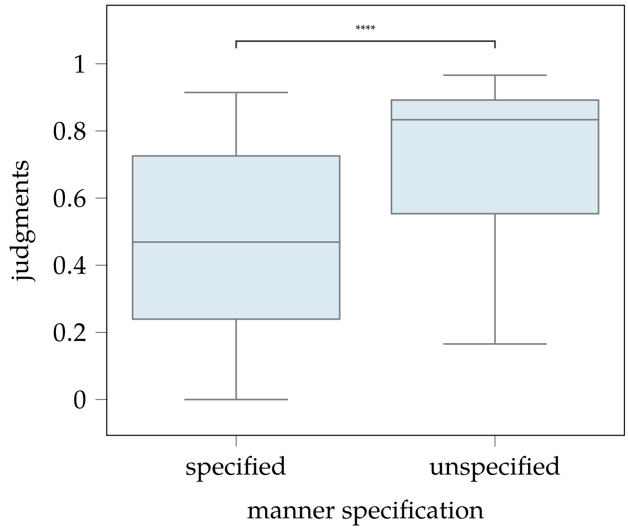

Figure 8.14: Effect of manner specification on normalized acceptability judgments about object drop in Italian.

8.3.3 Joint effect of predictors

Mirroring what I did in Section 8.2.3 about English, I will now report the results of three linear-mixed effects models (one for each different measure of semantic selectivity) I computed to account for the joint effect of my five predictors of object drop on the acceptability ratings provided by native speakers of Italian. These models are reported in:

- Table 8.4 (conditional $R^2 = 0.47$, marginal $R^2 = 0.14$), where semantic selectivity is measured with Resnik's SPS;
- Table 8.5 (conditional $R^2 = 0.47$, marginal $R^2 = 0.14$), where semantic selectivity is measured with Computational PISA;
- Table 8.6 (conditional $R^2 = 0.47$, marginal $R^2 = 0.15$), where semantic selectivity is measured with Behavioral PISA.

It appears that, in Italian as in English, the three models have the same total explanatory power (as quantified by the conditional R^2), but Behavioral PISA is the best measure of semantic selectivity if compared with Resnik's SPS and Computational PISA because it contributes to determine, *ceteris paribus*, the best linear mixed-effects model when only considering the fixed effects.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	4.677	0.255	18.367	0.000	4.178	5.176
tel[atelic]	0.950	0.279	3.402	0.001	0.403	1.498
perf[imperf]	0.308	0.041	7.610	0.000	0.229	0.388
iter[iter]	0.061	0.041	1.495	0.135	-0.019	0.140
spec[nospec]	0.550	0.279	1.976	0.048	0.004	1.096
sps	-0.093	0.131	-0.709	0.478	-0.351	0.164
subject Var	0.403	0.090				
verb Var	0.489	0.115				

Table 8.4: Linear mixed-effects model of the five predictors of object drop in Italian as fixed effects, with verb and participant subject as random effects, measuring semantic selectivity with Resnik's SPS.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	4.674	0.266	17.550	0.000	4.152	5.196
tel[atelic]	0.970	0.288	3.369	0.001	0.406	1.535
perf[imperf]	0.308	0.041	7.610	0.000	0.229	0.388
iter[iter]	0.061	0.041	1.495	0.135	-0.019	0.140
spec[nospec]	0.533	0.286	1.863	0.062	-0.028	1.094
cpisa	-0.033	0.141	-0.237	0.813	-0.310	0.243
subject Var	0.403	0.090				
verb Var	0.498	0.117				

Table 8.5: Linear mixed-effects model of the five predictors of object drop in Italian as fixed effects, with verb and participant subject as random effects, measuring semantic selectivity with Computational PISA.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	4.800	0.280	17.150	0.000	4.252	5.349
tel[atelic]	0.830	0.311	2.668	0.008	0.220	1.439
perf[imperf]	0.308	0.041	7.610	0.000	0.229	0.388
iter[iter]	0.061	0.041	1.495	0.135	-0.019	0.140
spec[nospec]	0.454	0.281	1.614	0.106	-0.097	1.005
bpisa	0.140	0.155	0.903	0.367	-0.164	0.443
subject Var	0.403	0.090				
verb Var	0.483	0.114				

Table 8.6: Linear mixed-effects model of the five predictors of object drop in Italian as fixed effects, with verb and participant subject as random effects, measuring semantic selectivity with Behavioral PISA.

The three models all converge, and for the most part they yield comparable results. In more detail, they show that:

- the effect of (a)telicity and (im)perfectivity is statistically significant and positive;
- the effect of manner (non-)specification is positive, but only statistically significant in the model quantifying semantic selectivity with Resnik's SPS;
- the effect of iterativity is statistically non-significant and positive;
- the effect of semantic selectivity, which is never statistically significant, is slightly negative in the models using Resnik's SPS and Computational PISA, but positive in the model using Behavioral PISA (consistently with everything I observed in this Chapter about the three different models of semantic selectivity in Italian).

As noted before about English, these models of the joint effect of the five predictors of object drop on the acceptability judgments in Italian support the creation of a Stochastic Optimality Theoretic model of the implicit object construction, despite the different (sometimes absent) statistical significance of the individual linguistic factors.

8.4 Closing remarks

The univariate analysis of the binary predictors of indefinite object drop showed that all of them discriminate significantly between the two groups of stimuli determined by each predictor (the group where the feature is present, and the group where it is absent), with the exception of iterativity. Moreover, there is a significant correlation between semantic selectivity (when measured with PISA models in English and with Behavioral PISA in Italian) and the acceptability judgments provided by participants to the behavioral experiment.

At this point, does it make sense to compute a model of object drop considering the joint effect of all five predictors? I answered positively to this question by means of linear mixed-effects models for English and Italian, which all converged on significant results. These results also point out that the models for English are consistently, albeit just slightly, better than the models for Italian computed with the same set of predictors. This may depend on the corpora of choice (ukWaC for English, itWaC for Italian), on the way the participants to the experiments behaved when providing judgments (despite the strict protocol in Section 7.5), and also on idiosyncratic characteristics of the two languages under scrutiny. I will also engage in similar considerations in Chapter 9.

The convergence of the linear mixed-effect models proves that a model of object drop considering the joint effect of all five predictors is indeed able to account for a non-negligible amount of variance in the data. Given these results, what would the added value of a Stochastic Optimality Theoretic model (in the novel linear, non-gaussian way introduced in Chapter 5) of indefinite object drop be? The best answer to this question was already provided by Medina (2007) herself, the ideator of the original model. This thesis (and hers) concerns itself with creating a *linguistic* model of the grammaticality of the implicit indefinite object construction, with explicit constraints whose violations determine which output will be the favored one among a set of possible candidates for a given input. A linear mixed-effects model provides a *statistical* model that computes the relative grammaticality of a candidate as a sum of weighted variables, whereas the Stochastic Optimality Theoretic model by Medina computes it as the sum of the probabilities of constraint orderings. Most importantly, Medina's mathematical model keeps the input, the candidate set, the constraints, and the probabilities of constraint re-ordering explicit and knowable in every step of the computation, while in the mixed model they are all collapsed together in the weights the model computes under the hood. I will come back to the relationship between regression models and Medina's stochastic model in Section 9.3.2.

Predicting the grammaticality of implicit objects 9

In this Chapter I will model the grammaticality of the implicit object construction (refer to Chapter 2) in a Stochastic Optimality Theoretic fashion inspired by Medina (2007) (refer to Chapter 4 and Chapter 5), using five aspectual and semantic factors (refer to Chapter 3 and Chapter 6) as constraints in several models of human acceptability judgments (refer to Chapter 7) in English and Italian. Based on the results of these two behavioral experiments described in Chapter 8, I present a linguistically-motivated probabilistic model of object drop considering the joint effect of all five predictors, which is able to account for the behavioral data I collected.

In particular, I will outline the models in Section 9.1, I will delve into the finer details of the full English and Italian models of object drop as a function of Behavioral PISA (introduced in Section 6.1.3) in Section 9.2, and I will draw some conclusions about relevant linguistic and mathematical aspects of these models in Section 9.3.

9.1 Introduction

9.1.1 Models

In this thesis, I build upon the foundations laid by Medina (2007), which I detailed in Chapter 5. In a nutshell, her Stochastic Optimality Theoretic analysis of the implicit object construction was focused on English, and used a set of only three predictors (semantic selectivity, telicity, and perfectivity) as constraints in the model. Moreover, she measured the verbs' semantic selectivity using the Selectional Preference Strength values originally computed by Resnik (1993, 1996), which poses clear limitations in the choice of transitive verbs to include in the model and which also suffers from some computational drawbacks due to being a taxonomy-based measure (more on this in Section 6.1.4).

Expanding on Medina's successful model of object drop, I bring several new ideas to the table:

- quantifying semantic selectivity with two similarity-based measures, i.e., a novel computational measure I contributed to develop in Cappelli and Lenci (2020) (Computational PISA, see Section 6.1.2), and a behavioral measure that improves on Medina's measure of Object Similarity (Behavioral PISA, see Section 6.1.3);
- modeling the implicit object construction both in English and in Italian, comparing the performance of the two models and possible language-dependent differences in the constraint re-ranking;
- computing increasingly more complex Stochastic Optimality Theoretic models of object drop, starting with Medina's three-predictor model, adding iterativity as a predictor in an intermediate model,

and computing the full five-predictor model also including manner specification among the predictors.

These additions to Medina's setting resulted in a grand total of 18 models of the implicit object construction, which are summarized in Table 9.1 for the reader's convenience.

Table 9.1: The 18 Stochastic Optimality Theoretic models of object drop I computed for English and Italian.

	SPS	Comp PISA	Behav PISA
StOT basic	eng | ita	eng | ita	eng | ita
StOT +iter	eng | ita	eng | ita	eng | ita
StOT +iter +spec	eng | ita	eng | ita	eng | ita

The basic Stochastic Optimality Theoretic model of English judgments using Resnik's SPS as a measure of semantic selectivity is, as recalled earlier in this Section, a replication of the model by Medina (2007) employing the same constraints and acceptability judgments based on the same experimental protocol (but with different target verbs and an updated computational preprocessing pipeline, as explained in Chapter 7 and Chapter 8). The other 17 models are instead new.

Naturally, one could ask why it is iterativity, and not manner specification, the predictor of object drop to be included in the intermediate Stochastic Optimality Theoretic models. After all, the main effect of iterativity was shown to be non-significant in Figure 8.6 and Figure 8.13 both for English and Italian, unlike the very significant main effect of manner specification in both languages (refer back to Figure 8.7 and Figure 8.14). Crucially though, I am creating probabilistic models considering the *joint* effect of five linguistic factors on the grammaticality of the implicit object construction. Since the linear mixed-effects models for English (see Section 8.2.3) revealed a highly significant effect of iterativity and a non-significant effect of manner specification, while the two predictors were almost equally non-significant in the mixed models for Italian (see Section 8.3.3), it appeared that iterativity plays a larger role in determining the grammaticality of object drop when considered in combination with all the other linguistic factors involved.

9.1.2 Input, output, and constraints

A very short summary of the lengthy explanation of (Stochastic) Optimality Theory I provided in Chapter 4, and especially of the explanation of the novel variant by Medina (2007) in Chapter 5, is in order. In particular, I am going to retrace the way the input to the optimization process maps to the output, and I will introduce my two novel constraints after looking back on Medina's original set.

As shown in (3) in Section 5.1, the input to the syntactic optimization operated by the model has to include all the relevant lexical and semantic information that will be mapped to syntactically well-formed output forms, and nothing else. Thus, the input to my basic Stochastic Optimality Theoretic models (and Medina's) will look like (1-a), the input to the intermediate models will look like (1-b), and the input to the full models with all five predictors will look like (1-c). All inputs in (1) contain a transitive verb with a subject and an unspecified direct object (since the model deals with indefinite, not definite, object drop), a numerical value for semantic selectivity (be it Resnik's SPS, Computational PISA, or Behavioral PISA), the [+Past] feature since all verbs in the stimuli are in the past tense, and the features of the predicate relative to all the binary predictors that are relevant in the model (two in the basic model, three in the intermediate model, four in the full model).

(1) a. verb (x,y), x = subject, y = unspecified, semantic selectivity = *numerical value*, [+ Past], [± Telic], [± Perfective]
b. verb (x,y), x = subject, y = unspecified, semantic selectivity = *numerical value*, [+ Past], [± Telic], [± Perfective], [± Iterative]
c. verb (x,y), x = subject, y = unspecified, semantic selectivity = *numerical value*, [+ Past], [± Telic], [± Perfective], [± Iterative], [± Manner-Specified]

Given these inputs, the GEN component of the Optimality Theoretic grammar (see Section 4.1) generates two outputs, i.e., one with an overt (unspecified) direct object and one with an implicit (namely, omitted) direct object.
As soon as the grammar yields a complete candidate set to evaluate, the model has to pick a winner (or, in our case, assign gradient grammaticality to the implicit object output in a probability space) based on the re-ranking of the relevant constraints. For the basic model, these are the ones in (2) (adapted from Medina's ones in (12), introduced in Section 5.3).

(2) a. *INT ARG (*INTERNAL ARGUMENT STRUCTURE)
The output must not contain an overt direct object.
b. FAITH ARG (FAITHFULNESS TO ARGUMENT STRUCTURE)
All arguments in the input must be present in the output.
c. TELIC END (TELIC ENDPOINT)
Telic predicates must be bounded by an object in the output.
d. PERF CODA (PERFECTIVE CODA)
Perfective predicates must have a direct object in the output.

I also designed the two novel constraints in (3), based on theoretical observations on iterativity and manner specification first introduced in Chapter 3 and explored further in Chapter 6. NON-ITER ARG is active both in the intermediate and in the full model, while MANN-SPEC ARG is only active in the full model.

(3) a. NON-ITER ARG (NON-ITERATIVE ARGUMENT)
Non-iterative predicates must occur with a direct object in the output.
b. MANN-SPEC ARG (MANNER-SPECIFIED ARGUMENT)
Manner-specified verbs must occur with a direct object in the output.

In all my Stochastic Optimality Theoretic models, *INT ARG is a markedness constraint that gets violated when there is an overt direct object in the output, directly conflicting with all the other constraints, which are *faithfulness* constraints penalizing implicit objects. This conflict between markedness and faithfulness constraints is the very core of an Optimality Theoretic grammar (refer back to Chapter 4).
In the specific case of this thesis, an implicit object output will be (probabilistically) grammatical whenever *INT ARG is ranked above all

1: As explained in detail in Section 5.3, a constraint is vacuously satisfied when no candidate in the candidate set can violate it.

the other constraints that are active (i.e., not vacuously satisfied[1]) for a given input. For instance, the full, five-predictor model would only favor object-dropping telic, perfective, non-iterative, manner-specified candidates if *Int Arg were ranked above Faith Arg, Telic End, Perf Coda, Non-Iter Arg, and Mann-Spec Arg. The same model would only require *Int Arg to be ranked above Faith Arg to allow for object drop in atelic, imperfective, iterative, manner non-specified candidates. As explained in Chapter 5, Medina (2007) uses semantic selectivity not as a constraint itself, due to it being a continuous factor, but as a way to re-rank the other constraints with respect to *Int Arg. I am going to go back on this line of reasoning (and the underlying math) in Section 9.2.2.

9.1.3 Model comparison

Let us put aside the inner workings of Medina-inspired Stochastic Optimality Theoretic models for the time being, and let us consider whether increasing the number of predictors (and thus the number of constraints) actually determines a better understanding of the nature of the implicit object construction. I am going to come back to the mathematical details of the model in Section 9.2, where I will focus on the two best-performing models, one for English and one for Italian. Unfortunately, it would be impossible to provide a complete account of all 18 models in Table 9.1 due to space constraints, but the interested reader can find graphical summaries of their results in Appendix E.

2: The Pearson's r coefficient measuring the correlation between two variables can vary between -1 and 1. The strength of the correlation is judged by considering the absolute value of the coefficient, so that it is non-existent when r is 0, very strong when it is closer to (-)1. If r is positive the two variables are directly proportional one to another, while if it is negative they are indirectly proportional.

English An initial step to assess the absolute performance of each model, and hence to compare them and gauge their performance relative to each other, would be to compute Pearson correlations[2] between the actual acceptability judgments provided by native speakers and the predicted grammaticality values yielded by each model. The Pearson's r coefficients for English, all highly significant ($p < 0.001$), are collected in Table 9.2. These results show that the predicted values correlate quite well with the human-generated values in each model, going from 0.661 for the basic model using Resnik's SPS as a measure of semantic selectivity to 0.700 for the full model using Behavioral PISA as a measure of semantic selectivity.

Table 9.2: Pearson correlations between actual and predicted values for the nine Stochastic OT models of object drop in English.

	SPS	Comp PISA	Behav PISA
StOT basic	0.661	0.686	0.693
StOT +iter	0.664	0.689	0.696
StOT +iter +spec	0.670	0.691	0.700

However, the correlation coefficient only serves to quantify the strength of the linear relationship between actual and predicted judgments, without providing any information on how well the independent variables in the model (i.e., the predictors) explain the variance in the dependent variable (i.e., the acceptability judgments). I gleaned this information by computing the adjusted R^2 value for each model, obtaining the results in Table 9.3 relative to English.

The R-squared value, also known as "coefficient of determination", can be computed as the squared Pearson's r or, alternatively, as in Equation

9.1 (with the Summed Squared Error computed as in Equation 9.2 and the Total Sum of Squares computed as in Equation 9.3).

$$R^2 = 1 - \frac{\text{Summed Squared Error}}{\text{Total Sum of Squares}} \tag{9.1}$$

The Summed Squared Error, computed with Equation 9.2, is defined as the sum of the squared difference between each actual judgment and its corresponding judgment predicted by the model, for all judgments in the sample.

$$SSE = \sum_{i=1}^{n}(y_i - \widehat{y})^2 \tag{9.2}$$

The Total Sum of Squares is defined as the sum of the squared difference between each acceptability judgment in the sample and the average acceptability judgment, for all judgments in the sample. This is shown in Equation 9.3.

$$TSS = \sum_{i=1}^{n}(y_i - \overline{y})^2 \tag{9.3}$$

R^2 varies between 0 and 1, and it can be thought of as a percentage indicating the goodness of fit of a statistical model. However, it always increases when using additional predictors in a model, regardless of the usefulness of these variables in predicting the dependent variable. One could always add yet one more parameter to the model, overfit it to the data, and claim to have a very successful model due to a very high R^2 value. In order to overcome this major drawback of R^2, it is recommended to compute an *adjusted* R^2 value that only increases when adding relevant parameters, while it decreases when adding useless ones. It is defined as in Equation 9.4, where n is the number of acceptability judgments to be predicted and k is the number of independent variables (i.e., predictors) in the model, and it can be thought of as the percentage of variance explained by the sole indipendent variables that have an actual effect on the dependent variable.

$$\text{adjusted } R^2 = 1 - \frac{(1 - R^2)(n - 1)}{n - k - 1} \tag{9.4}$$

Let us close this much needed statistical parenthesis and go back to the summary of the English results. Oddly enough, Medina (2007, p. 147) limited her model assessment to the computation of the Summed Squared Error instead of also using it to compute the (adjusted) R^2 value, which makes it impossible to compare mathematically her results and the ones I obtained in the basic model using Resnik's SPS as a measure of semantic selectivity. Moreover, the Summed Squared Error does not have an intrinsic meaning (unlike R^2), since it just increases whenever the total number of stimuli in the experiment increases (provided there is some difference between the actual and predicted values).

According to the adjusted R^2 values in Table 9.3, the nine models explain between 42.1% (intermediate model with Resnik's SPS) and 46.8% (full model with Behavioral PISA) of the variation in the data. Given the complex nature of the implicit object construction and the interaction between all the predictors, these results, though modest in absolute terms, are quite encouraging.

Table 9.3: Adjusted R^2 values for the nine Stochastic OT models of object drop in English.

	SPS	Comp PISA	Behav PISA
StOT basic	0.422	0.457	0.467
StOT +iter	0.421	0.456	0.466
StOT +iter +spec	0.425	0.454	0.468

Let us inspect Table 9.3 in more detail. Looking at it horizontally, i.e., comparing the performance of each type of model (basic, intermediate, full) on varying the measure of semantic selectivity, it emerges that models using Behavioral PISA have a better explanatory power than models using Computational PISA, which in turn are better than models using Resnik's SPS, regardless of the number of predictors in the model. Looking at the table vertically, i.e., comparing the performance of the three increasingly rich models of object drop based on the same measure of semantic selectivity, it results that the intermediate model is consistently worse (albeit imperceptibly) than the basic model regardless of the measure of semantic selectivity, while the addition of manner specification as a predictor in the full model makes it a better fit when using Resnik's SPS and Behavioral PISA, but a slightly worse fit than the intermediate model when using Computational PISA. Moreover, consistently with conclusions drawn in Section 8.2.1, the difference in performance between the models based on Computational PISA and those based on Behavioral PISA is much lower than the difference between either of those and the models based on Resnik's SPS. In general, it is possible to conclude that a full, five-predictor model is an appropriate choice to model the grammaticality of the implicit object construction in English, and the best model among the three full models is the one using Behavioral PISA to quantify semantic selectivity. I will provide a thorough analysis of this model in Section 9.2.

Italian Let us now examine the performance of the nine Stochastic Optimality Theoretic models of object drop in Italian and compare it to the results for English I just discussed. The Pearson correlations between actual and predicted grammaticality judgments in Table 9.4, all highly significant ($p < 0.001$), show varying degrees of reliability ranging from 0.621 for basic and intermediate models using Computational PISA to 0.694 for the full model using Resnik's SPS. As for English, I evaluated the goodness-of-fit of the nine models for Italian by computing the adjusted R^2 values in Table 9.5. These coefficients show that some models are a fairly poor fit (especially the intermediate model using Computational PISA, which only explains 36.5% of the variance in the data), and only two of them have a performance comparable with the English models (the full model using Resnik's SPS explains 45.8% of the variance, the full model using Behavioral PISA explains 45.5%).

Table 9.4: Pearson correlations between actual and predicted values for the nine Stochastic OT models of object drop in Italian.

	SPS	Comp PISA	Behav PISA
StOT basic	0.637	0.621	0.655
StOT +iter	0.637	0.621	0.655
StOT +iter +spec	0.694	0.655	0.692

Taking a closer look at Table 9.5, severals conclusions can be drawn.

	SPS	Comp PISA	Behav PISA
StOT basic	0.391	0.370	0.414
StOT +iter	0.386	0.365	0.410
StOT +iter +spec	0.458	0.404	0.455

Table 9.5: Adjusted R^2 values for the nine Stochastic OT models of object drop in Italian.

Looking at it line-by-line, it appears that Computational PISA-based models are consistently the worst for each type of Stochastic Optimality Theoretic model and Behavioral PISA-based models are the best (it actually loses to Resnik's SPS in the case of the full models, but only by a negligible 0.3% difference). Looking at it column-by-column, results show that the intermediate model is consistently worse than the basic model, which in turn is consistently worse than the full model, indicating that iterativity alone is not a good addition to the basic three-predictor model devised by Medina (2007), but iterativity and manner specification together provide the model with a much stronger explanatory power. Interestingly, there is a stark difference between the performance of Behavioral PISA-based models on one hand, and the performance of models using corpus-based measures of semantic selectivity (Resnik's SPS and Computational PISA) on the other hand. This state of affairs mirrors closely the conclusions I drew in Section 8.3.1 about the way these measures of semantic selectivity were computed, also in contrast with English results. All in all, we can conclude that it makes sense to compute a five-predictor model to understand the factors regulating the implicit object construction in Italian, and it is best to implement semantic selectivity in such a model using Behavioral PISA despite the remarkable performance of the full SPS-based model (given all the drawbacks of Resnik's SPS which I pointed out throughout this Section and in Chapter 8). These results should not surprise, considering that computational models are by their very nature approximations of human judgments relative to selectional preferences of verbs.

Comparing English and Italian By looking at Table 9.3 and Table 9.5 in particular, it is evident that any given model using the same set of predictors and the same measure of semantic selectivity fits English data better than Italian data, with this difference being way more noticeable for models employing corpus-based measures of semantic selectivity than for models using Behavioral PISA (based on human similarity ratings) for the same purpose. As observed several times here and in Chapter 8, this may be most likely due to the better overall quality of the ukWaC corpus I used to model English if compared to itWaC (refer back to Section 6.1.2 for more details on the two corpora).
Another intriguing difference between English and Italian relative to object drop that emerges by comparing Table 9.3 and Table 9.5 is that the addition of manner specification to the model determines a veritable qualitative leap in the case of Italian, where the full models are way better than the basic and intermediate ones, while the same is not true of English, where the performance of full models is quite similar (although slightly better) to that of basic and intermediate models. Given that all other factors in the models, as well as the stimuli used in the experiments, are identical in all respects but the language itself, we can surmise that this difference between English and Italian models has to be ascribed to manner specification itself. We could be tempted to seek an explanation in the well-known distinction between verb-framed and satellite-framed

languages proposed by Talmy (1991, 2000) with respect to motion verbs. Verb-framed languages, such as Italian, typically encode the Path of motion in the verb root, while they (optionally) express the manner of motion via additional lexical material (e.g., *uscì correndo*, lit. 'he/she went-out running'). Satellite-framed languages, such as English, typically encode the manner of motion in the verb root and make use of particles to encode the Path of motion (e.g., *to go in*, *to fall down*). Looking at Table 9.3 and Table 9.5 through Talmy-styled lenses, it would seem that Italian speakers are much more sensitive to manner being encoded in the verb root than English speakers due to Italian being a verb-framed language, despite there being no framing-dependent differences in the surface form of the verbs used in the stimuli (e.g., Eng. *to devour* / It. *divorare*). On the flip side, the distinction between verb- and satellite-framed languages seems to have much less hold outside the domain of motion verbs (see Mastrofini (2013) about manner-of-speaking verbs in English and Italian), and therefore it is possible that the explanation for the spike in the Italian (and not in the English) adjusted R^2 values due to the manner specification parameter has to be found elsewhere.

9.2 A full account of the full models

In this Section, I will only discuss two of the 18 models in Table 9.1, namely the best-performing model of object drop in English and the best model for Italian. The interested reader can find a summary of the other models in Appendix E.
Based on Table 9.3, the best-performing model of the implicit object construction in English is the full model making use of Behavioral PISA to measure semantic selectivity. As for Italian, I would have to choose the full model quantifying semantic selectivity with Resnik's SPS based on the results in Table 9.5, but I will instead present and discuss the full model using Behavioral PISA thanks to the negligible R^2 difference between this model and the one using Resnik's SPS in Italian. Crucially, this choice will make it possible to compare the English and the Italian models, since it minimizes the differences between them.

9.2.1 A quick recap

I will now go over the logic behind Medina's variant of Stochastic Optimality Theory (first introduced in Chapter 5) which I am using to compute the full models of the gradient grammaticality of object drop in English and Italian. In a nutshell,

1. the probability of *Int Arg dominating each of the other constraints[3] is expressed as a function of the input verb's semantic selectivity (which I compute using Resnik's SPS, Computational PISA, and Behavioral PISA);
2. the values of the function are used to compute the relative probabilities of each of the 16 possible re-rankings of *Int Arg with respect to the five other constraints at play (see Table 9.6)[4] ;
3. these relative probabilities determine the relative probability (and thus grammaticality) of the implicit object output for a given input,

3: *Int Arg has to be ranked above all the other active constraints for a given input in order to obtain an implicit object output (see Section 5.3 and Section 9.1.2)

4: In order to save space, in Table 9.6 Faith Arg, Telic End, Perf Coda, Non-Iter Arg, and Mann-Spec Arg are referred to as F, T, P, N, and M, respectively.

depending on the input's semantic selectivity score and binary aspectual features.

*Int Arg ≫ {F, T, P, N, M}
T ≫ *Int Arg ≫ {F, P, N, M}
P ≫ *Int Arg ≫ {F, T, N, M}
{T, P} ≫ *Int Arg ≫ {F, N, M}
M ≫ *Int Arg ≫ {F, T, P, N}
{M, T} ≫ *Int Arg ≫ {F, P, N}
{M, P} ≫ *Int Arg ≫ {F, T, N}
{M, T, P} ≫ *Int Arg ≫ {F, N}
N ≫ *Int Arg ≫ {F, T, P, M}
{N, T} ≫ *Int Arg ≫ {F, P, M}
{N, P} ≫ *Int Arg ≫ {F, T, M}
{N, T, P} ≫ *Int Arg ≫ {F, M}
{M, N} ≫ *Int Arg ≫ {F, T, P}
{M, N, T} ≫ *Int Arg ≫ {F, P}
{M, N, P} ≫ *Int Arg ≫ {F, T}
{M, N, T, P} ≫ *Int Arg ≫ F

Table 9.6: Set of the 16 possible re-rankings of *Int Arg with respect to Faith Arg, Telic End, Perf Coda, Non-Iter Arg, and Mann-Spec Arg, these being unordered with respect one to another.

As explained in Chapter 5 relative to Medina's (2007) model, each re-ranking in Table 9.6 yields an implicit object output if *Int Arg is ranked above *all* the active constraints for a given input. Thus, for instance, the first re-ranking always yields an implicit object output regardless of the aspectual features of the input, because *Int Arg outranks all the other constraints. The second re-ranking, T ≫ *Int Arg ≫ {F, P, N, M}, only yields an implicit object output for atelic inputs (because they vacuously satisfy the Telic End constraint). For the same reason, the re-ranking {N, T, P} ≫ *Int Arg ≫ {F, M} would yield an implicit object output only for inputs where Non-Iter Arg, Telic End, and Perf Coda are vacuously satisfied, namely, iterative, atelic, imperfective inputs. Finally, the last re-ranking would only yield an implicit object output for manner-unspecified, iterative, atelic, imperfective inputs, given that *Int Arg only outranks Faith Arg.
As shown in Table 5.7, the actual computational steps needed to model object drop go backwards with respect to the three-step logic I summed up just now. Recalling the summary of Medina's computational reasoning in Section 5.5, I will follow the exact same procedure:

1. the grammaticality of the indefinite object drop is quantified via an acceptability judgment survey (refer back to Chapter 7 for the experimental setting and to Chapter 8 for the results), the results thereof are equated to the probability of an implicit object output for a given input;
2. the probability of each of the 16 possible constraint orderings in Table 9.6 can be estimated via the probability of an implicit object output (i.e., the average judgment for a given input, normalized between 0 and 1);
3. knowing the probability of each constraint ordering, it is possible to estimate the probability of *Int Arg dominating each constraint and, finally, the probability of obtaining an implicit object output with each type of input.

In particular, I will assess the mathematical procedure in Section 9.2.2, while I will discuss the probability of *Int Arg dominating each of the other five constraints and the probability of each type of input resulting in an implicit object output in Section 9.2.3 and Section 9.2.4, respectively.

9.2.2 Fitting the model

In the full Stochastic Optimality Theoretic model(s) I am going to compute, based on my four binary predictors, there are 16 different types of input. The combinatory logic behind this result is shown in Table 9.7.

Table 9.7: The four binary constraints in the full Stochastic Optimality Theoretic model give rise to 16 different types of inputs.

	telicity	perfectivity	iterativity	manner specification
input 1	+	+	+	+
input 2	+	+	+	-
input 3	+	+	-	+
input 4	+	+	-	-
input 5	+	-	+	+
input 6	+	-	+	-
input 7	+	-	-	+
input 8	+	-	-	-
input 9	-	+	+	+
input 10	-	+	+	-
input 11	-	+	-	+
input 12	-	+	-	-
input 13	-	-	+	+
input 14	-	-	+	-
input 15	-	-	-	+
input 16	-	-	-	-

As stated in Section 9.2.1, the probability of an implicit object output for each type of input in Table 9.7 is equal to the normalized acceptability rating attributed to that specific input. Then, this rating-as-probability is equated to the probability sum of all the rankings in Table 9.6 where *Int Arg is ranked above all the relevant, active constraints for the specific type of input under consideration. So, for instance, the probability of an implicit object output for a telic, perfective, non-iterative, manner-specified input is computed as in Equation 9.5, since this input violates all the five faithfulness constraints at play, making it necessary to have *Int Arg outranking all of them for an implicit object output to be licensed by the model. The probability of an implicit object output for an atelic, perfective, non-iterative, manner-specified input, instead, is computed as in Equation 9.6 because the Telic End constraint is vacuously satisfied by atelic inputs. For the same reason, the probability of an implicit object output for atelic, imperfective, iterative, manner-unspecified inputs is computed as in Equation 9.7, i.e., as the probability sum of all the rankings in Table 9.6 (since these inputs vacuously satisfy all the constraints in the model with the exception of Faith Arg).

$$p(\text{implicit})_{\text{Tel Perf Non-Iter Spec}} = p(*I \gg F, T, P, N, M) \tag{9.5}$$

$$\begin{aligned} p(\text{implicit})_{\text{Atel Perf Non-Iter Spec}} &= p(*I \gg F, T, P, N, M)+ \\ &+ p(T \gg *I \gg F, P, N, M) \end{aligned} \tag{9.6}$$

$$\begin{aligned} p(\text{implicit})_{\text{Atel Imperf Iter Non-Spec}} &= p(*I \gg F, T, P, N, M)+ \\ &+ p(T \gg *I \gg F, P, N, M) + p(P \gg *I \gg F, T, N, M)+ \\ &+ p(T, P \gg *I \gg F, N, M) + p(M \gg *I \gg F, T, P, N)+ \\ &+ p(M, T \gg *I \gg F, P, N) + p(M, P \gg *I \gg F, T, N)+ \\ &+ p(M, T, P \gg *I \gg F, N) + p(N \gg *I \gg F, T, P, M)+ \\ &+ p(N, T \gg *I \gg F, P, M) + p(N, P \gg *I \gg F, T, M)+ \\ &+ p(N, T, P \gg *I \gg F, M) + p(M, N \gg *I \gg F, T, P)+ \\ &+ p(M, N, T \gg *I \gg F, P) + p(M, N, P \gg *I \gg F, T)+ \\ &+ p(M, N, T, P \gg *I \gg F) \end{aligned} \tag{9.7}$$

Knowing the probability of an implicit object output for each type of input (i.e., the normalized judgment for that type of input), and knowing the computation of the probability sums of relative rankings which give rise to it (as just shown briefly in Equation 9.5 to Equation 9.7), it is now possible to compute the probability of *Int Arg dominating *each* of the other five constraints, which will be used later to determine the parameters of the model itself. Limiting my examples to two types of input to avoid encumbering the reader with unnecessary details, the probability of an implicit object output for a telic, perfective, non-iterative, manner-specified input[5] (computed before in Equation 9.5) can be unpacked as in Equation 9.8, while the probability of an implicit object output for an atelic, perfective, non-iterative, manner-specified input[6] (computed before in Equation 9.6) can be unpacked as in Equation 9.9. The reason behind this calculation was illustrated in Section 5.4.3, where I explained that, in Medina's (2007) original model, the probability of each individual re-ranking ordering is equal to the joint probabilities of the independent pairwise orderings that comprise it. Thus, for instance, the probability of *Int Arg outranking all the other five constraints at play (see Equation 9.5) is equal to the joint probabilities (refer to Page 80) of *Int Arg outranking Faith Arg, *Int Arg outranking Telic End, *Int Arg outranking Perf Coda, *Int Arg outranking Non-Iter Arg, and *Int Arg outranking Mann-Spec Arg (see Equation 9.8).

5: Such as the stimulus sentence *Betty had beheaded*.

6: Such as the stimulus sentence *Paul had doodled*.

$$\begin{aligned} p(\text{implicit})_{\text{Tel Perf Non-Iter Spec}} &= p(*I \gg F) \cdot p(*I \gg T) \cdot p(*I \gg P)\cdot \\ &\cdot p(*I \gg N) \cdot p(*I \gg M) \end{aligned} \tag{9.8}$$

$$\begin{aligned} p(\text{implicit})_{\text{Atel Perf Non-Iter Spec}} &= p(*I \gg F) \cdot p(*I \gg T) \cdot p(*I \gg P)\cdot \\ &\cdot p(*I \gg N) \cdot p(*I \gg M) + p(*I \gg F) \cdot [1 - p(*I \gg T)]\cdot \\ &\cdot p(*I \gg P) \cdot p(*I \gg N) \cdot p(*I \gg M) \end{aligned} \tag{9.9}$$

As introduced in Section 5.4, the main innovation by Medina (2007) within the landscape of Stochastic Optimality Theory is the definition of the ranking of *Int Arg with respect to each of the other constraints at play as a (linear) function of the input verb's semantic selectivity. Such a function takes the shape of Equation 9.10 (updating Equation 5.1 with the use of Behavioral PISA instead of Resnik's SPS). In these equations, the γ and δ parameters are, respectively, the values the function takes at mininum and maximum Behavioral PISA. As explained in Chapter 5 and in the rest of this Section, creating a model of indefinite object drop in this framework boils down to estimating the γ and δ parameters (and minimizing the Summed Squared Error between actual and predicted judgments).

$$p(\text{*Int Arg} \gg \text{con}) = \frac{\delta_k - \gamma_k}{\text{bPISA}_{max} - \text{bPISA}_{min}} \cdot (\text{bPISA}_i - \text{bPISA}_{min}) + \gamma_k \tag{9.10}$$

In particular, the five linear functions involved in my full models are shown in Equation 9.11 to Equation 9.15 (Equation 9.11 to Equation 9.13 are Medina's original Equation 5.2 to Equation 5.4).

$$p(\text{*Int Arg} \gg \text{F}) = \frac{\delta_1 - \gamma_1}{\text{bPISA}_{max} - \text{bPISA}_{min}} \cdot (\text{bPISA}_i - \text{bPISA}_{min}) + \gamma_1 \tag{9.11}$$

$$p(\text{*Int Arg} \gg \text{T}) = \frac{\delta_2 - \gamma_2}{\text{bPISA}_{max} - \text{bPISA}_{min}} \cdot (\text{bPISA}_i - \text{bPISA}_{min}) + \gamma_2 \tag{9.12}$$

$$p(\text{*Int Arg} \gg \text{P}) = \frac{\delta_3 - \gamma_3}{\text{bPISA}_{max} - \text{bPISA}_{min}} \cdot (\text{bPISA}_i - \text{bPISA}_{min}) + \gamma_3 \tag{9.13}$$

$$p(\text{*Int Arg} \gg \text{N}) = \frac{\delta_4 - \gamma_4}{\text{bPISA}_{max} - \text{bPISA}_{min}} \cdot (\text{bPISA}_i - \text{bPISA}_{min}) + \gamma_4 \tag{9.14}$$

$$p(\text{*Int Arg} \gg \text{M}) = \frac{\delta_5 - \gamma_5}{\text{bPISA}_{max} - \text{bPISA}_{min}} \cdot (\text{bPISA}_i - \text{bPISA}_{min}) + \gamma_5 \tag{9.15}$$

It is now possible to compute the probability of an implicit object output for any type of input in terms of a polynomial function computed as the product of several linear functions whose independent variable is the verb's Behavioral PISA score. This result can be obtained by plugging Equation 9.11 to Equation 9.15 into the computations of the joint probabilities of *Int Arg dominating each of the other five constraints. So, for instance, the probability of an implicit object output for a telic, perfective, non-iterative, manner-specified input (Equation 9.8) can be computed as in Equation 9.16.

$$
\begin{aligned}
&p(\text{implicit})_{\text{Tel Perf Non-Iter Spec}} = \\
&= [\frac{\delta_1 - \gamma_1}{\text{bPISA}_{max} - \text{bPISA}_{min}} \cdot (\text{bPISA}_i - \text{bPISA}_{min}) + \gamma_1] \cdot \\
&\cdot [\frac{\delta_2 - \gamma_2}{\text{bPISA}_{max} - \text{bPISA}_{min}} \cdot (\text{bPISA}_i - \text{bPISA}_{min}) + \gamma_2] \cdot \\
&\cdot [\frac{\delta_3 - \gamma_3}{\text{bPISA}_{max} - \text{bPISA}_{min}} \cdot (\text{bPISA}_i - \text{bPISA}_{min}) + \gamma_3] \cdot \\
&\cdot [\frac{\delta_4 - \gamma_4}{\text{bPISA}_{max} - \text{bPISA}_{min}} \cdot (\text{bPISA}_i - \text{bPISA}_{min}) + \gamma_4] \cdot \\
&\cdot [\frac{\delta_5 - \gamma_5}{\text{bPISA}_{max} - \text{bPISA}_{min}} \cdot (\text{bPISA}_i - \text{bPISA}_{min}) + \gamma_5] \qquad (9.16)
\end{aligned}
$$

All the variables in such equations are known, with the exception of γs and δs, which are the values the linear functions take at bPISA$_{min}$ and bPISA$_{max}$, respectively. Based on Medina's method (see Page 86), the computational model of the indefinite object construction takes as input the acceptability judgments (normalized between 0 and 1) and the Behavioral PISA scores for all the target stimuli, and optimizes the relevant polynomial functions (such as the one in Equation 9.16) so that:

- δ_i and γ_i fall between 0 and 1;
- the Summed Squared Error[7] between the actual judgments and the ones predicted by the model are minimized.

7: Refer to Equation 9.2 for a definition of Summed Squared Error.

In practice, the script creates the model by associating to each input stimulus sentence the correct equation of the type illustrated in Equation 9.16, according to the aspectual features of the sentence. The probability of obtaining an implicit object output with that type of input corresponds to the (normalized) acceptability judgment human participants provided in the Likert-scale experiment.
Medina (2007, p. 135) made use of Excel Solver to estimate the γs and δs of the linear functions, while I did so by coding a custom Python script that makes use of the `curve_fit` method of the `optimize` function of the SciPy library (Virtanen et al. 2020). My script, which is fully documented and commented for the convenience of future researchers, can be perused and downloaded from my GitHub profile[8] . The interested reader can also download the raw data I used as input in the model from here[9] .

8: https://github.com/giuliacappelli/MedinaStochasticOptimalityTheory

9: https://github.com/giuliacappelli/dissertationData

9.2.3 Parameters of the linear functions

In this Section, I am going to discuss the probability of *Int Arg outranking each of the other five constraints at play in my full models of object drop in English and Italian. As explained in Chapter 4 and in Section 9.2.2, these probabilities stem from the estimation of the γs and δs of the linear functions in Equation 9.11 to Equation 9.15, whose product yields a different polynomial function (such as Equation 9.16) for each type of input among the 16 that are being modeled here (Table 9.7).

English The parameters of the five linear functions used to estimate the probability of *Int Arg being ranked above the other five constraints in English are reported in Table 9.8. The values taken by the γs and δs show that the re-ranking probability of *Int Arg depends indeed on the semantic selectivity of verbs (here measured via Behavioral PISA), albeit in different degrees depending on the constraints. Moreover, it always shows a directly proportional relation to Behavioral PISA values, meaning that the re-ranking probability is higher for verbs with a higher semantic selectivity.

Table 9.8: Values of unknown parameters δ_i and γ_i in the full Stochastic Optimality Theoretic model of the implicit object construction in English.

	γ	δ
p(*Int Arg $\gg$ Faith Arg)	0.902	1.000
p(*Int Arg $\gg$ Telic End)	0.490	1.000
p(*Int Arg $\gg$ Perf Coda)	0.802	0.904
p(*Int Arg $\gg$ Non-Iter Arg)	0.922	0.999
p(*Int Arg $\gg$ Mann-Spec Arg)	0.889	1.000

These results are also represented graphically in Figure 9.1, in order to make their relation to one another more evident. First of all, it emerges that the effect of semantic selectivity on the re-ranking probability of *Int Arg is much stronger for Telic End than for the other constraints, since the curve connecting the corresponding γ and δ values is steeper than any other curve in the figure.

Moreover, while all the five curves start from different points (their γ), four of them (all but the curve relative to Perf Coda) have a δ of 1, which is the maximum possible value given the constraints on the function optimization. This means that for verbs having a Behavioral PISA score equal to 1, it would be impossible for *Int Arg to be ranked below any of Faith Arg, Telic End, Non-Iter Arg, or Mann-Spec Arg.

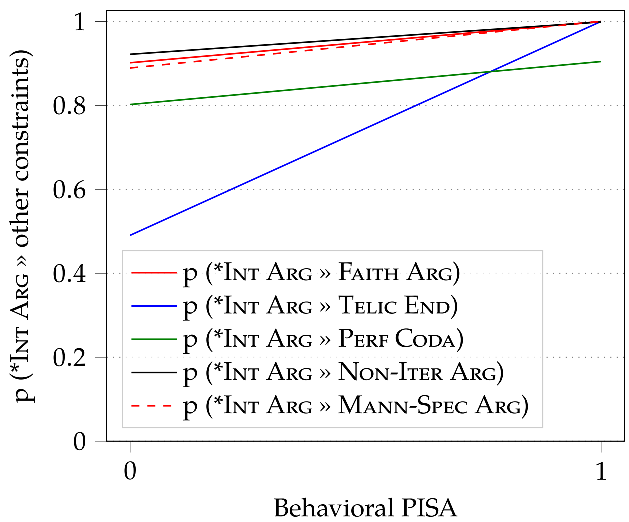

Figure 9.1: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Behavioral PISA (English full model).

Another relevant conclusion stemming from Figure 9.1 is that, regardless of semantic selectivity, *Int Arg is always more likely to rank above Non-Iter Arg than above Faith Arg, above Faith Arg more than above Mann-Spec Arg, and above Mann-Spec Arg more than above Perf Coda, since the curves associated with these rankings never cross. The very high γs and δs for these functions go to show that, in the model of English grammar hereby described, *Int Arg is quite likely to rank above all those constraints. The only exception to this trend is the probability of

*INT ARG outranking TELIC END, which is higher than the probability of *INT ARG outranking PERF CODA when the Behavioral PISA score of the input verb is higher than 0.765, lower when the Behavioral PISA score is lower than 0.765, and exactly the same (88%) if Behavioral PISA is 0.765. This depends on the fact that the curves for the re-ranking of *INT ARG with respect to TELIC END and PERF CODA cross at (0.765, 0.880), i.e., when the Behavioral PISA score of the verb in the input is 0.765 and the re-ranking probability of *INT ARG is 88%.

Italian The γ and δ parameters of the linear functions used to estimate the probability of *INT ARG being ranked above the other five constraints in Italian are in Table 9.9.

	γ	δ
p(*INT ARG ≫ FAITH ARG)	0.773	0.892
p(*INT ARG ≫ TELIC END)	0.802	0.471
p(*INT ARG ≫ PERF CODA)	0.803	0.895
p(*INT ARG ≫ NON-ITER ARG)	0.964	1.000
p(*INT ARG ≫ MANN-SPEC ARG)	0.635	1.000

Table 9.9: Values of unknown parameters δ_i and γ_i in the full Stochastic Optimality Theoretic model of the implicit object construction in Italian.

These results are also shown in Figure 9.2.

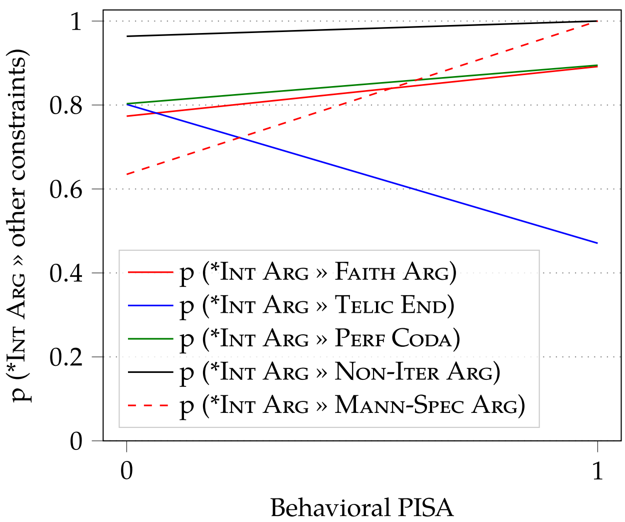

Figure 9.2: Probability of *INT ARG being ranked above each of the other constraints, varying in accordance with Behavioral PISA (Italian full model).

These re-ranking probabilities paint a complex picture. As in English grammar, also in Italian the re-ranking probability of *INT ARG varies depending on the semantic selectivity of the verb in the input. The γs are lower than the δs for the functions relative to FAITH ARG, PERF CODA, NON-ITER ARG, and MANN-SPEC ARG, meaning that the probability of *INT ARG outranking these constraints is directly proportional to the Behavioral PISA score of the input verb. On the contrary, the re-ranking probability of *INT ARG with respect to TELIC END is inversely proportional to Behavioral PISA, against expectations (refer back to Chapter 5). This unexpected result can be easily explained by looking at the relation between judgments and Behavioral PISA in Italian, shown in Figure 9.3 (same visualization as in Figure 8.10, but here the verbs are marked differently based on their telicity feature). We will remember that TELIC END is only active for telic inputs, while it is vacuously satisfied by atelic inputs, and that *INT ARG has to outrank all active constraints for a given input in order for an implicit object output to be grammatical (see Section 9.1.2). Thus, clearly, the model predicts a lower probability of *Int Arg outranking Telic End for verbs with a higher Behavioral PISA score because it is fed input data where the least semantically selective verb (*rubare*, 'to steal') is also the most grammatical one without a direct object, based on human acceptability judgments. Necessarily, all the other telic verbs have higher Behavioral PISA scores and lower acceptability judgments, thus determining the inverse re-ranking trend I observed.

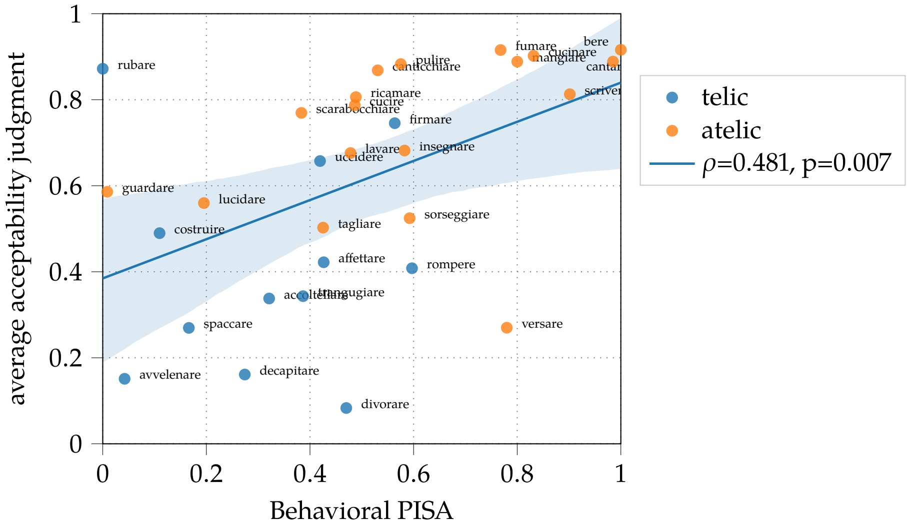

Figure 9.3: Correlation between Behavioral PISA and normalized acceptability judgments on object drop in Italian (highlighting telicity).

Going back to Figure 9.2, it is also possible to observe that the effect of Behavioral PISA on the re-ranking probability of *Int Arg with respect to the other constraints is stronger for Telic End (albeit inversely) and Mann-Spec Arg, less noticeable for Perf Coda and Faith Arg, and almost irrelevant for Non-Iter Arg. Moreover, the very high γ and δ values for the function relative to Non-Iter Arg make it so that (almost) regardless of semantic selectivity, *Int Arg will be most likely (between 96.4% and 100%) to outrank Non-Iter Arg.

The second constraint which is more likely to be outranked by *Int Arg is Perf Coda, followed by Faith Arg. The situation is complicated by relevant interactions between the function relative to Mann-Spec Arg and those relative to the other four constraints[10] . In more detail, it results that when Behavioral PISA is higher than 0.240, *Int Arg is more likely to outrank Mann-Spec Arg than Telic End (unlike when Behavioral PISA is lower), when Behavioral PISA is higher than 0.6 approximately, *Int Arg is more likely to outrank Mann-Spec Arg than Perf Coda and Faith Arg (unlike when Behavioral PISA is lower), and finally, when Behavioral PISA is 1, it is certain (probability of 100%) that *Int Arg will outrank both Non-Iter Arg and Mann-Spec Arg.

10: There is also a small interaction between the function relative to Telic End and the functions for Perf Coda and Faith Arg, but I am not describing it into detail given how close it is to the minimum possible Behavioral PISA score.

9.2.4 Predicted grammaticality of an implicit object output

At last, this Section will present the grammaticality of implicit object outputs in English and in Italian as computed by the full Stochastic Optimality Theoretic model. The grammaticality of object drop is equated to the predicted probability of an implicit object output (depending on semantic selectiviy and the four binary predictors) as determined via plugging the γs and δs of the five separate linear functions (reported in Section 9.2.3) into the 16 polynomial functions of the type exemplified in Equation 9.16, one for each type of input in the computation (obtained combinatorily as shown in Table 9.7).

English Figure 9.4 shows how the probability (hence, the grammaticality) of an implicit object output varies depending on semantic selectivity, measured with Behavioral PISA, for the 16 different types of input in the model of English grammar.
The grammaticality of object drop is gradient, since it has different probabilities depending on the type of input under consideration, and it is also shown to vary based on semantic selectivity —if it were not so, the 16 curves in the figure would all be still separate from one another, but all horizontally flat. Consistently with the values of the γs and δs shown in Figure 9.1, the probability of an implicit object output in the English grammar is always in a positive (non-linear) relation with Behavioral PISA. Moreover, it appears that all kinds of input warrant the possibility of dropping the direct object at least to some degree, given that no function in the figure ever reaches zero (the lowest value is around 0.3).
Let us look more closely at the different inputs and their probability of yielding an implicit object output. Figure 9.4 presents four distinct bundles of curves, corresponding to atelic imperfective inputs (the black lines), atelic perfective inputs (the green lines), telic imperfective inputs (the blue lines), and telic perfective inputs (the red lines). These bundles are arranged so that the two bundles for atelic inputs move along a different direction than the one of the two bundles for telic inputs, and so that atelic imperfective inputs always favor object drop more than atelic perfective inputs, which favor it more than telic imperfective inputs (with a caveat I will discuss in a short while), which finally favor it more than telic perfective inputs. This is consistent with the steep function corresponding to the probability of *Int Arg outranking Telic End depicted in Figure 9.1, in contrast with the much milder slopes of the functions associated with Perf Coda and with the other three constraints.

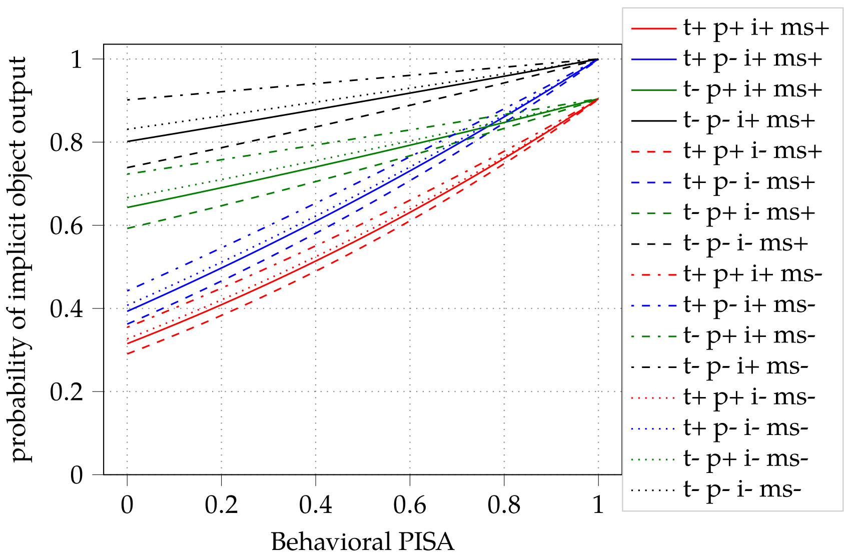

Figure 9.4: Probability of an implicit object output for each aspectual type, as a function of Behavioral PISA (English full model).

Within each bundle in Figure 9.4, the four sub-types of inputs are ordered in the same way. In particular, iterative manner-unspecified inputs (the dot-dashed lines) are the most likely to drop their direct object in the output, followed by non-iterative manner-unspecified inputs (the dotted lines), iterative manner-specified inputs (the full lines), and finally non-iterative manner-specified inputs (the dashed lines). Once again, this depends on the different probabilities of *Int Arg outranking the other constraints discussed in Section 9.2.3, which was (slightly) steeper for Mann-Spec Arg than for Non-Iter Arg.

It is also possible to uncover here the interaction between the effects of telicity and perfectivity shown in Figure 9.1 in the shape of the intersection between the functions relative to the re-ranking of *Int Arg with respect to Telic End and Perf Coda. Indeed, the bundle of atelic perfective inputs and the bundle of telic imperfective inputs in Figure 9.4 cross when Behavioral PISA is about 0.8, so that telic imperfective inputs are more likely to yield an implicit object output than atelic perfective inputs when Behavioral PISA is higher than 0.8 approximately.

Finally, it is possible to observe that while the 16 different types of input all have different probabilities of dropping the direct object in the output when Behavioral PISA is close to zero, these differences become much smaller the higher Behavioral PISA gets. Eventually, when Behavioral PISA is equal to one, all imperfective inputs are sure to drop their direct object (probability of 100%), while all perfective inputs are about 90% likely to do so.

The full picture is indeed consistent with the hypothesis, based on the expected relation between the five predictors and the likelihood of object drop (see Chapter 3 and Chapter 6). Atelic imperfective iterative manner-unspecified inputs, having all the aspectual features which the literature pinpoints as likely to favor object drop, are the most likely to yield an implicit object output (probability between 90% and 100%, depending on Behavioral PISA). On the contrary, telic perfective non-iterative manner-specified inputs, bearing aspectual features that are resistant to object drop, are the least likely to allow it (probability between 30% and 90% depending on Behavioral PISA). Nevertheless, while such inputs are the most resistant to object drop if compared to all the 15 other types, they still guarantee quite a wide margin for maneuver in an absolute sense, since object drop is at least 30% probable even at the most unlikely. In general, semantic selectivity (modeled via Behavioral PISA) appears to facilitate object drop, in accordance with the re-ranking of the constraints at play.

Italian Figure 9.5 shows how the probability (hence, the grammaticality) of an implicit object output varies depending on semantic selectivity, measured with Behavioral PISA, for the 16 different types of input in the model of Italian grammar.

As in English grammar, also in Italian grammar the acceptability of object drop is shown to be gradient across 16 different types of input, and to vary according to the input verb's semantic selectivity (measured with Behavioral PISA). However, unlike in English, the probability of an implicit object output is not always in a positive relation with Behavioral PISA, due to the negative effect of semantic selectivity on the probability of *Int Arg outranking Telic End shown in Figure 9.2. Despite this glaring difference between the two grammars, all 16 kinds of input in

Italian can yield implicit object outputs to varying extents, with their probabilities ranging from little more than 30% to approximately 90% depending on Behavioral PISA and the aspectual features of the input verb. Crucially, it is never the case that an implicit object output candidate is always ungrammatical for a given input, since no function in this figure ever reaches zero.

Let us assess these intricate results in more detail. Based on the direction of the 16 curves defined by the polynomial functions associated with the corresponding types of input, a pattern emerges that results in four bundles of curves —one for atelic manner-unspecified inputs (the four black and green dotted and dot-dashed lines), one for atelic manner-specified inputs (the four black and green full and dashed lines), one for telic manner-unspecified inputs (the four red and blue dotted and dot-dashed lines), and one for telic manner-specified inputs (the four red and blue full and dashed lines). Ignoring the several intersections between the functions, which I will tackle later in this Section, the general trend has the atelic-unspecified bundle yield implicit object outputs more likely than the atelic-specified bundle, followed by the telic-unspecified bundle, and finally by the telic-specified bundle. This state of affairs reflects the results shown in Figure 9.2, where Behavioral PISA is shown to have a much greater effect on the re-ranking probability of *Int Arg with respect to Telic End (negatively) and Mann-Spec Arg (positively), than with respect to the other constraints at play.

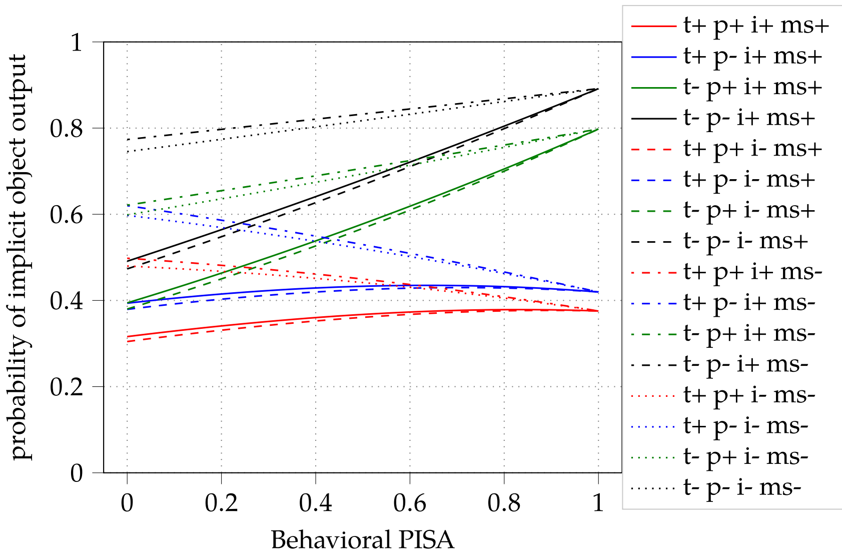

Figure 9.5: Probability of an implicit object output for each aspectual type, as a function of Behavioral PISA (Italian full model).

The direction of the curves in Figure 9.5 also stems directly from the parameters of the five linear functions estimated in Section 9.2.3 and shown in Figure 9.2. In particular,

- Behavioral PISA has a strong positive effect on the bundle of atelic manner-specified inputs, since they vacuously satisfy Telic End (which has a decreasing probability of being outranked by *Int Arg based on Behavioral PISA) but they violate Mann-Spec Arg (whose probability of being outranked by *Int Arg has a strong positive correlation with Behavioral PISA);
- Behavioral PISA has a mild positive effect on the bundle of atelic manner-unspecified inputs, since they vacuously satisfy both Telic End and Mann-Spec Arg, only violating constraints whose proba-

bility of being outranked by *Int Arg is very mildly influenced by Behavioral PISA;
- ▶ Behavioral PISA has a strong negative effect on the bundle of telic manner-unspecified inputs, since they violate Telic End while vacuously satisfying Mann-Spec Arg;
- ▶ Behavioral PISA has a negligible effect on the bundle of telic manner-specified inputs, since they violate both Telic End and Mann-Spec Arg (whose highly negative and highly positive effects, respectively, get canceled by their interaction).

Within each bundle, the iterative input is always more likely to drop the object in the output than the non-iterative input, as is the imperfective input if compared to the perfective input. The stronger effect of perfectivity than iterativity on the probability of an implicit object output[11] is easily explained, once again, by the greater steepness of the curve associated with the probability of *Int Arg outranking Perf Coda in Figure 9.2, if compared to the almost-horizontal curve associated with the probability of *Int Arg outranking Non-Iter Arg in the same plot.

11: Within each bundle, the two curves associated to imperfective inputs are always above the two curves associated to perfective inputs. The effect of iterativity is only visible within these sub-bundles, so that the iterative curve is above the non-iterative curve.

Something needs to be said about the starting and ending point of the 16 functions in Figure 9.5. The functions associated to the 8 sub-bundles I just discussed (i.e., the "minimal pairs" of functions varying only in their iterativity feature) tend to have approximately 5 different values when Behavioral PISA is zero, which is about 0.8 for atelic imperfective manner-unspecified inputs, about 0.6 for perfective manner-unspecified (telic and atelic alike) inputs, about 0.5 for atelic imperfective manner-specified inputs and telic perfective manner-unspecified inputs, about 0.4 for atelic perfective manner-specified inputs and telic imperfective manner-specified inputs, and about 0.3 for telic perfective manner-specified inputs. Due to the effect of Behavioral PISA and to the complex interactions between the constraints I highlighted before in this Section, the functions have a much tighter distribution of ending points when Behavioral PISA is maximum (i.e., equal to 1). It appears that in this case, atelic imperfective inputs are about 90% likely to drop their objects in the output, atelic perfective inputs are about 80% likely, telic imperfective inputs are little more than 40% likely, and telic perfective inputs are little less than 40% likely. This shows that for verbs with high Behavioral PISA scores, there is a major effect of telicity on the probability of them yielding an implicit object output (with atelic verbs being way more likely than telic verbs to drop their object), a secondary effect of perfectivity (with imperfective verbs being more likely than perfective verbs to drop their object), and no effect of iterativity or manner specification.

Figure 9.5 also shows several relevant intersections between the functions associated with the 16 inputs, which once again stem directly from the interactions between the probabilities of *Int Arg outranking each of the other constraints shown in Figure 9.2. In particular, the atelic manner-specified bundle crosses the telic manner-unspecified bundle because of the major interaction between the curves associated with Telic End and Mann-Spec Arg in Figure 9.2, and the telic imperfective manner-specified sub-bundle crosses the telic perfective manner-unspecified sub-bundle because of that major interaction plus the interaction between the curves associated with Mann-Spec Arg and Perf Coda.

All things considered, the overall picture is consistent with expectations. Indeed, as in English grammar, in Italian too the atelic imperfective iterative manner-unspecified inputs are the most likely to yield an implicit object output (between 80% and 90%, more or less), since they bear all the aspectual features which are known to favor object drop (refer back to Chapter 3 and Chapter 6). Conversely, telic perfective non-iterative manner-specified inputs are the most unlikely to result in an implicit object output (between 30% and almost 40%, approximately), given that they bear object-drop resistant aspectual features. It is also interesting to note that these inputs (especially the second one) are the least affected by Behavioral PISA if compared to the ones belonging to the two other bundles in Figure 9.5, meaning that the binary predictors play a much stronger role in determining their likelihood of dropping the direct object than semantic selectivity.

9.2.5 Model assessment

In this Section, I will comment on the reliability of the full Stochastic Optimality Theoretic models of object drop in English and Italian I presented in Section 9.2.4, which boils down to measuring the distance between the actual judgments provided by human participants to the experiment (designed and conducted as in Chapter 7) and the acceptability values predicted by the probabilistic model.
I will do so globally, by considering the adjusted R squared values of the two models (which I already discussed in some detail in Section 9.1.3), and locally, by computing the individual squared error of each stimulus in both experiments. I will also compute the Pearson correlations between the actual and predicted judgments for each type of input, in addition to the overall Pearson correlation relative to all the sentences in the stimuli sets. This method follows closely the analysis in Medina (2007, pp. 146–154), with a relevant difference pertaining to the global assessment of the model. Instead of computing the (adjusted) R^2 of her model, Medina evaluated the overall performance by means of the Summed Squared Error between actual and predicted judgments —a choice that, as I will argue in Section 9.3.1, can lead to unintended conclusions.

English As shown in Table 9.3, the adjusted R^2 of the full model of the implicit object construction in English (modeling semantic selectivity with Behavioral PISA) is 0.468. This means that the model explains 46.8% of the variance in the data and, thus, that it has a non-negligible explanatory power. This result, together with the highly significant Pearson correlation between actual and predicted values shown in Table 9.2 ($r = 0.700$, $p < 0.0001$), proves that the full model captures at least some of the relevant factors determining the grammaticality of object drop in English.
Let us look more closely at the results. Table 9.10 collects the Pearson correlations between actual and predicted values relative to the main input types in the model, i.e., the stimuli grouped by aspectual features in isolation, without interactions (presence/absence of telicity, perfectivity, iterativity, and manner specification). The correlations are all highly significant ($p < 0.0001$) and go from 0.412 for atelic inputs to 0.815 for manner-specified inputs. This observation is critical for the interpretation of the results, because it shows that the overall picture is quite more complex than it may seem by only considering general R^2 and Pearson values.

Table 9.10: Pearson correlations between actual and predicted judgments in the full StOT model of object drop in English modeling semantic selectivity with Behavioral PISA (main input types).

input type	Pearson r
telic	0.487
atelic	0.412
perfective	0.636
imperfective	0.730
iterative	0.654
non-iterative	0.741
manner-specified	0.815
manner-unspecified	0.546

Let us deepen the analysis by considering Table 9.11, which presents the Pearson correlations between actual and predicted values relative to all the 16 input types in the model. These results come with a caveat, namely, that only a small, variable number of stimuli feature within each input type, given that the experiment has 120 target stimuli divided into 16 different input types (refer back to Chapter 7 for more details on the experimental design). For this reason, the vast majority of the correlations in the table turned out to be statistically non-significant.
However, interesting conclusions can be drawn by the statistically significant results, which are the correlations between actual and predicted values relative to:

- the 8 telic perfective non-iterative manner-specified inputs;
- the 8 telic imperfective non-iterative manner-specified inputs;
- the 8 telic imperfective iterative manner-specified inputs (which approach statistical significance without reaching it, at p = 0.072).

The same 8 verbs are involved in these three correlations, i.e., *to chop*, *to swig*, *to sign*, *to slice*, *to poison*, *to behead*, *to knife*, *to devour*. Moreover, they all are among the least likely inputs to yield an implicit object output, based on the results depicted in Figure 9.4. This observation is compatible with the conclusions drawn by Medina (2007, pp. 150–152), where the author argues that the model is quite apt at detecting the features conditioning the grammaticality of object drop with telic inputs, while it does not appear to have the same ability with respect to atelic inputs.

Table 9.11: Pearson correlations between actual and predicted judgments in the full StOT model of object drop in English modeling semantic selectivity with Behavioral PISA (all 16 input types).

telicity	perfectivity	iterativity	manner	r	p
telic	perfective	iterative	specified	0.52	ns
telic	imperfective	iterative	specified	0.665	(0.072)
atelic	perfective	iterative	specified	0.015	ns
atelic	imperfective	iterative	specified	-0.521	ns
telic	perfective	non-iterative	specified	0.904	0.002
telic	imperfective	non-iterative	specified	0.789	0.02
atelic	perfective	non-iterative	specified	-0.455	ns
atelic	imperfective	non-iterative	specified	-0.458	ns
telic	perfective	iterative	unspecified	0.289	ns
telic	imperfective	iterative	unspecified	0.297	ns
atelic	perfective	iterative	unspecified	0.454	ns
atelic	imperfective	iterative	unspecified	0.407	ns
telic	perfective	non-iterative	unspecified	0.204	ns
telic	imperfective	non-iterative	unspecified	-0.03	ns
atelic	perfective	non-iterative	unspecified	0.213	ns
atelic	imperfective	non-iterative	unspecified	0.317	ns

However, Medina's conclusion about the model being better at handling telic inputs than atelic ones does not live up to a closer analysis of the results, this time looking at the squared error for each sentence in the stimuli set. The full list of results is available in Appendix F and, in CSV format, in a repository[12] on my GitHub profile. Incidentally, one could also wonder as to why this analysis (following Medina) takes into consideration squared errors instead of the more straightforward absolute value[13] of the difference between actual and predicted judgments. The reason is that $x^2 < |x|$ for $x \in (-1, 1)$ while $x^2 > |x|$ when $|x| > 1$ (as depicted in Figure 9.6), so that, compared to absolute error, squared error is more lenient towards small errors and more penalizing towards large errors.

12: https://github.com/giuliacappelli/dissertationData

13: The absolute value, or modulus, of a real number is its distance from 0, i.e., its numerical value regardless of its sign.

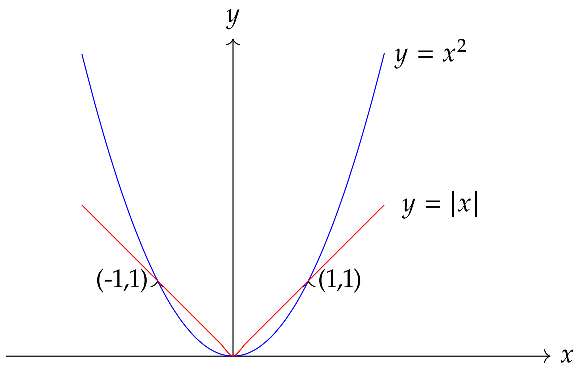

Figure 9.6: Relation between squared error and absolute error, visualized as the intersection between $y = x^2$ and $y = |x|$, respectively.

With that said, the analysis of the squared errors yielded by the full, Behavioral PISA-based model of object drop in English contradicts Medina's conclusions about atelic verbs being clumsily handled by the model. Indeed, the 13 best-performing stimuli (whose squared error is even zero) all feature atelic verbs, while they vary with respect to their other aspectual features and Behavioral PISA scores. These are collected in Table 9.12.

verb	tel	perf	iter	spec	bPISA	actual	model	sq. error
drink	-	-	+	-	0.608	0.977	0.961	0
polish	-	-	+	+	0.267	0.83	0.852	0
polish	-	-	-	+	0.267	0.795	0.803	0
polish	-	+	-	+	0.267	0.66	0.666	0
sew	-	-	+	-	0.404	0.952	0.941	0
sew	-	+	-	-	0.404	0.761	0.756	0
sing	-	-	+	-	1	0.993	1	0
sing	-	-	-	-	1	1	0.999	0
sing	-	+	-	-	1	0.91	0.903	0
smoke	-	-	+	-	0.736	0.992	0.974	0
teach	-	+	+	-	0.607	0.839	0.831	0
wash	-	-	+	+	0.475	0.882	0.893	0
wash	-	-	-	+	0.475	0.859	0.856	0

Table 9.12: The 13 best-performing stimuli in the full model of object drop in English (modeling semantic selectivity with Behavioral PISA).

Let us now consider the 5 worst-performing stimuli based on their squared errors, collected in Table 9.13. These results can be easily explained by considering the Behavioral PISA scores of the input verbs in the five stimuli.

The verb *to cut* has a (normalized between 0 and 1) Behavioral PISA score of 0.356, which would call for a similar normalized acceptability judgments at about one-third of the 0-1 scale. However, the participants to the experiment provided an average rating of 0.286, which is substantially lower than expected based on Behavioral PISA alone. The model,

Table 9.13: Squared errors of the 5 worst-performing stimuli in the full model of object drop in English (modeling semantic selectivity with Behavioral PISA).

verb	tel	perf	iter	spec	bPISA	actual	model	sq. error
cut	-	+	-	-	0.356	0.286	0.745	0.211
steal	+	+	-	-	0	0.83	0.327	0.253
steal	+	-	+	-	0	0.962	0.442	0.27
steal	+	-	-	-	0	0.927	0.407	0.27
steal	+	+	+	-	0	0.945	0.355	0.348

presented with an atelic, iterative input with a not-too-low Behavioral PISA score, which are all features of a good candidate for object drop, predictably assigns to such an input a sensible, theory-abiding rating of 0.745 (i.e., a 74.5% probability of yielding an implicit object output).

On the contrary, all the target stimuli with the verb *to steal* get a much lower predicted rating than the actual rating provided by human subjects. This happens, once again, because of the clash between Behavioral PISA (and telicity) and the judgments obtained in the experiment with native speakers. The verb *to steal* is telic (a feature that usually goes hand-in-hand with object drop) and it has the lowest Behavioral PISA score among the 30 target verbs[14] , therefore the model is quite keen on predicting a low probability of object drop for such an input. However, despite the minimum Behavioral PISA score, human subjects provided very high acceptability ratings for all the stimuli featuring this telic, manner-unspecified verb (ranging from 0.830 for the perfective non-iterative stimulus to 0.962 for the imperfective iterative stimulus, consistently with what I noted so far about the role of the binary predictors). This unexpected clash between Behavioral PISA and human judgments is easily explained. As real-world knowledge suggests, it is indeed possible to steal a wide array of items from their legitimate possessor, as well as someone's breath, or heart, or thunder, or equally essential things in common-use idioms[15] . The broad spectrum of stealable items, i.e., the low semantic density of the direct objects of the verb *to steal*, determines the low Behavioral PISA score of this verb. On the other hand, native speakers of English found it quite acceptable to use the verb intransitively even in the most object-drop resistant context (i.e., the perfective non-iterative stimulus *Diana had stolen.*), since when processing such utterances they are typically more focused on the anti-social behavior shown by the Agent than on the concrete reality of the object the Agent stole (refer to Goldberg (2005a, pp. 21–28) for a similar effect of world knowledge facilitating object drop with some so-called "taboo" verbs).

14: Please refer to Appendix C.3 for the raw Behavioral PISA scores and to Figure 8.3 for a graphical representation of the normalized scores.

15: Refer back to Page 96 for the detailed description of how I avoided idiom-dependent artifacts in my Computational and Behavioral PISA calculations.

Italian As argumented in Section 9.1.3, the full Stochastic Optimality Theoretic model of object drop in Italian, making use of Behavioral PISA to quantify semantic selectivity, achieves a satisfactory level of reliability in its results. In particular, it explains 45.5% of the variance in the data (based on the adjusted R^2 computed in Table 9.5), and there is a highly significant, large correlation between actual and predicted acceptability ratings (Pearson r = 0.692, p value < 0.0001, as shown in Table 9.4).

A very high and significant correlation with comparable Pearson rs and the same p value is also found between actual and predicted judgments when dividing the full set of stimuli into subgroups based on the main input types, as in Table 9.14. Among these 8 input types, only atelic and manner-unspecified inputs yield predicted ratings that vary remarkably from those elicited from human subjects (Pearson rs are respectively 0.374 and 0.470, statistically significant in both cases). The overall results show that, despite these two inconsistencies, so far the general picture stays true to the one depicted by the adjusted R^2 and total Pearson correlation without a great deal of nuance.

input type	Pearson r	p value
telic	0.640	0.000
atelic	0.374	0.001
perfective	0.631	0.000
imperfective	0.730	0.000
iterative	0.692	0.000
non-iterative	0.694	0.000
manner-specified	0.749	0.000
manner-unspecified	0.470	0.000

Table 9.14: Pearson correlations between actual and predicted judgments in the full StOT model of object drop in Italian modeling semantic selectivity with Behavioral PISA (main input types).

Is this also valid when considering all the 16 input types hereby modeled? Let us look at the Pearson correlations in Table 9.15. Once again, as noted about the English stimuli, large p values are to be expected in such a computation, given that the 16 slots for the input types are populated by a meager 120-sentence total. However, while in English the correlation turned out to be statistically significant for two types of input (and approaching significance for a third type), this is not the case in Italian. I interpret these results to mean that the overall picture made up by the Italian data is more nuanced that the English one, where at least some clear-cut distinctions between the aspectual types emerge.

telicity	perfectivity	iterativity	manner	r	p
telic	perfective	iterative	specified	0.329	ns
telic	imperfective	iterative	specified	0.630	ns
atelic	perfective	iterative	specified	0.466	ns
atelic	imperfective	iterative	specified	0.022	ns
telic	perfective	non-iterative	specified	0.558	ns
telic	imperfective	non-iterative	specified	0.478	ns
atelic	perfective	non-iterative	specified	0.099	ns
atelic	imperfective	non-iterative	specified	0.282	ns
telic	perfective	iterative	unspecified	0.486	ns
telic	imperfective	iterative	unspecified	0.751	ns
atelic	perfective	iterative	unspecified	0.429	ns
atelic	imperfective	iterative	unspecified	0.436	ns
telic	perfective	non-iterative	unspecified	0.552	ns
telic	imperfective	non-iterative	unspecified	0.775	ns
atelic	perfective	non-iterative	unspecified	0.419	ns
atelic	imperfective	non-iterative	unspecified	0.404	ns

Table 9.15: Pearson correlations between actual and predicted judgments in the full StOT model of object drop in Italian modeling semantic selectivity with Behavioral PISA (all 16 input types).

Finally, a more in-depth analysis of the results concerns the squared error of each stimulus in the experiment about object drop in Italian (reported in full here in Appendix F and online[16] on my GitHub profile. Comparably with English, there are 12 stimuli in the full set of 120 stimuli whose squared error is null (which equates to perfect model performance), but, unlike in English, they do not belong to any input aspectual type in particular as a group. These 12 best-performing stimuli are collected in Table 9.16.

16: https://github.com/giuliacappelli/dissertationData

Let us now consider the five worst-performing stimuli, collected in Table 9.17. Once again, the inconsistent model predictions are fully explained by the interaction between the binary aspectual features, Behavioral PISA, and the actual judgments provided by human participants.

Table 9.16: The 12 best-performing stimuli in the full model of object drop in Italian (modeling semantic selectivity with Behavioral PISA).

verb	tel	perf	iter	spec	bPISA	actual	model	sq. error
accoltellare	+	-	+	+	0.321	0.445	0.424	0
accoltellare	+	+	-	+	0.321	0.337	0.345	0
accoltellare	+	+	+	+	0.321	0.358	0.353	0
cantare	-	+	+	-	0.985	0.785	0.795	0
insegnare	-	-	+	-	0.583	0.827	0.842	0
lavare	-	-	+	+	0.478	0.661	0.672	0
lavare	-	+	+	+	0.478	0.575	0.569	0
rompere	+	+	-	-	0.597	0.419	0.431	0
scrivere	-	-	-	-	0.901	0.895	0.877	0
spaccare	+	-	-	+	0.166	0.378	0.399	0
trangugiare	+	-	+	+	0.387	0.438	0.428	0
trangugiare	+	+	+	+	0.387	0.349	0.359	0

Table 9.17: Squared errors of the 5 worst-performing stimuli in the full model of object drop in Italian (modeling semantic selectivity with Behavioral PISA).

verb	tel	perf	iter	spec	bPISA	actual	model	sq. error
firmare	+	+	-	+	0.564	0.789	0.365	0.179
versare	-	-	+	-	0.780	0.432	0.866	0.188
versare	-	-	-	-	0.780	0.299	0.859	0.313
versare	-	+	+	-	0.780	0.183	0.757	0.330
versare	-	+	-	-	0.780	0.165	0.751	0.343

The model assigns to the perfective, non-iterative stimulus featuring the telic, manner-specified verb *firmare* 'to sign' a low probability of licensing object drop (36.5%), since it has an intermediate Behavioral PISA score (0.564 on a 0-1 scale) and it also presents four out of four binary features that penalize object drop (more on this in Chapter 3 and Chapter 6). However, there is considerable distance between the rating predicted by the model and the one provided by native speakers of Italian, according to whom this verb is 78.9% likely to drop its object in the sentence *Sara aveva firmato* 'Sara had signed'. This situation is quite similar to the one I discussed on Page 152 with respect to the poor performance of the model relative to the target stimuli with the verb *to steal*, due to very high human ratings against very low model predictions. In that case, as well as in the case of the stimulus with *firmare* 'to sign' in Table 9.17, it would be possible to ascribe the rating mismatch to the fact that, regardless of semantic selectivity and aspectual features, people are quite likely to accept implicit objects in sentences assumed to be maximally resistant to object drop if they feel a culture-induced pressure to focus on the action itself performed by the Agent instead of on the Theme or Patient involved in the event. The sociolinguistic pressure has the native speaker of English accept object-less stimuli with *to steal* because the act itself of stealing is repulsive in our property-dominated culture. In the case of the verb *to sign*, Italian speakers are probably prone to find it quite acceptable when used intransitively because of Agent affectedness (discussed in Section 2.4.2 and in Section 3.1.2), since documents are usually signed to obtain something in return. An alternative explanation of the unexpected grammaticality of indefinite object drop with Italian *firmare* 'to sign' may be found in the low affeceness of its object (a low-transitivity trait among the ones defined by Hopper and Thompson (1980), as discussed in Section 2.1). However, this analysis seems to be less ideal when used to account for the unexpected grammaticality of intransitive *to steal* in English. Thus, it may be best to explain both these ratings by broadly resorting to a communication need to focus on the activity itself rather than on its results (an intransitivization mechanism described in more

detail on Page 23).
The other stimuli where the full model of the implicit object construction in Italian fails to predict judgments close to the human-elicited ones all feature the verb *versare* 'to pour'. In this case, a verb that would be the perfect candidate for object drop (due to its very high Behavioral PISA score, atelicity, and lack of manner specification) gets predictably assigned high probabilities of dropping the direct object by the model, but it is judged as variably unlikely to favor object drop by human participants. In particular, they found it extremely unlikely in both perfective stimuli, quite unlikely with the imperfective non-iterative stimulus, and almost halfway likely with the imperfective iterative stimulus (all according to the literature on the matter). The very same behavior of the verb *to pour* (in English) is found in Medina (2007, p. 148), who ascribes the mismatch between human judgments and model predictions to the very high semantic selectivity score of the verb (the highest in her 30-verb set), which "forces the model to assign a very high grammaticality to the implicit object output". This explanation can also be used to account for the judgments-predictions mismatch for the Italian stimuli featuring the verb *versare* 'to pour'. Another explanation, not contradicting the previous one, is possible considering that while the computation of Behavioral PISA is based on a single sense of the verb (refer back to Section 6.1.3 for the details), i.e., the literal sense relative to liquid pouring, the verb *versare* in Italian is also used to refer to the act of depositing liquid assets (such as one's paycheck) in a bank account. Thus, the verb gets assigned a high Behavioral PISA score because all *versare*-able objects in the computation are actual liquids one can pour, but native speakers participating in the Likert-scale experiment may well have been thinking of the other sense of the verb when providing their judgments, given that the sentences were presented on the screen one by one without additional intra- or extra-linguistic context (refer back to Chapter 7). For instance, when presented with the object-less stimulus sentence *Marta stava versando* (lit. *Marta was pouring*), they may have imagined a scenario where Marta was depositing her paycheck in her bank account, rather than pouring something from a pitcher.

9.2.6 Comparing the English and Italian models

Parameters of the linear functions and predicted grammaticality Several similarities, as well as some crucial differences, between the English and the Italian grammars with respect to their tolerance for the implicit object construction jumped to the eye throughout the analysis of both full Stochastic Optimality Theoretic models in this Chapter. Let us compare them in all relevant respects step by step.
The first glaring difference between the grammars of English and Italian emerges by comparing the parameters of the linear functions used to estimate the probability of *Int Arg outranking the other five constraints at play[17] , as collected in Table 9.8 and Table 9.9 (and graphically represented in Figure 9.1 and Figure 9.2). Based on these parameters, it appears that the role of the five predictors of object drop is not the same in the two grammars, given that:

- in English *Int Arg is more likely to outrank Faith Arg than Perf Coda, while in Italian the opposite holds;

17: Refer back to Section 9.2.2 for the full account of this computation.

- the function describing the probability of *Int Arg outranking Mann-Spec Arg as a function of semantic selectivity is much steeper in Italian than in English, and in Italian it interacts noticeably with all the other functions;
- in English the probability of *Int Arg outranking Telic End is a positive function of semantic selectivity, just like all the other functions taken into consideration here, while in Italian it actually *decreases* with the increase of Behavioral PISA.

Two aspects of the English and Italian grammars are instead the same in both, i.e., the re-ranking probability of *Int Arg is indeed a function of semantic selectivity (here computed via Behavioral PISA), and *Int Arg is always most likely to outrank Non-Iter Arg (with a negligible effect of semantic selectivity).
As observed in Section 9.2.4, these differences in the re-rankings of *Int Arg yield predictable differences in the predicted grammaticality of an implicit object output for the 16 input types in the two languages, as depicted in Figure 9.4 for English and Figure 9.5 for Italian. In particular, a comparison between the two figures shows that:

- the 16 curves can be divided into four main bundles in both languages, but these are defined by telicity and perfectivity in English, by telicity and manner specification in Italian;
- the values of the 16 functions in English always increase when Behavioral PISA increases, while in Italian they decrease for telic, manner-unspecified inputs (regardless of their perfectivity and iterativity);
- while the probabilities of licensing an implicit object output are more or less the same in English and in Italian for verbs with a low Behavioral PISA score, they are much higher in English than in Italian for verbs with a high Behavioral PISA score (or, more accurately, the high probabilities shown by all English verbs are only reached by atelic verbs in Italian);
- imperfective inputs in English always drop their object for maximally semantically selective verbs (i.e., Behavioral PISA = 1), while no Italian input ever has a 100% probability of yielding an implicit object output.

However, several commonalities between the two languages still emerge:

- the object-drop probability range is comparable in both languages (30-100% in English, 30-90% in Italian);
- the effect of Behavioral PISA is different in the two languages, but it is still relevant nevertheless;
- the four binary predictors act according to expectations in both languages, i.e., atelic verbs are more likely than telic verbs to drop their object, imperfective verbs more than perfective verbs, iterative verbs more than non-iterative verbs, and manner-unspecified verbs more than manner-specified verbs.

Evaluation of the models' performance Despite the different way English and Italian re-rank *Int Arg with respect to the other constraints and the different probabilities of dropping the object they assign to the 16 input types under examinations, the performances of the two full Stochastic Optimality Theoretic models are strikingly similar (as just

discussed in Section 9.2.5). In general, they are almost equally able to model the grammaticality of implicit object outputs based on the five predictors, given that the model for English has an adjusted R^2 equal to 0.468 and a 0.700 Pearson r ($p < 0.0001$) for the correlation between actual and predicted acceptability ratings, while the model for Italian has a 0.455 adjusted R^2 and a 0.692 Pearson r ($p < 0.0001$).
Moreover, the close similarity between the two grammars still holds when performing a much finer-grained comparison, namely, considering the squared error of the ratings the model predicted for each stimulus. The stimuli having a null squared error (up to three decimal places) are 13 in English and 12 in Italian, out of 120 stimuli. On the opposite side of the spectrum, the five worst-performing stimuli prediction-wise in English (see Table 9.13) and in Italian (see Table 9.17) feature different verbs in the two languages, but they have comparable squared errors. In particular, the squared errors for the five English stimuli range between 0.211 and 0.348, while for Italian they range between 0.179 and 0.343.
Based on these observations and other conclusions drawn in this Chapter and in Chapter 8, it would finally be worth mentioning that the model of the grammaticality of object drop in English is consistently better than the Italian model, albeit slightly. The reason for this small discrepance cannot be found in noisy corpus data, since Behavioral PISA is not a corpus-based semantic selectivity measure like Computational PISA, nor in the experimental design, since it is the same in both cases. I would attribute it to a more clear-cut effect of the predictors in English than in Italian, as argued in this Chapter, and also to semantic differences between the English target verbs and their Italian translations. Indeed, despite putting all the care possible in minimizing polysemy effects and in finding the most appropriate translation to build comparable verb sets (refer back to Page 106), some differences due to language-internal idiosyncracies still emerge. For instance, English *to smoke* and Italian *fumare* both typically refer to the act of inhaling the byproduct of tobacco combustion, but the former can also refer to a technique used to cure meat and fish (which would correspond to Italian *affumicare*). To give another example, English *to break* and Italian *rompere* both refer to the act of destroying the physical integrity of something, but they also participate in a variety of language-specific idioms (e.g., *break a leg!* in English and *rompere le palle* in Italian, lit. 'to break the balls', which is often used intransitively to convey the same pragmatic intention in a less colorful fashion). As is to be expected, the intuitions of native speakers pertaining to the grammaticality of the implicit object construction are influenced by the "semantic cluster" of meanings each verb activates in their minds, which thus lead to somewhat different judgments in English and Italian. This interpretation is consistent with Fillmore (1969, p. 100), who argued that "with polysemous transitive verbs, in other words verbs with several different senses, it is rather certain types of the senses and not the predicates *per se* that permit leaving out the object".

9.3 Final remarks

9.3.1 Reproducing Medina

In Section 9.2 I provided a full account of the inner workings and results yielded by the full Stochastic Optimality Theoretic models of object drop in English and in Italian. I motivated the choice to delve into the details of these two models, instead of any of the 16 other ones (collected in Table 9.1 and explained in Section 9.1.1), by making reference to the overall performance of the 18 models (see Section 9.1.3). However, a question was left unanswered. Since the models hereby discussed were prompted by the original study by Medina (2007), it would indeed be interesting to compare her results with the ones I obtained in the comparable models among my 18 ones, namely, the basic 3-predictor models. In this Section I will first compare Medina's model to my basic model quantifying semantic selectivity with Resnik's SPS, since it is the same measure of semantic selectivity Medina employed, and then I will comment on the two PISA-based models.
As I observed in Section 9.1.3, it is impossible, unfortunately, to provide a thorough, statistically-motivated comparison of my models against Medina's model, because she chose to evaluate her model's performance by means of the Summed Squared Error[18] of the predicted ratings she obtained in the model, instead of relying on more robust measures such as the (adjusted) R^2. Thus, I can only compare the parameters of the three linear functions and probabilities of an implicit object output for each aspectual type by Medina (2007, pp. 143–144) with the same kind of results I obtained with my own basic model(s).
Medina's original results, which I reproduced graphically and discussed in Section 5.5.5 (γ and δ parameters in Figure 5.5, probability of object drop in Figure 5.6), are reported again here in Figure 9.7 Figure 9.8.

18: The Summed Squared Error of a model increases by increasing the number of datapoints to predict, similarly to how non-adjusted R squared increases when increasing the number of predictors in the model. Therefore, such measures are heavily influenced by factors external to the quality itself of the model. Moreover, a model can have higher SSE than another even if its adjusted R squared is lower —e.g., my full Behavioral PISA-based model of English has an adjusted R squared equal to 0.468 and a SSE of 4.966, but the same model for Italian yields a 0.455 adjusted R squared and a 4.516 SSE.

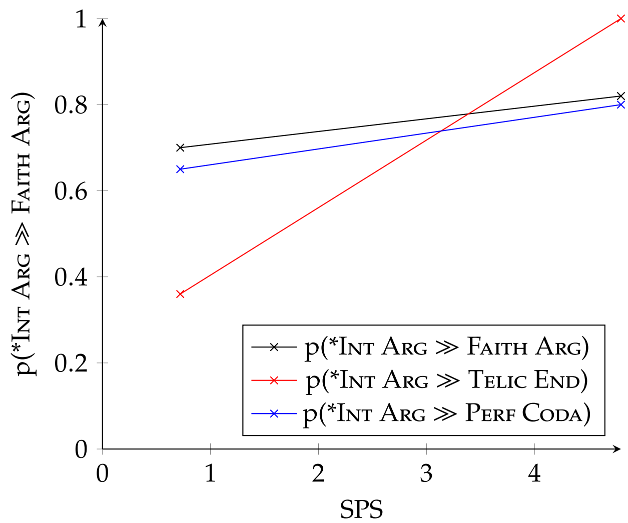

Figure 9.7: Graphical representation of the values of γ and δ in Medina's model (reproduction of Figure 5.5).

The same results I obtained in my basic model with semantic selectivity computed via Resnik's SPS are reported here in Figure 9.9 and Figure 9.10.

As is made evident by comparing Figure 9.7 and Figure 9.9, my SPS-based basic model fails to reproduce Medina's findings. In Medina's model

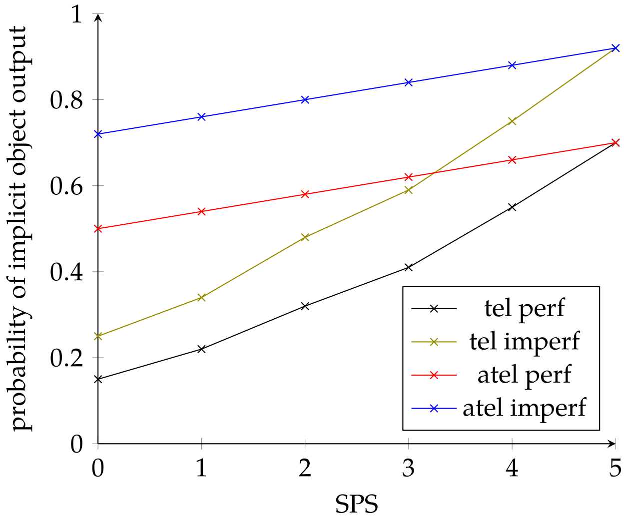

Figure 9.8: Representation of the relationship between semantic selectivity and the probability of an implicit object output in Medina's model, based on computed γ and δ values (reproduction of Figure 5.6).

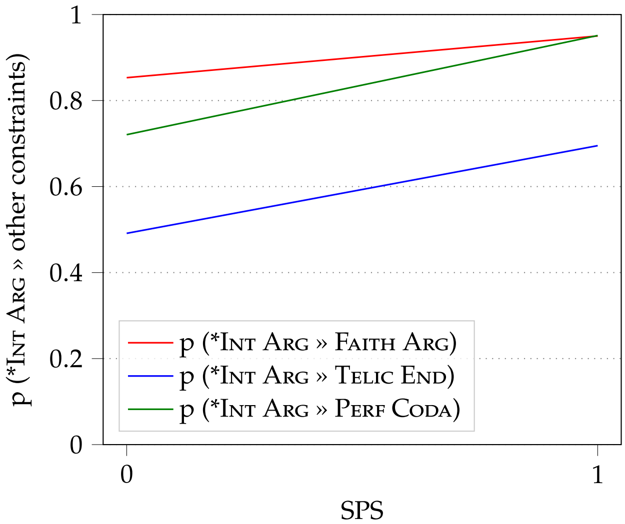

Figure 9.9: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Behavioral PISA (English full model).

there is a prominent interaction between the probability of *Int Arg outranking Telic End and the probability of it outranking the two other constraints at play (i.e., Faith Arg and Perf Coda), while this interaction is absent from my SPS-based model. In particular, the relative re-ranking probabilities are ordered the same way in both models when Resnik's SPS (raw in Medina's plot, normalized in mine) is very low, i.e., *Int Arg is most likely to outrank Faith Arg, then Perf Coda, then Telic End. However, while in my model this relation also holds for high values of SPS, in Medina's model *Int Arg becomes more likely to outrank Telic End than other constraints for mid-to-high values of SPS.

This state of affairs is reflected in the probability of licensing an implicit object output for each separate aspectual type, reported here in Figure 9.9 for Medina's model and in Figure 9.10 for my SPS-based basic model. Indeed, while in both models the object-dropping probabilities for the four aspectual inputs are ordered the same way for low-SPS verbs (atelic imperfective $\gg$ atelic perfective $\gg$ telic imperfective $\gg$ telic perfective), in Medina's model telic imperfective inputs are actually more likely to drop their object than atelic perfective inputs for mid-to-high-SPS verbs. In my model these probabilities vary according to SPS, but they never interact with one another.

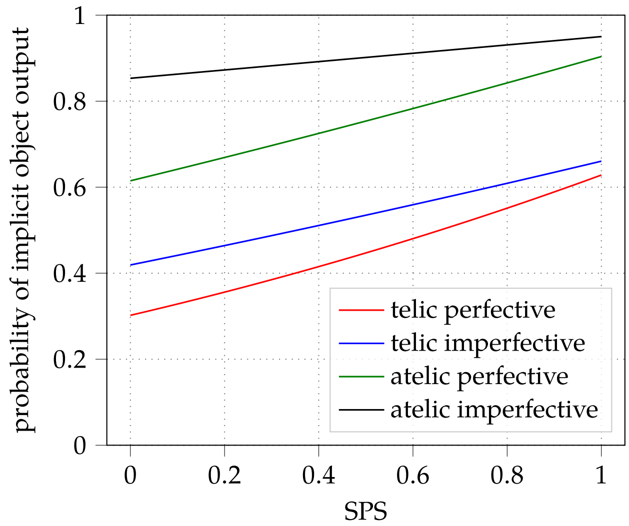

Figure 9.10: Probability of an implicit object output for each aspectual type, as a function of Behavioral PISA (English full model).

However, it is interesting to note that I reproduced Medina's findings in my basic model using Computational PISA and, a bit worse, in my basic model using Behavioral PISA. I will not report here the four figures to avoid cluttering these pages, but the interested reader can compare Medina's figures with my own ones relative to the Computational PISA model (Figure E.3 and Figure E.4) and to the Behavioral PISA model (Figure E.5 and Figure E.6), collected in Appendix E. Since Medina considered a different set of verb than I did, and recruited 15 different participants than the 30 ones who participated in my experiment, it is possible to conclude that the grammar of English with respect to object drop defined by Medina is indeed true to the way native speakers of English re-rank the constraints to judge the grammaticality of object drop in their language. Crucially, Medina's results are not an artifact of the specific verbs she picked (the same as in Resnik (1993, 1996), for evident computational reasons). Why are her findings reproduced quite closely by my PISA-based basic models but not by my SPS-based basic model, which after all uses the very same measure of semantic selectivity Medina used? I would motivate these results by making reference to the shortcomings of Resnik's SPS, which I discussed extensively in Section 6.1, and in particular to its need for both a taxonomy (such as WordNet) and for a corpus (to extract the frequencies needed in the computation, as in Section 6.1.1). It is thus unsurprising that a model making use of a taxonomy-based measure yields results of fleeting reproducibility, even more so considering that the corpus upon which Resnik and Medina based their calculations (the Brown corpus) is much smaller than the one I employed to obtain Computational PISA scores (the ukWaC corpus).

9.3.2 On regression models

I would finally like to echo Medina's (2007) concerns about the possibility of using a statistical regression model as a linguistically-informed model of language. In her thesis (Medina 2007, pp. 132–133), she concluded that her Stochastic Optimality Theoretic model shares the property of additivity with linear regression models, since the former models the probability of an implicit object output as the sum of the probabilities of the relevant constraint re-rankings, while the latter models it as the sum of weighted variables. The main difference between the two kinds of model lies in the fact that the linguistic model keeps the input, the constraints, and the constraint re-rankings explicit, while in a regression model they are collapsed into weighted variables. In general, Pater, Potts, and Bhatt (2006) observe that harmonic grammars[19] translate into linear systems of equations and, thus, are solvable as such.

I second Medina's conclusions, based on the results I obtained in my own study (refer back to Chapter 8 for the linear regressions and to the current Chapter for the linguistic models). In particular, I am going to compare the linear mixed-effects model[20] in Table 8.3 for English and in Table 8.6 for Italian with the results of my full linguistic model of object drop in Figure 9.4 for English and in Figure 9.5 for Italian. The mixed models clearly capture the main aspects of the linguistic models, e.g., the prominent role of telicity and perfectivity in jointly determining the grammaticality of the implicit object construction and the statistically less relevant effect of manner specification and iterativity. The role of Behavioral PISA determines another major divide between the two kinds of model, since in the regression it is assigned a weight just like any other predictor, be it continuous or binary, while in the Stochastic Optimality Theoretic model it is used as the independent variable in the computation of separate re-ranking functions for each constraint at play.

It is worth noting that Medina's model makes explicit use of a tenet of regression models, i.e., the requirement for the model to minimize the Summed Squared Error, which the reader can find in Medina's formulation on Page 86 (and again in Section 9.2.2) and in a more mathematically intense fashion in Bates et al. (2015, p. 13) with regards to linear mixed-effects models. In a sense, a theoretical and computational method bridging the gap between linear regressions, which are linguistically naive, and Medina's model, which is not thoroughly defined as a linguistically-informed regression model, is Linear Optimality Theory (Keller 2000, 2006), a stochastic variant of Optimality Theory which represents the constraint rankings as numerical weights and has the grammaticality of any linguistic structure be proportional to the sum of the weights of the constraints it violates. While both Keller's Linear Optimality Theory and Medina's variant of Stochastic Optimality Theory rely on the Least Square Estimation algorithm to estimate parameters, the two models differ significantly in their implementation because Linear Optimality Theory defines grammaticality in terms of the relative harmony of two candidates in the same candidate set, while Medina's model evaluates the grammaticality of null-object candidates across candidate sets (for reasons explained in Section 4.1.4).

19: That is to say, grammars selecting the most harmonic output among the candidates generated on the basis of the input, where "harmony" is the satisfaction of a set of weighted constraints. Refer to Chapter 4 for more details.

20: Which, I will remember, is simply a linear regression model taking both fixed and random effects into account.

Conclusions

Conclusions and open questions 10

10.1 Final comments

10.1.1 Recap of main findings

I presented my models of the implicit indefinite object construction in English and in Italian in Chapter 9, together with a discussion of their performance and results. Here I will provide a short summary of the main findings of the two full models, namely, the five-predictor models I computed using Behavioral PISA (introduced in Section 6.1.3) to quantify semantic selectivity, as explained in Section 9.1.1.

In general, the models perform comparably well, explaining almost half of the variance in the data (adjusted R^2 is 0.468 for the English model and 0.455 for the Italian model, as reported in Section 9.1.3). As I argued in Chapter 9, the better performance of the English model with respect to the Italian model may depend on semantic differences between the target verbs included in the behavioral experiments, and also on a more clear-cut role the predictors play in English than in Italian. These results, especially in the light of the fact that the full models improve on the performance of the reduced models (namely, Medina's three-predictor model and my own four-predictor model not considering manner specification[1]), are indeed encouraging. However, even the full models fall short of explaining all the variance in the data, demonstrating that there is still room for improvement. In Section 10.2, I will propose some ideas to expand upon these models and improve their computation in future research.

Let us look more closely at the constraint re-rankings and subsequent grammaticality predictions of object drop in the two models. Both in English and in Italian, the probability of *Int Arg outranking Telic End[2] varies strongly depending on Behavioral PISA, so that the curves described by the functions associated to this probability in the two languages are the steepest among all the curves associated to the re-ranking probabilities (as shown in Section 9.2.3). However, while in English the probability of *Int Arg outranking Telic End is directly proportional to semantic selectivity, in Italian the relation is one of inverse proportionality, for reasons discussed while commenting Figure 9.3.

Moreover, while in both languages the probability of *Int Arg outranking Telic End varies greatly depending on Behavioral PISA, there are differences with respect to the other predictors in the two languages. In particular, in English there is an interaction between the functions associated with the re-ranking probabilities of Telic End and Perf Coda, because telic imperfective verbs are more likely to drop their object than atelic perfective verbs for high Behavioral PISA values, while the opposite holds for Behavioral PISA scores lower than 0.8, approximately. In Italian, instead, there is an interaction between the function associated to the re-ranking probability of Mann-Spec Arg and the functions associated

1: As discussed in Section 9.1.1, the four-predictor models include Medina's three predictors and iterativity, while manner specification is only added in the full five-predictor models.

2: Required to have grammatical implicit indefinite objects with telic verbs, as explained in Chapter 5 and Chapter 9.

to the re-ranking probabilities of all the other constraints[3] . Thus, both models show a main effect of telicity on the probability that the object-less use of a transitive verb is considered grammatical, but the second most relevant factor in the model is perfectivity in English and manner specification in Italian.

3: The interaction with NON-ITER ARG is trivial, because it happens when Behavioral PISA is 1, i.e., the maximum value.

In addition to the adjusted R^2 values and the main role played by telicity, the two models are also comparable for:

- the range of predicted object-dropping probabilities (30-100% in English, 30-90% in Italian);
- the relevance of semantic selectivity in determining the grammaticality of object drop (with re-ranking probabilities that are always directly proportional to Behavioral PISA, with the exception of TELIC END in Italian);
- the fact that the predictors perform consistently with theoretical literature on object drop (refer to Chapter 2, Chapter 3, and Chapter 6).

Indeed, in both models atelic imperfective iterative manner-specified verbs are the most likely to drop their object (between 80% and 90%), while telic perfective non-iterative manner-unspecified verbs are the least likely (between 30% and 40%). Moreover, atelic verbs are more likely to occur with implicit objects than telic verbs, imperfective verbs more than perfective verbs, iterative verbs more than non-iterative verbs, and manner-unspecified verbs more than manner-specified verbs, as expected.

Even though semantic selectivity plays an active role in both models, the range of predicted grammaticality across different input types is not the same in English and in Italian. Indeed, while it is comparable for low-Behavioral PISA verbs in the two languages (ranging from 30% to 80% in English and from 30% to 90% in Italian), it is much narrower and higher in English (between 90% and 100%) than in Italian (between 40% and 90%), as shown and discussed in Section 9.2.4.

10.1.2 Comments on iterativity and manner specification

In this dissertation, I modeled the implicit indefinite object construction following Medina's (2007) steps. A major element of novelty I added is the inclusion of two novel constraints in the model, i.e., NON-ITER ARG and MANN-SPEC ARG (refer to Chapter 9), which are based on the role played by iterativity and manner specification (refer to Chapter 3), respectively, in facilitating indefinite object drop.

As I argued in Section 9.1.3, the addition of these two new constraints proved to be beneficial to the performance of the model when applied both to English and to Italian data. Indeed, the full five-predictor models explain the variance in the data better than the three- and the four-predictor models regardless of the chosen measure of semantic selectivity (Resnik's SPS, Computational PISA, or Behavioral PISA). However, the four-predictor models (including iterativity in addition to Medina's telicity, perfectivity, and semantic selectivity) do not perform better than the three-predictor models, with basically identical adjusted R^2 values in English and slightly smaller adjusted R^2 values in Italian. Taken together, these results mean that iterativity alone is not a sufficient addition to Medina's model (rather, it makes the model needlessly more complicated, since it does not explain more variance in the data), but models including iterativity and manner specification together have a stronger explanatory power.

The lower performance of the models including iterativity without manner specification echoes observations drawn in Section 9.2.3 relative to the probability of *Int Arg outranking Non-Iter Arg, that was shown to be very high both in a relative sense (since it is the highest among all the re-ranking probabilities) and in an absolute sense (92-100% in English, 96-100% in Italian). Since Non-Iter Arg is vacuously satisfied by iterative inputs, and varies almost imperceptibly according to semantic selectivity with non-iterative inputs (for which it is an active constraint), it stands to reason that it has no noticeable effect on the predicted grammaticality of object drop.

The same analysis would also explain the significant effect of the full models, where manner specification is also included among the predictors. In particular, in Section 9.1.3 I showed that the addition of manner specification determines a much stronger qualitative leap in the full models of Italian than in English, where the increase in the performance of the models is rather modest. Once again, these results are related to the probability of *Int Arg outranking the relevant constraint, i.e., Mann-Spec Arg. As shown in Section 9.2.3, the curve described by the function associated to this re-ranking probability is quite steep in Italian and it intersects all the other curves, while in English it is not steep at all and it has no interactions with the other curves. Thus, manner specification in Italian interacts in meaningful ways with semantic selectivity (as shown by the steepness of the curve) and the other binary predictors of object drop, making it an important factor in an expanded model of indefinite object drop. In English it has an effect too, but less evident.

10.1.3 Comments on semantic selectivity

In addition to the presence of additional constraints in the models, the other dimension of variation highlighted in Table 9.1 in Section 9.1.1 is the measure used to quantify semantic selectivity in the models. In particular, I computed three families of models, each based on Resnik's SPS (as in Medina's original model), on Computational PISA, or on Behavioral PISA. In Section 9.1.3, I provided adjusted R^2 scores of each model in each family, showing that in English SPS-based models are the worst-performing, while PISA-based models are noticeably better (with Behavioral PISA being better than Computational PISA). In Italian, instead, Computational PISA-based models are the worst-performing, followed by SPS-based models and, lastly, by Behavioral PISA-based models.

In Chapter 9, I provided a possible explanation of these facts, together with the correlations between semantic selectivity and average Likert grammaticality ratings presented in Section 8.2.1 and Section 8.3.1, by making reference to the way each measure of semantic selectivity is defined and computed (as detailed in Section 6.1). Let us consider the different performance of the three families of models in the light of Section 6.1.4, where I evaluated the correlations between the three measures of semantic selectivity both in English and in Italian. I argued that the English facts (SPS-based models being worse than PISA-based models, and SPS correlating poorly with PISAs while Behavioral PISA and Computational PISA correlate well with each other) may depend on the nature of these measures, considering that both PISA measures are based on the computation of pairwise similarity scores (distributional cosine similarity for Computational PISA, Likert-scale human judgments of similarity for Behavioral PISA), while SPS suffers from all the problems of taxonomy-based measures, as discussed in Section 6.1. The Italian picture is different, since in this case Computational PISA-based models perform worse than SPS-based models. Interestingly, as shown in Section 8.3.1, Computational PISA in Italian correlates very well with SPS (even better than with Behavioral PISA). I take this to mean that even after the manual cleansing I performed to purge any artifacts from the corpus data (recounted and motivated on Page 96), the itWaC corpus, upon which I based the computation of SPS and Computational PISA relative to Italian, has a stronger effect on the resulting scores than the ukWaC corpus, which I used to model the computational measures of semantic selectivity in English.

However, there may also be some undesirable side effect generated by the choice of ukWaC for English, given that I was able to reproduce Medina's findings relative to indefinite object drop in English with my three-predictor PISA-based models, but not with SPS, which is the measure she used (refer to Section 9.3.1). This may depend either on the corpus of choice or on the set of target verbs (or on both), but I am confident I can take the verbs off the suspect list because I made sure to include strictly transitive, as much as possible monosemous, verbs in my set (refer to Section 7.2), while Medina used the same verbs included in Resnik's (1993) original computation of SPS, a set also including verbs such as *to do, to get, to have*. Thus, I conclude that I was not able to reproduce Medina's model using SPS because of the corpus I used. It would be possible to test this hypothesis by computing again Medina's model using her verb set and ukWaC, and my verb set and the corpus Resnik (and thus Medina) based his computation on. To conclude, I observe that both in English and in Italian the best-performing models are based on Behavioral PISA. This result does not surprise at all, since this measure, being based on human similarity judgments, can be considered a benchmark model for semantic selectivity.

10.1.4 Is Optimality Theory the optimal choice?

In Chapter 4, I provided several arguments in favor of the use of a probabilistic model of the implicit indefinite object construction, whose grammaticality cannot be reduced to a simple matter of being binarily acceptable or unacceptable. Indeed, different transitive verbs allow indefinite object drop to varying degrees based on their semantic selectivity, and any given transitive verb may show different degrees of grammaticality when used intransitively based on other semantic, aspectual, and pragmatic features. Thus, models where an object-less candidate output can either be the winner or not (i.e., not accounting for gradience), such as Yankes's (2021 [2022]) standard Optimality Theoretic model of indefinite object drop in English, are doomed to only account for overly simplified observations about this phenomenon.

Therefore, probabilistic models of grammar are needed to account for the full range of grammaticality shown by the implicit indefinite object construction. Having evaluated several such models with respect to their applicability to the problem at hand, I argued in particular in favor of Stochastic Optimality Theory, since it has appealing properties the other models under examination lack. For instance, Harmonic Grammar was argued by Kuhn (2002) to be mathematically too powerful to be respectful of linguistic theory, since constraint weights can be re-adjusted until they yield results compatible with the observed typology. In a linguistically motivated model this should be done via modifications of the constraint set, rather than via numerical tweaks, and for this exact reason standard Optimality Theory (insufficient, as previously argued, for studies on indefinite object drop) was created on the basis of Harmonic Grammar. Keller's (1997) Extended Optimality Theory is a first step towards modeling gradient grammaticality, since it expresses the grammaticality of candidates as a ranking, instead of having a single winner and several, equally ungrammatical, losers. However, these rankings are not comparable across candidate sets (since they only have ordinal meaning), and gradient grammaticality judgments are only used to provide the candidates with a ranking.

Stochastic Optimality Theory, "the best motivated and most thoroughly probabilistic extension to Optimality Theory" according to Manning (2003, p. 25), defines the gradient grammaticality of a candidate as a function of the number and type of constraint re-rankings returning it as optimal. Crucially, the simultaneous optimality of multiple candidates is achieved by having constraints float on a continuous numerical scale. In its classic definition, presented in Section 4.2.4, Stochastic Optimality Theory assigns to each constraint a probability distribution (a normal distribution of values defined by a Gaussian curve), and the overlap between two such distributions determines the probability of the two constraints being re-ranked with respect to one another. Such a framework is indeed a well-thought-out choice to model the gradient grammaticality of indefinite object drop. However, classic Stochastic Optimality Theory cannot account for the effect of semantic selectivity straight out of the box, if one intends to measure it properly as a gradient property[4] of verbs. Indeed, it is not possible to force gradient semantic selectivity into a binary constraint without losing significant explanatory power. For this reason, Medina (2007) devised a variant of Stochastic Optimality Theory (presented in Chapter 5) where the constraints re-rank based on the input verb's semantic selectivity rather than on the overlap of *normal* probability distributions. In particular, the probability of each faithfulness constraint, penalizing object drop, re-ranking with respect to the *Int Arg markedness constraint, favoring object drop, is expressed as a *linear* function of the verb's semantic selectivity.

To conclude, Medina's variant of Stochastic Optimality Theory, which I adopted in this dissertation, appears to be the optimal choice to model indefinite object drop, since it leverages the many attractive features of the classic version of the theory while also accounting for the effect of gradient semantic selectivity on the gradient grammaticality of null objects. More accurately, it is the optimal choice for a model of indefinite object drop *so far*. As I will sketch out in Section 10.2.2, future research may reveal that non-linear functions can yield better predictions than linear functions.

4: As discussed in Chapter 5 and Chapter 6, Medina (2007) modeled it via Resnik's (1993) Selectional Preference Strength taxonomy-based measure, while in this dissertation I also used Computational PISA (a novel similarity-based measure stemming from distributional semantics I contributed to define in Cappelli and Lenci (2020)) and Behavioral PISA (another similarity-based measure obtained via human judgments).

10.2 Future directions

10.2.1 Expanding the model

Additional predictors Among the linguistic factors facilitating indefinite object drop presented in Chapter 3, I only picked five to serve as predictors in my models (detailed in Chapter 6), namely: semantic selectivity, telicity, perfectivity (all three from Medina's (2007) original model), manner specification, and iterativity (two novel additions), for reasons detailed in Section 3.5. The way these predictors are implemented in the (linear) Stochastic Optimality Theoretic models was explained in Chapter 5 relative to Medina's model and in Chapter 9 relative to mine. Future research may expand upon my models in the same way I expanded upon Medina's, namely, by introducing additional predictors in the model based on theoretical literature. A relevant area of interest, which I only brushed against by including iterativity (a broadly-intended pragmatic factor) in my model, is that of pragmatic and discourse factors (refer to Section 3.3). Since out-of-context utterances only happen in laboratory environments, research on pragmatic and discourse factors will provide much more ecological data to studies on indefinite object drop. However, this should not be intended as a potshot at models based on no-context stimuli, such as Medina's and the ones proposed in this dissertation. Indeed, given that the same semantic and aspectual factors determine indefinite object drop both in context-rich and in no-context utterances, it makes sense to model these factors first and to add contextual factors later on. Moreover, a word of caution is needed regarding the possible addition of pragmatics to experiment on indefinite object drop, since sufficient context may make virtually any object recoverable and, thus, any transitive verb acceptable when used intransitively. Thus, experiments including intra- and extra-linguistic contexts will have to be carefully calibrated in order to quantify the exact role of each type of context, and to avoid having context-external factors confound the experiment.

Given that indefinite object drop challenges prototypical transitivity, it would also be interesting to include in the model the neglected parameters described by Hopper and Thompson (1980) (refer back to Table 2.1 in Section 2.1), in particular affirmation, mode, agency (strictly related to Agent affectedness, discussed on Page 25 and in Section 3.1.2), and affectedness of the object (tackled in Chapter 2).

Corpus frequencies As shown in Boersma (2004) and Boersma and Hayes (2001), Stochastic Optimality Theory can be used to model corpus data just as well as grammaticality judgments, namely, via the evaluation of constraints that get re-ranked along a continuous numerical scale (refer to Chapter 4). Rather than trivially duplicating the results I obtained and discussed in this dissertation, new models of corpus frequencies are sure to shed a different light on the implicit indefinite object construction. As I anticipated in Section 3.4, neither Resnik (1993, 1996) nor Medina (2007) found a precise correlation between corpus frequencies and the gradient grammaticality judgments provided in behavioral experiments about implicit indefinite objects.

Indeed, linguistic research has long since shown that there is no clear-cut correspondence between ratings elicited from native speakers and corpus frequencies (Manning 2003). In particular, it is often the case that low-frequency utterances (or other linguistic items) receive mid-to-high acceptability judgments in behavioral experiments (Bader and Häussler 2010; Bermel and Knittl 2012; Boersma 2004; Keller and Asudeh 2002; Kempen and Harbusch 2005). There is also no strict relation between the relative grammaticality of a linguistic structure with respect to another and their relative corpus frequencies, since, for instance, Bader and Häussler (2010, pp. 315–316) report that they found no pairs of syntactic structures in their study where a member of the pair was judged as more grammatical than the other but occurred with a smaller frequency in the corpus, while Boersma (2004) argues in favor of the opposite. Moreover, Bader and Häussler (2010) experimental results show both a "ceiling mismatch"[5] (meaning that two syntactic structures may be judged as maximally grammatical, but occur with different frequencies in the corpus) and a "floor mismatch" (meaning that two syntactic structures may never or almost-never occur in the corpus, but receive different acceptability judgments).

5: The authors also observe that this is not a measurement artifact due to the use of a capped scale, such as binary or 7-point Likert ratings, because it is also found with Magnitude Estimation ratings, which are open-ended both at the top and at the bottom.

A common worry about linguistic research based on corpus material is that frequencies are less reliable than human judgments because there is no way to control language production as one controls an experimental design. This line of reasoning would surely curb easy enthusiasm about the replication of the current study to model corpus frequencies of indefinite object drop, if Kempen and Harbusch (2008) and Schütze (1996 [2016]) did not observe that acceptability ratings are too "contaminated by performance factors", that is to say, biased by other tasks the raters perform in addition to the one they are explicitly asked to carry out (e.g., they judge the similarity between the target sentence and the one they consider its "ideal delivery" paraphrase). Thus, if linguistics gladly relies on acceptability judgments (and, oftentimes, the results of one's own introspection), provided they are based on a rigorous experimental design, there should be no qualms about modeling corpus frequencies, provided they are interpreted in the light of the factors possibly influencing them. In general, given that no experimental method is error-free, it is good practice to compare the results obtained with different methods. In the specific case of studies on indefinite object drop, there may be a trade-off between the analysis of easily computable[6] frequencies extracted from non-manipulable[7] corpus utterances, and the analysis of acceptability judgments provided by human subjects on easily manipulable experimental stimuli.

6: Provided the corpus is annotated in such a way as to make data extraction easy, of course.

7: Also, possibly under-representative of the language one intends to study, given that even in corpora that are not genre-specific it is difficult to obtain complete coverage of language uses and contexts.

Modeling corpus frequencies of indefinite null objects using the very same model(s) defined in this dissertation may present additional challenges if compared to modeling acceptability ratings, since it is impossible to manipulate aspectual and discourse factors in a corpus study as in a behavioral experiment. However, the possible absence (or very low frequency) of a given object-less verb in a given aspect may well be considered an interesting, modelable datum in itself, provided one adjusts the model to account for such findings. Alternatively, it would be possible to design a production experiment to design an *ad-hoc* corpus to model the frequency of indefinite null objects in controlled speech or writing. It is also important to note that it would be possible, if not even easy, to include discourse and world-knowledge context (somewhat ancillary to semantics and aspectual factors in this dissertation) in a model of object drop based on frequencies extracted from a large corpus, given that these null objects appear in sentences which are part of larger documents with explicit context information. Moreover, a corpus study of object drop may provide an answer to a question foreshadowed by Kempen and Harbusch (2005) and Medina (2007) (refer to Section 3.4), namely, whether a "production threshold" exists blocking mid-to-low grammaticality structures from ever being uttered and, if so, which numerical value has to be assigned to this threshold.

Other implicit complements of verbs Direct object of optionally transitive verbs are far from being the only NP complements of verbs participating in syntactic omissions. For instance, the literature mentions:

- Agents of passives, as in *The ship was sunk* $\varnothing_{Agent}$ (Bhatt and Pancheva 2017; Lasersohn 1993; Ruppenhofer and Michaelis 2014);
- Source (Gillon 2006), Goal (Lasersohn 1993; Ruppenhofer and Michaelis 2014), and Path (Recanati 2002) locative phrases occurring with motion verbs[8] , as in *Bill left* $\varnothing_{Source}$, *Hilary arrived* $\varnothing_{Goal}$, and *The cow jumped over* $\varnothing_{Path}$;
- Themes of reflexive (*Peter shaved (himself)*) and reciprocal (*Mary and Peter divorced (from each other)*) predicates (Németh 2014);
- Recipients of three-argument verbs, as in *The mayor donated $300* $\varnothing_{Recipient}$ (Ruppenhofer 2005);
- Instruments, as in *The executioner beheaded the prisoner* $\varnothing_{Instrument}$ (Koenig, Mauner, and Bienvenue 2002, 2003; Koenig, Mauner, Bienvenue, and Conklin 2007; Rissman 2010; Rissman and Rawlins 2017; Rissman, Rawlins, and B. Landau 2015).

8: Interestingly, Gillon (2006, pp. 9–12) and Ruppenhofer, Gorinski, and Sporleder (2011, p. 333) observe that in some pairs of near-synonym verbs, such as *to leave, to vacate* and *to arrive, to reach*, only one member of the pair allows for the omission of the locative phrase. This is consistent with the literature on the role of manner specification in argument omission (refer to Section 3.1.3).

Among all these *syntactically* optional complements of verbs, Instruments stand out because they can also be *semantically* optional. Indeed, Koenig, Mauner, and Bienvenue (2002, 2003) and Koenig, Mauner, Bienvenue, and Conklin (2007) divided Instrument-taking verbs into two classes, i.e., the Require-Instrument class (*to chop, to slice, to write*) and the Allow-Instrument class (*to eat, to break, to open*). In Cappelli and Lenci (2020), I computed Computational PISA scores of Instrument-taking English verbs (together with transitive verbs, as discussed in Section 6.1.2), showing that this measure of semantic selectivity can reliably tell apart Require- and Allow-Instrument verbs. I argue that this is a promising starting point in a possible computational model of the factors regulating the syntactic optionality of Instruments, given the insight this method provided in the study of indefinite object drop.
Modeling Instruments, as well as the other implicit complements listed in this Section, will provide useful information to further theoretical and experimental research on syntactic optionality.

Indefinite object drop diachronically As argued in Section 3.4, with reference to Glass (2020), Goldberg (2001, 2005b), and Lorenzetti (2008), verbs appearing in generic contexts with a habitual interpretation (e.g., *Pat drinks; Pat smokes; Chris sings; Sam bakes*, from Goldberg (2001, p. 518)) are likely to participate in the implicit indefinite object construction, due to Goldberg's principle of Omission under Low Discourse Prominence. Diachronically, the frequent use of transitive verbs in such contexts

probably led to the grammaticalization of their intransitive use in episodic contexts as well, often with a specialized meaning (e.g. 'to drink alcohol' in *Pat drinks*, 'to bake pastries' in *Sam bakes*).
A model of indefinite object drop in historical and contemporary texts (a gap in the literature first observed by Goldberg (2001)) would substantiate this hypothesis and shed some more light on the mechanisms regulating the role of semantic and aspectual predictors in addition to discourse factors. Moreover, diachronic change also affects semantic selectivity (which, as discussed in Chapter 2, Chapter 3, and Chapter 6 is a major predictor of indefinite object drop) and other facets of verb meaning. For instance, verbs may undergo semantic changes expanding the range of their meaning (e.g., Vulgar Latin **adripare* 'to reach the shore' gave rise to Italian *arrivare* 'to arrive'), shifting it from a concrete to a metaphorical interpretation (e.g., *to broadcast* originally meant 'to cast seeds widely' on a field, while the rise of communication technologies in the 20th century shifted its meaning to 'spread a message or news widely'), or restricting it (e.g., Latin *cubare* 'to lie, to rest' became Italian *covare* 'to brood', referring to the lying act performed by egg-laying animals to nurse their eggs).
Several corpora of English and Italian are available for a diachronic study on indefinite object drop, each focusing on different text types within different time spans. Among the diachronic corpora of English[9] , relevant ones to inquire about null objects may be:

- the Helsinki Corpus of English Texts (Rissanen 1993), spanning over Old (V-XII centuries) to Early Modern (late 15th - late 17th centuries) English, and covering many different genres (such as chronicles, handbooks, laws, and Bible excerpts);
- the Corpus of Late Modern English Texts (De Smet 2005; De Smet et al. 2015), covering public domain British English texts from 1710 to 1920;
- the Corpus of Historical American English (Davies 2010, 2012), recently purged of inconsistent lemmas and malformed tokens (Alatrash et al. 2020), a large-scale (around 400 million words) corpus covering different genres (newspapers, fiction and non-fiction books, magazines) from 1810 to 2009.

9: Refer to Hilpert and Gries (2016) for an introduction to quantitative approaches to diachronic corpus linguistics mentioning several corpora of English.

As for Italian, relevant diachronic corpora may be:

- DiaCORIS (Onelli et al. 2006), a 100-million-word corpus of texts written between 1861 (year of the National Unification) and 1945 covering genres such as newspapers, fiction, and academic prose;
- the MIDIA corpus (Gaeta et al. 2013; Iacobini, De Rosa, and Schirato 2014), a 7-million-word corpus of documents written between the 13th and the 20th centuries;
- the OVI corpus (Larson, Artale, and Dotto 2005), a half-million word corpus of texts from the 12th to the 14th centuries;
- a diachronic corpus of newspaper articles published on "L'Unità" (Basile et al. 2020), the official newspaper of the Italian Communist Party from 1924 to the dissolution of the Party in 1991, published between 1924 and 2015.

While not assembled with linguistic research in mind, unlike the aforementioned corpora, the collection of over 1000 literary works written throughout the whole history of Italian literature offered in the Biblioteca Italiana Zanichelli[10] may provide useful insight as to indefinite object

10: Available as Windows-only software on a DVD distributed online at https://www.zanichelli.it/ricerca/prodotti/biblioteca-italiana-zanichelli.

drop in Italian literature.

In theory, the wider the time span covered by a given diachronic corpus, the better. A corpus ranging over several centuries of written language would indeed provide a broader perspective on the possible grammaticalization of null objects outside habitual contexts, rather than a corpus covering a shorter span of time. However, it may also be the case that changes in grammar happened much faster in the last century, when distant communication became possible, literacy was not an upper-class privilege anymore, and, in the last 30 years or so, English became the main language of the internet. As for Italian, Basile et al. (2020) observe that deep changes occurred in this language during the second half of the 20th century[11] . Thus, it is possible that use-dependent pressures towards the grammaticalization of null objects are more evident in corpora focusing on the last century, than on previous time periods. Moreover, subtle changes of this kind are surely more frequent and observable in corpora based on spoken language, or language written to be read shortly after (such as the one used in newspapers and other mass media). Thus, both broad diachronic corpora and 20th-century corpora may be of use to understand the history of indefinite object drop.

11: Without detracting, of course, from the several transitions the grammar of Italian underwent in the several centuries of its history.

Typologically different languages In this dissertation, I modeled the implicit indefinite object construction in English and Italian, two typologically close languages. As discussed in Chapter 9, several differences between the two emerged with respect to their licensing of indefinite null objects, meaning that this phenomenon depends on much finer-grained aspects than what typology alone would warrant. Nevertheless, a theory of grammar should not be a theory of English and English-like languages alone, and therefore a comprehensive model of indefinite object drop should consider a variety of typologically different languages.

As noted by Jackendoff (2003, p. 134), languages such as Korean and Japanese (to which we may add Chinese) allow for null arguments more easily than English, so that they would pose "no justification for distinguishing between obligatorily and optionally expressed semantic arguments". However, in the same paragraph he also argues in favor of the distinction between definite and indefinite null objects being a fully idiosyncratic lexical property of verbs or, at most, of semantic verb classes, a position which I argued against throughout Chapter 2 and Chapter 3. Thus, it is not to be excluded that languages such as Korean, Japanese, and Chinese may show different degrees of acceptability of indefinite object drop with different optionally transitive verbs, even if they are free in their Topic-drop-based licensing of definite null objects (just like English and typologically similar languages are, as argued on Page 13). Another aspect differentiating these languages from languages such as English and Italian, with respect to aspects playing a role in object drop, is their treatment of non-culminating accomplishments. Indeed, while in English the sentence **I burned it but it didn't burn* is utterly ungrammatical due to being contradictory, its Japanese equivalent (*moyashita keredo moenakatta*, as reported in Radden and Seto (2003, p. 236)) is grammatical because the verb *moyasu* 'to burn' is less focused on the result than its English counterpart —it is, thus, a non-culminating accomplishment. The existence of such verbs made Ikegami (1991) distinguish between DO-languages such as English, focusing on the Agent, and BECOME-languages such as Japanese, focusing on the process. Indeed, given the considerations provided in Section 2.4.2 relative to indefinite object drop as a mechanism driven by the need to focus on the activity rather than on its effects on the Patient, it stands to reason that BECOME-languages should allow object drop more often and in a wider variety of contexts than DO-languages. Another typologically different family, namely, Slavic languages, may shed light on the role played by grammatical aspect in the implicit indefinite object construction. As mentioned in Section 3.2.2, simplifying a very long tradition of studies in a way that will surely disgruntle many of those who fostered research in this area, in Slavic languages perfectivity is embedded in the lexicon rather than being expressed morphologically (as in English and Italian). To be more precise, in Slavic languages the opposition between so-called perfective and imperfective verbs actually derives from the encoding (or lack thereof) of telicity in the verb (Bertinetto 2001; Bertinetto and Delfitto 2000; Bertinetto and Lentovskaya 2012). In a diachronic perspective, discussed in Bertinetto and Lentovskaya (2012) relative to Russian, the loss of the overt aspectual markers in the passage between Old Slavonic to (Northern) Slavic languages gave rise to a syncretic system merging lexical aspect and grammatical aspect. The opposition between prefixed and simple verbs in Russian is interpreted as "unmistakable evidence" of the original distinction being one of lexical, not grammatical, aspect. Crucially, since so-called imperfective Slavic verbs can be used in perfective contexts, and given that so-called perfective verbs are always ungrammatical with null objects (Sopata 2016; Tsimpli and Papadopoulou 2006), experiments relative to indefinite object drop in Slavic languages (be they corpus-based or judgment-based) should only focus on the realization of implicit objects with imperfective transitive verbs. I take this state of affairs to mean that Perf Coda acts as a hard constraint in Slavic languages, being always re-ranked above *Int Arg for perfective inputs (refer to Chapter 5 and Chapter 9), instead of being a soft, re-rankable constraint as it is in English and Italian. Indeed, as discussed in Section 3.2.3, in these two languages telicity, perfectivity and tense are intertwined, so that the interpretation of one factor partially depends on the others. This explains why in my models telicity and (secondarily) perfectivity both play a gradient role in favoring object drop, depending on each verb's semantic selectivity, despite the behavioral experiments being carefully designed to isolate the effect of each factor at play. Based on previous considerations, I hypothesize that a model of indefinite object drop in Slavic languages would paint quite a different picture.

10.2.2 Different math

In this dissertation I followed Medina (2007) in providing a Stochastic Optimality Theoretic model of indefinite object drop (see Chapter 5) where the binary predictors described in Chapter 6 are used to define four faithfulness constraints re-ranking with respect to *Int Arg (a markedness constraint). Semantic selectivity, modeled along a continuous numerical scale, cannot give rise to a binary constraint itself. Instead, Medina implements it in the model by defining the probability of each faithfulness constraint re-ranking with respect to *Int Arg as a (linear) function of the semantic selectivity of the input verb, deviating from the Stochastic Optimality Theoretic norm of associating fixed normal curves to constraints (refer to Section 4.2.4). However, as Medina (2007, p. 110)

herself comments, there is no compelling reason why these re-ranking functions should be necessarily *linear* functions. Indeed, she argues that linear functions are "a reasonable place to begin to explore the relative contribution of Semantic Selectivity to the implicit object construction", and I followed in her steps to obtain models comparable to hers. Future research on indefinite object drop may benefit from employing non-linear functions, defining a more complex algorithm than the one used here to determine the best function and its parameters.
Going back to linear Stochastic Optimality Theoretic models of object drop, in Section 9.3.2 I commented on the differences between them and linear mixed-effects models, which are linear regression models including both fixed and random effects in the computation. Mixed models, by their very nature, are able to account both for the effect of the predictors of object drop on the grammaticality of indefinite null objects (the fixed effects), and for the effect of the source of random variability in the data, i.e., the target verbs and the participants to the experiment (the aptly-named random effects). Medina's model and my own, instead, are more like classic linear regression model in that they only account for fixed effects. I minimized any effect the human participants may have had on the results by modeling their normalized ratings (refer to Section 8.1.2), but this pre-processing adaptation of the Likert ratings is more of a quick fix than the kind of solid method I endorse for future research. Indeed, ideally the model should take raw ratings as input, and not only account for the different use the participants made of the Likert scale, but also quantify the amount of variance in the data depending on the participants alone. The same goes for random effects depending on the target verbs, which my models are not able to compute.

10.2.3 A follow-up on recoverability and prototypicality

In Section 2.3.2, I argued, with reference to the literature, that indefinite null objects refer to the prototypical Patients of a given transitive verb, as recovered by speakers via world knowledge. However, this prototypicality is not to be intended as a monolith. Rather, it depends on extra- and intra-linguistic context, as it is possible to argue based on Rice's (1988) examples in (1). In (1-a), the act of smoking is intended to refer to cigarettes, because they are the most common object of smoking in contemporary Western society, but different context (such as a mention to olden times, or a Middle-East setting) may induce a reading where the omitted object refers to a pipe or a waterpipe. In (1-b), the act of drinking is linked to alcohol assumption for reasons made clear in Chapter 2 and Section 3.1.2, but it would necessarily refer to some other prototype were the subject a toddler or a teetotaler, let alone a non-human participant. Similarly, John may be understood to drive a bike in (1-c) and to read a newspaper in (1-d), provided slight differences in the available context (e.g., a downtown location to drive to, or a different reading time).

(1) a. John smokes (cigarettes / *Marlboros / *a pipe / *SMOKING MATERIALS).
b. John drinks (alcohol / *gin / *water / *coffee / *LIQUIDS).
c. When he goes to Boston, John drives (a car / *a Toyota / *a motorcycle / *A VEHICLE).

d. Each afternoon, John reads (a book / *Ulysses / *the newspaper / *PRINTED MATTER).

Therefore, possibly enriching some additional models I envisioned in Section 10.2.1, it may be useful to design an experiment targeted at the prototypicality[12] intrinsic to object recoverability. I would imagine a cloze-test experiment where subjects have to fill in the gap in sentences such as *John smokes* ___, manipulating context as to have no-context sentences, common-sense contexts, and uncommon contexts. I hypothesise that there would be much greater agreement between participants relative to the fillers of common-context stimuli and no-context stimuli (where the context is inherently provided via world knowledge), than the agreement relative to uncommon-context stimuli (where it is difficult to imagine prototypicality).

12: This task tackles a different problem than the one this thesis focused on (i.e., modeling the grammaticality of object-less transitive verbs).

Appendix

Verbs used in the stimuli | A

This appendix collects the English and Italian verbs used in the stimuli of the behavioral experiment eliciting acceptability judgments about the implicit object construction, as explained in full detail in Section 7.2. These data are also available on my GitHub profile*. The stimuli used in the behavioral experiment are listed in the same GitHub repository and here in Appendix D.

A.1 Target verbs

A.1.1 Matching English and Italian verbs

English	Italian
behead	decapitare
break	rompere
build	costruire
chop	spaccare
clean	pulire
cook	cucinare
cut	tagliare
devour	divorare
doodle	scarabocchiare
drink	bere
eat	mangiare
embroider	ricamare
hum	canticchiare
kill	uccidere
knife	accoltellare
poison	avvelenare
polish	lucidare
pour	versare
sew	cucire
sign	firmare
sing	cantare
sip	sorseggiare
slice	affettare
smoke	fumare
steal	rubare
swig	trangugiare
teach	insegnare
wash	lavare
watch	guardare
write	scrivere

* https://github.com/giuliacappelli/dissertationData

A.1.2 English

verb	frequency	Zipf scores
behead	1674	2.9418
break	196609	5.0116
build	479945	5.3992
chop	15330	3.9036
clean	53629	4.4474
cook	36378	4.2789
cut	158274	4.9174
devour	3447	3.2555
doodle	350	2.2621
drink	56215	4.4679
eat	136063	4.8518
embroider	2689	3.1476
hum	2714	3.1516
kill	140951	4.8671
knife	494	2.4118
poison	4710	3.3910
polish	6360	3.5215
pour	26960	4.1487
sew	4141	3.3351
sign	168608	4.9449
sing	75238	4.5945
sip	3090	3.2080
slice	9389	3.6906
smoke	21213	4.0446
steal	41619	4.3373
swig	229	2.0779
teach	198500	5.0158
wash	41347	4.3345
watch	170952	4.9509
write	634329	5.5203

A.1.3 Italian

verb	frequency	Zipf scores
accoltellare	1933	3.0860
affettare	5539	3.5432
avvelenare	8732	3.7409
bere	58875	4.5697
cantare	77281	4.6879
canticchiare	2119	3.1259
costruire	282558	5.2509
cucinare	15135	3.9798
cucire	7029	3.6467
decapitare	3987	3.4004
divorare	10850	3.8352
firmare	126382	4.9015
fumare	28974	4.2618
guardare	358839	5.3547
insegnare	111884	4.8486

verb	frequency	Zipf scores
lavare	28971	4.2618
lucidare	2561	3.2082
mangiare	117137	4.8685
pulire	34400	4.3364
ricamare	4727	3.4744
rompere	65876	4.6185
rubare	38715	4.3877
scarabocchiare	563	2.5503
scrivere	855506	5.7320
sorseggiare	3145	3.2974
spaccare	14799	3.9700
tagliare	78147	4.6927
trangugiare	652	2.6140
uccidere	156043	4.9930
versare	91025	4.7590

A.2 Filler verbs

A.2.1 Matching English and Italian verbs

English	Italian
clap	applaudire
fast	digiunare
knock	bussare
laugh	ridere
limp	zoppicare
rest	riposarsi
scream	urlare
sleep	dormire
smile	sorridere
stagger	barcollare

B Behavioral PISA

This appendix collects the English and Italian stimuli for the experiments I ran to compute the Behavioral PISA scores. These data, together with the full results provided by the participants to the experiment, are also available on my GitHub profile[1] . The Behavioral PISA scores determined by these stimuli are listed in the same GitHub repository and here in Appendix C.3.
The code used to generate these stimuli and to compute the final scores is available on my GitHub profile in a dedicated repository[2] .

1: https://github.com/giuliacappelli/dissertationData

2: https://github.com/giuliacappelli/behavioralPISA

B.1 English stimuli

verb	noun 1	noun 2
behead	enemy	prisoner
behead	enemy	woman
behead	member	man
behead	prophet	family
behead	son	prisoner
behead	woman	member
break	cork	aircraft
break	fibre	doll
break	phone	flag
break	timber	drill
break	valve	cord
break	vegetation	handbag
build	bunkhouse	tarmac
build	cloud	arena
build	complexity	stall
build	foundation	intention
build	manufactory	repeater
build	radar	phone
chop	ginger	coriander
chop	leek	apricot
chop	nut	banana
chop	spice	bar
chop	sprig	dill
chop	tomato	peel
clean	book	bike
clean	boot	carpet
clean	lab	log
clean	lock	cloth
clean	square	area
clean	wound	printer
cook	cod	kipper
cook	drumstick	blackberry
cook	fries	fennel

verb	noun 1	noun 2
cook	mixture	vinegar
cook	plant	garlic
cook	side	dish
cut	box	content
cut	card	bullet
cut	hull	waffle
cut	nostril	shoelace
cut	pine	houseplant
cut	pine	plant
devour	animal	slug
devour	fish	cock
devour	goat	octopus
devour	pasture	cod
devour	people	hamburger
devour	sweet	aphid
doodle	animal	passage
doodle	character	design
doodle	idea	picture
doodle	image	design
doodle	picture	drawing
doodle	triangle	design
drink	allergen	sauce
drink	gel	drop
drink	infusion	syrup
drink	poison	supplies
drink	soda	coke
drink	wine	booze
eat	flesh	mollusc
eat	grub	food
eat	kinsman	supplement
eat	plant	grass
eat	serving	honey
eat	vegetable	protein
embroider	banner	panel
embroider	blouse	shirt
embroider	embellishment	strip
embroider	flag	crest
embroider	garment	tapestry
embroider	lace	shawl
hum	chorus	track
hum	line	word
hum	music	tune
hum	passage	tune
hum	selection	riff
hum	track	theme
kill	blacksmith	publisher
kill	flora	rainforest
kill	heathen	philistine
kill	merchant	sir
kill	officer	astrologer
kill	sister	pal
knife	bride	man

verb	**noun 1**	**noun 2**
knife	man	wife
knife	people	man
knife	pig	people
knife	pig	rival
knife	rival	people
poison	body	husband
poison	husband	pigeon
poison	mother	hamster
poison	stream	body
poison	supplies	meat
poison	water	brandy
polish	brass	silver
polish	electrode	lens
polish	fingernail	tongue
polish	frame	copper
polish	helmet	lamp
polish	phone	machine
pour	acid	rice
pour	alloy	sugar
pour	oil	substance
pour	sherry	champagne
pour	soup	beverage
pour	wine	soda
sew	cast	purse
sew	fabric	costume
sew	flower	seed
sew	label	tag
sew	mask	pattern
sew	work	tusk
sign	alliance	convention
sign	apprenticeship	application
sign	comment	document
sign	enquiry	survey
sign	memorial	theory
sign	receipt	bill
sing	choir	falsetto
sing	harmony	melody
sing	hymn	anthem
sing	pop	rap
sing	selection	repertoire
sing	song	carol
sip	beer	gin
sip	cocoa	mead
sip	concoction	beverage
sip	latte	espresso
sip	whiskey	vodka
sip	wine	vinegar
slice	chunk	piece
slice	clay	butter
slice	garlic	turnip
slice	image	element
slice	root	bulb

verb	noun 1	noun 2
slice	skin	joint
smoke	cannabis	mushroom
smoke	cocaine	crack
smoke	hemp	cocaine
smoke	packet	dose
smoke	pot	joint
smoke	tobacco	cigar
steal	chattel	kid
steal	nappy	bell
steal	raft	patent
steal	sugar	silver
steal	supply	videotape
steal	wave	bucket
swig	ale	brew
swig	brew	potion
swig	can	drop
swig	champagne	alcohol
swig	cider	brew
swig	pint	drink
teach	coding	recording
teach	industry	obedience
teach	mechanic	premise
teach	theology	prayer
teach	topic	information
teach	tutorial	recipe
wash	building	surface
wash	cotton	clothing
wash	fridge	microwave
wash	hob	sieve
wash	nursery	drain
wash	packaging	belongings
watch	channel	chore
watch	coding	blossom
watch	pizza	enhancement
watch	product	curling
watch	screen	curve
watch	skating	childhood
write	justification	dissertation
write	missive	opera
write	paragraph	essay
write	review	assertion
write	synopsis	analysis
write	topography	license

B.2 Italian stimuli

verb	noun 1	noun 2
accoltellare	amico	autore
accoltellare	cognato	agente
accoltellare	controllore	ispettore
accoltellare	marocchino	connazionale

verb	**noun 1**	**noun 2**
accoltellare	persona	tifoso
accoltellare	ragazzino	uomo
affettare	cipolla	carciofo
affettare	cipolla	motociclista
affettare	oliva	zucca
affettare	pane	carne
affettare	porro	cuore
affettare	prosciutto	salame
avvelenare	banana	donna
avvelenare	batterio	animale
avvelenare	gatto	polmone
avvelenare	insetto	figlio
avvelenare	sposo	criminale
avvelenare	tessuto	sangue
bere	bibita	aperitivo
bere	carburante	fluido
bere	latte	bibita
bere	sostanza	medicina
bere	the	succo
bere	veleno	tisana
cantare	brano	storia
cantare	preghiera	messa
cantare	rima	epopea
cantare	singolo	coro
cantare	storiella	leggenda
cantare	testo	orazione
canticchiare	canzone	motivo
canticchiare	canzonetta	musica
canticchiare	hit	ritmo
canticchiare	incantesimo	poesia
canticchiare	strofa	ninna
canticchiare	tema	canzoncina
costruire	carro	scettro
costruire	cinta	istituto
costruire	gommone	diga
costruire	lanterna	croce
costruire	roccaforte	imballaggio
costruire	scaletta	gabinetto
cucinare	boccone	banchetto
cucinare	carne	filetto
cucinare	cibo	ripieno
cucinare	grigliata	prelibatezza
cucinare	pesce	tonno
cucinare	zampa	pollo
cucire	bambola	pezza
cucire	biancheria	reggiseno
cucire	divisa	corredo
cucire	filo	stendardo
cucire	margine	cappotto
cucire	veste	giacca
decapitare	banda	membro
decapitare	corpo	donna

verb	noun 1	noun 2
decapitare	famiglia	cittadino
decapitare	ostaggio	americano
decapitare	uomo	mostro
decapitare	venditore	morto
divorare	abitante	uomo
divorare	albero	radice
divorare	capo	cuore
divorare	figlio	nemico
divorare	raccolto	pascolo
divorare	topo	malcapitato
firmare	appello	proposta
firmare	email	revoca
firmare	introduzione	trattativa
firmare	querela	annuncio
firmare	rivista	libretto
firmare	sceneggiatura	nota
fumare	erba	canapa
fumare	filtro	sigaretta
fumare	fumo	gomma
fumare	oppio	cocaina
fumare	pacchetto	sigaro
fumare	paglia	incenso
guardare	asino	parabola
guardare	emisfero	basket
guardare	freccia	gioco
guardare	preferenza	copertura
guardare	programma	mantello
guardare	risorsa	veduta
insegnare	derivata	meccanismo
insegnare	regia	psicologia
insegnare	regola	strategia
insegnare	stile	sistema
insegnare	teologia	chirurgia
insegnare	tradizione	culto
lavare	barattolo	scala
lavare	barba	guancia
lavare	calzino	lana
lavare	cisterna	stalla
lavare	costume	prodotto
lavare	seno	collo
lucidare	cerchio	catena
lucidare	legno	superficie
lucidare	moto	vetro
lucidare	mouse	posata
lucidare	pistola	armatura
lucidare	telaio	mobile
mangiare	banana	patata
mangiare	carota	antipasto
mangiare	formaggio	pasto
mangiare	panettone	topo
mangiare	trota	corvo
mangiare	zucchina	verdura

verb	noun 1	noun 2
pulire	ambiente	foresta
pulire	cenere	cisterna
pulire	ingresso	passaggio
pulire	macchia	sangue
pulire	pennello	occhio
pulire	radice	terra
ricamare	carattere	logo
ricamare	contorno	trama
ricamare	coperta	lenzuolo
ricamare	logo	opera
ricamare	sfondo	copertina
ricamare	velo	mantello
rompere	ginocchio	polso
rompere	maglia	filo
rompere	marmo	cristallo
rompere	rivestimento	telecomando
rompere	sottomarino	barattolo
rompere	vetrina	recinto
rubare	album	bottiglia
rubare	archivio	ala
rubare	bicchiere	elemosina
rubare	fetta	milione
rubare	manufatto	informazione
rubare	progetto	migliaio
scarabocchiare	appunto	pensiero
scarabocchiare	foglio	quaderno
scarabocchiare	libro	carta
scarabocchiare	nome	firma
scarabocchiare	pagina	disegno
scarabocchiare	pagina	frase
scrivere	capoverso	giallo
scrivere	domanda	intervento
scrivere	email	intestazione
scrivere	narrazione	monologo
scrivere	saggio	opera
scrivere	sottotitolo	questionario
sorseggiare	acqua	the
sorseggiare	amaro	grappa
sorseggiare	bevanda	brodo
sorseggiare	bibita	composto
sorseggiare	nettare	bibita
sorseggiare	sorso	succo
spaccare	anello	tazza
spaccare	balla	ceppo
spaccare	cemento	pallina
spaccare	chitarra	ghiaccio
spaccare	finestra	tavolo
spaccare	ponte	asse
tagliare	bacca	ala
tagliare	carne	pomodoro
tagliare	giacca	grembiule
tagliare	legame	ormeggio

verb	**noun 1**	**noun 2**
tagliare	legatura	lepre
tagliare	torta	pero
trangugiare	birra	cappuccino
trangugiare	boccone	pasto
trangugiare	cena	minestra
trangugiare	pizza	panettone
trangugiare	pozione	bevanda
trangugiare	pozione	veleno
uccidere	appartenente	equipaggio
uccidere	bastardo	parassita
uccidere	carne	muffa
uccidere	prigioniero	religioso
uccidere	segretario	israeliano
uccidere	tifoso	pregiudicato
versare	carburante	soluzione
versare	inchiostro	vernice
versare	polpa	colla
versare	residuo	prodotto
versare	succo	sciroppo
versare	veleno	bile

Verb-dependent predictors of object drop

C

This appendix collects the semantic selectivity scores, telicity feature, and manner specification of each English and Italian verb of interest. These data are also available on my GitHub profile*. The stimuli used in the Behavioral PISA experiment are listed in the same GitHub repository and here in Appendix B.
The scripts used to compute the semantic selectivity scores are available on GitHub here† for Behavioral PISA and here‡ for Resnik's SPS and Computational PISA.

C.1 Resnik's SPS scores

C.1.1 English

verb	value
behead	5.41951429109159
break	1.6793306972271176
build	1.0444365443830292
chop	3.438656094289755
clean	1.4409848073158913
cook	3.3312852756730833
cut	2.022539900328247
devour	3.2412081031584195
doodle	3.6684580190582055
drink	3.0733070499044426
eat	2.9666860095348273
embroider	4.676864553416348
hum	3.3698350402403054
kill	3.6808371012416345
knife	0
poison	4.279653745683352
polish	2.296068943151817
pour	2.7724980420575305
sew	2.833979840980959
sign	1.5293797499155062
sing	2.642727020790552
sip	3.3363686807974804
slice	3.5758202575701117
smoke	5.229644753166431
steal	1.8768283472927347
swig	2.8662435216484727
teach	1.6629568673957216
wash	1.8464744191425093

* https://github.com/giuliacappelli/dissertationData
† https://github.com/giuliacappelli/behavioralPISA
‡ https://github.com/ellepannitto/PISA

verb	value
watch	2.3395754700199665
write	2.239589940207992

C.1.2 Italian

verb	value
accoltellare	3.890054972777093
affettare	4.171865123063593
avvelenare	2.787554655603073
bere	3.0177796446611502
cantare	2.362766569982537
canticchiare	2.8050620027560345
costruire	0.7875067657846383
cucinare	3.71382520298977
cucire	3.0016072216417995
decapitare	3.5252440727001564
divorare	3.091973944651725
firmare	1.3281521941994174
fumare	3.1709384554039595
guardare	0.7338795569704666
insegnare	2.0545094983958077
lavare	2.638599775346278
lucidare	2.497534218645052
mangiare	3.042929318310735
pulire	1.5520676487000922
ricamare	2.3608646075158055
rompere	2.244588368992655
rubare	1.7026267535600168
scarabocchiare	2.439636158138708
scrivere	1.8178364364916102
sorseggiare	2.923331966414907
spaccare	2.575603929377065
tagliare	2.5737633430353495
trangugiare	2.9701312620677065
uccidere	3.230657237621352
versare	3.1375585181456485

C.2 Computational PISA scores

C.2.1 English

verb	value
behead	0.2665590109890108
break	0.1672398532974804
build	0.14259006152197046
chop	0.250876263920088
clean	0.1701836735398874
cook	0.31215897191016945
cut	0.18466699313592272
devour	0.18176671598824243

verb	value
doodle	0.22332499999999994
drink	0.2691043316531047
eat	0.21487358165625178
embroider	0.2410483984674331
hum	0.37887409183108933
kill	0.18732028544846346
knife	0.2619611111111111
poison	0.1850712748876043
polish	0.20191239599454402
pour	0.24835831137116315
sew	0.22664585367147644
sign	0.20564031937795324
sing	0.3141075723850854
sip	0.37466938967136204
slice	0.25784927497097393
smoke	0.2608563767121237
steal	0.15744235473302504
swig	0.4285276995305164
teach	0.18015142467099565
wash	0.19233420757156694
watch	0.15607432511726602
write	0.20329813297609786

C.2.2 Italian

verb	value
accoltellare	0.3079597819593789
affettare	0.46626135248993866
avvelenare	0.22779692116092695
bere	0.3234907686569406
cantare	0.3205721869843004
canticchiare	0.37776087472201647
costruire	0.17287654835175703
cucinare	0.4350583056740942
cucire	0.29180270654834617
decapitare	0.2499725954666117
divorare	0.256135131008793
firmare	0.20475470353876601
fumare	0.3025446626439931
guardare	0.16237313468251074
insegnare	0.1959324326755706
lavare	0.2624123234829927
lucidare	0.26015092597804695
mangiare	0.29584647583123846
pulire	0.21079758204545984
ricamare	0.22094988730807189
rompere	0.23340662139696455
rubare	0.18519251247467175
scarabocchiare	0.26218470824949713
scrivere	0.20467814921049665
sorseggiare	0.4077989860406971
spaccare	0.2667972822445544

verb	value
tagliare	0.24275482460977424
trangugiare	0.3105916099773242
uccidere	0.2148306899420238
versare	0.33561646678650436

C.3 Behavioral PISA scores

C.3.1 English

verb	value
behead	0.29055436999022555
break	0.12276836214623506
build	0.21762785567742657
chop	0.34440540442572826
clean	0.17292497041218488
cook	0.33185906739493143
cut	0.2871232879595378
devour	0.23809859297188518
doodle	0.5125231723017851
drink	0.47495494373653285
eat	0.3859282310020324
embroider	0.48347837937372223
hum	0.4604600976785315
kill	0.33436898147194566
knife	0.4052205744674658
poison	0.2141225350611907
polish	0.3294412144676424
pour	0.33449987699263023
sew	0.46365623305930415
sign	0.4745322863537071
sing	0.7019831506086257
sip	0.5707923807178529
slice	0.4286860332412498
smoke	0.6132821453149185
steal	0.09412671500790444
swig	0.5701913824589526
teach	0.4225079780942123
wash	0.4027036029805784
watch	0.12497312517080204
write	0.3776801508375687

C.3.2 Italian

verb	value
accoltellare	0.5886637800215203
affettare	0.5885110211271879
avvelenare	0.2953994090101937
bere	0.7764391268585936
cantare	0.7486440386129157
canticchiare	0.7521474918969137

verb	value
costruire	0.24323348807407993
cucinare	0.7835973989735566
cucire	0.6138248655755957
decapitare	0.49898373191487244
divorare	0.5727069147760088
firmare	0.4967896690588007
fumare	0.6936365143704156
guardare	0.19041248697926685
insegnare	0.5121774320314921
lavare	0.5202997492919127
lucidare	0.46611648860671046
mangiare	0.6240360335935481
pulire	0.5684277904658784
ricamare	0.6457925988588612
rompere	0.5456384270051742
rubare	0.22732823115800935
scarabocchiare	0.774598373604205
scrivere	0.575161142941159
sorseggiare	0.7606638064253161
spaccare	0.35800971947008
tagliare	0.4437135136473041
trangugiare	0.758835738433465
uccidere	0.41402756288438985
versare	0.6275661376317657

C.4 Telicity

C.4.1 English

verb	telicity	in-for	progressive	conjunction
behead	yes	yes	yes	yes
break	yes	yes	yes	yes
build	yes	yes	yes	yes
chop	yes	yes	yes	yes
clean	no	yes	no	no
cook	no	yes	no	no
cut	no	no	no	no
devour	yes	yes	yes	no
doodle	no	no	no	no
drink	no	no	no	no
eat	no	no	no	no
embroider	no	no	no	no
hum	no	no	no	no
kill	yes	yes	yes	yes
knife	yes	no	yes	yes
poison	yes	yes	yes	yes
polish	no	no	no	no
pour	no	no	no	no
sew	no	no	no	no
sign	yes	yes	yes	yes
sing	no	no	no	no

verb	telicity	in-for	progressive	conjunction
sip	no	no	no	yes
slice	yes	yes	yes	yes
smoke	no	no	no	no
steal	yes	yes	yes	yes
swig	yes	yes	yes	no
teach	no	no	yes	no
wash	no	no	no	no
watch	no	no	no	no
write	no	no	no	no

C.4.2 Italian

verb	telicity	in-for	progressive	conjunction
accoltellare	yes	no	yes	yes
affettare	yes	yes	yes	yes
avvelenare	yes	yes	yes	yes
bere	no	no	no	no
cantare	no	no	no	no
canticchiare	no	no	no	no
costruire	yes	yes	yes	yes
cucinare	no	yes	no	no
cucire	no	no	no	no
decapitare	yes	yes	yes	yes
divorare	yes	yes	yes	no
firmare	yes	yes	yes	yes
fumare	no	no	no	no
guardare	no	no	no	no
insegnare	no	no	yes	no
lavare	no	no	no	no
lucidare	no	no	no	no
mangiare	no	no	no	no
pulire	no	yes	no	no
ricamare	no	no	no	no
rompere	yes	yes	yes	yes
rubare	yes	yes	yes	yes
scarabocchiare	no	no	no	no
scrivere	no	no	no	no
sorseggiare	no	no	no	yes
spaccare	yes	yes	yes	yes
tagliare	no	no	no	no
trangugiare	yes	yes	yes	no
uccidere	yes	yes	yes	yes
versare	no	no	no	no

C.5 Manner specification

C.5.1 English

verb	manner specification	a manner of...
behead	yes	kill
break	no	
build	no	
chop	yes	cut
clean	no	
cook	no	
cut	no	
devour	yes	eat
doodle	yes	write
drink	no	
eat	no	
embroider	yes	sew
hum	yes	sing
kill	no	
knife	yes	cut
poison	yes	kill
polish	yes	clean
pour	no	
sew	no	
sign	yes	write
sing	no	
sip	yes	drink
slice	yes	cut
smoke	no	
steal	no	
swig	yes	drink
teach	no	
wash	yes	clean
watch	no	
write	no	

C.5.2 Italian

verb	manner specification	a manner of...
accoltellare	yes	tagliare
affettare	yes	tagliare
avvelenare	yes	uccidere
bere	no	
cantare	no	
canticchiare	yes	cantare
costruire	no	
cucinare	no	
cucire	no	
decapitare	yes	uccidere
divorare	yes	mangiare
firmare	yes	scrivere
fumare	no	
guardare	no	
insegnare	no	
lavare	yes	pulire
lucidare	yes	pulire
mangiare	no	

verb	**manner specification**	**a manner of...**
pulire	no	
ricamare	yes	cucire
rompere	no	
rubare	no	
scarabocchiare	yes	scrivere
scrivere	no	
sorseggiare	yes	bere
spaccare	yes	tagliare
tagliare	no	
trangugiare	yes	bere
uccidere	no	
versare	no	

Sentence stimuli for the behavioral experiment about object drop

D

This appendix collects the sentence stimuli for the behavioral experiment about object drop in English and Italian, each featuring 30 transitive verbs of interest and 10 intransitive filler verbs (listed in Appendix A). These data are also available on my GitHub profile*, together with the individual judgments provided by 30 participants per language on a 7-point Likert scale.
The Python scripts I wrote to analyse the results and create a Stochastic Optimality Theoretic model of object drop are available on my Github profile† in a dedicated repository.

D.1 English

D.1.1 Target sentences

verb	dObj	perfective	iterative	sentence
behead	no	no	no	Clara was beheading.
behead	no	no	yes	Tom was beheading again.
behead	no	yes	no	Betty had beheaded.
behead	no	yes	yes	Sarah had beheaded again.
break	no	no	no	Clara was breaking.
break	no	no	yes	Clara was breaking again.
break	no	yes	no	Sarah had broken.
break	no	yes	yes	Sam had broken again.
build	no	no	no	Sarah was building.
build	no	no	yes	Paul was building again.
build	no	yes	no	Paul had built.
build	no	yes	yes	Paul had built again.
chop	no	no	no	Sean was chopping.
chop	no	no	yes	Clara was chopping again.
chop	no	yes	no	Sean had chopped.
chop	no	yes	yes	Paul had chopped again.
clean	no	no	no	Tom was cleaning.
clean	no	no	yes	Sarah was cleaning again.
clean	no	yes	no	John had cleaned.
clean	no	yes	yes	Diana had cleaned again.
cook	no	no	no	Diana was cooking.
cook	no	no	yes	John was cooking again.
cook	no	yes	no	Betty had cooked.
cook	no	yes	yes	Clara had cooked again.
cut	no	no	no	John was cutting.
cut	no	no	yes	Sam was cutting again.
cut	no	yes	no	Sean had cut.

* https://github.com/giuliacappelli/dissertationData
† https://github.com/giuliacappelli/MedinaStochasticOptimalityTheory

verb	dObj	perfective	iterative	sentence
cut	no	yes	yes	Tom had cut again.
devour	no	no	no	Diana was devouring.
devour	no	no	yes	Sam was devouring again.
devour	no	yes	no	Paul had devoured.
devour	no	yes	yes	Betty had devoured again.
doodle	no	no	no	Sarah was doodling.
doodle	no	no	yes	Tom was doodling again.
doodle	no	yes	no	Paul had doodled.
doodle	no	yes	yes	Sam had doodled again.
drink	no	no	no	Clara was drinking.
drink	no	no	yes	Tom was drinking again.
drink	no	yes	no	Sarah had drunk.
drink	no	yes	yes	Sarah had drunk again.
eat	no	no	no	Betty was eating.
eat	no	no	yes	Paul was eating again.
eat	no	yes	no	Betty had eaten.
eat	no	yes	yes	Sean had eaten again.
embroider	no	no	no	Clara was embroidering.
embroider	no	no	yes	Sam was embroidering again.
embroider	no	yes	no	Sarah had embroidered.
embroider	no	yes	yes	Paul had embroidered again.
hum	no	no	no	Sam was humming.
hum	no	no	yes	Tom was humming again.
hum	no	yes	no	Mary had hummed.
hum	no	yes	yes	Diana had hummed again.
kill	no	no	no	Diana was killing.
kill	no	no	yes	Sam was killing again.
kill	no	yes	no	Sarah had killed.
kill	no	yes	yes	Tom had killed again.
knife	no	no	no	Tom was knifing.
knife	no	no	yes	John was knifing again.
knife	no	yes	no	Diana had knifed.
knife	no	yes	yes	Sarah had knifed again.
poison	no	no	no	Mary was poisoning.
poison	no	no	yes	Paul was poisoning again.
poison	no	yes	no	Paul had poisoned.
poison	no	yes	yes	Sean had poisoned again.
polish	no	no	no	Mary was polishing.
polish	no	no	yes	Mary was polishing again.
polish	no	yes	no	Paul had polished.
polish	no	yes	yes	Betty had polished again.
pour	no	no	no	Tom was pouring.
pour	no	no	yes	Betty was pouring again.
pour	no	yes	no	Tom had poured.
pour	no	yes	yes	Mary had poured again.
sew	no	no	no	Betty was sewing.
sew	no	no	yes	Tom was sewing again.
sew	no	yes	no	Betty had sewn.
sew	no	yes	yes	Mary had sewn again.
sign	no	no	no	Betty was signing.
sign	no	no	yes	Sarah was signing again.
sign	no	yes	no	John had signed.

verb	dObj	perfective	iterative	sentence
sign	no	yes	yes	Mary had signed again.
sing	no	no	no	Sam was singing.
sing	no	no	yes	Paul was singing again.
sing	no	yes	no	Sam had sung.
sing	no	yes	yes	Paul had sung again.
sip	no	no	no	Paul was sipping.
sip	no	no	yes	John was sipping again.
sip	no	yes	no	Betty had sipped.
sip	no	yes	yes	Sean had sipped again.
slice	no	no	no	John was slicing.
slice	no	no	yes	Diana was slicing again.
slice	no	yes	no	Tom had sliced.
slice	no	yes	yes	John had sliced again.
smoke	no	no	no	Sam was smoking.
smoke	no	no	yes	Diana was smoking again.
smoke	no	yes	no	Diana had smoked.
smoke	no	yes	yes	Tom had smoked again.
steal	no	no	no	Betty was stealing.
steal	no	no	yes	John was stealing again.
steal	no	yes	no	Diana had stolen.
steal	no	yes	yes	Sam had stolen again.
swig	no	no	no	Tom was swigging.
swig	no	no	yes	Sam was swigging again.
swig	no	yes	no	Paul had swigged.
swig	no	yes	yes	Sean had swigged again.
teach	no	no	no	Sarah was teaching.
teach	no	no	yes	Sarah was teaching again.
teach	no	yes	no	Clara had taught.
teach	no	yes	yes	Sean had taught again.
wash	no	no	no	Paul was washing.
wash	no	no	yes	John was washing again.
wash	no	yes	no	Tom had washed.
wash	no	yes	yes	Mary had washed again.
watch	no	no	no	Betty was watching.
watch	no	no	yes	Sarah was watching again.
watch	no	yes	no	Sam had watched.
watch	no	yes	yes	Diana had watched again.
write	no	no	no	John was writing.
write	no	no	yes	Diana was writing again.
write	no	yes	no	Paul had written.
write	no	yes	yes	Tom had written again.

D.1.2 Control sentences

verb	dObj	perfective	iterative	sentence
behead	yes	no	no	Tom was beheading a prisoner.
behead	yes	no	yes	John was beheading a prisoner again.
behead	yes	yes	no	Tom had beheaded a prisoner.
behead	yes	yes	yes	Diana had beheaded a prisoner again.
break	yes	no	no	Sarah was breaking a vase.

verb	dObj	perfective	iterative	sentence
break	yes	no	yes	Mary was breaking a vase again.
break	yes	yes	no	Paul had broken a vase.
break	yes	yes	yes	Tom had broken a vase again.
build	yes	no	no	Betty was building a house.
build	yes	no	yes	Paul was building a house again.
build	yes	yes	no	Betty had built a house.
build	yes	yes	yes	Diana had built a house again.
chop	yes	no	no	Sarah was chopping a log.
chop	yes	no	yes	Diana was chopping a log again.
chop	yes	yes	no	Tom had chopped a log.
chop	yes	yes	yes	Tom had chopped a log again.
clean	yes	no	no	Tom was cleaning a table.
clean	yes	no	yes	Sam was cleaning a table again.
clean	yes	yes	no	Sam had cleaned a table.
clean	yes	yes	yes	Tom had cleaned a table again.
cook	yes	no	no	Diana was cooking dinner.
cook	yes	no	yes	John was cooking dinner again.
cook	yes	yes	no	Mary had cooked dinner.
cook	yes	yes	yes	John had cooked dinner again.
cut	yes	no	no	Sam was cutting some paper.
cut	yes	no	yes	Betty was cutting some paper again.
cut	yes	yes	no	Mary had cut some paper.
cut	yes	yes	yes	Paul had cut some paper again.
devour	yes	no	no	Paul was devouring a roasted chicken.
devour	yes	no	yes	Sarah was devouring a roasted chicken again.
devour	yes	yes	no	Mary had devoured a roasted chicken.
devour	yes	yes	yes	Betty had devoured a roasted chicken again.
doodle	yes	no	no	Paul was doodling a stick man.
doodle	yes	no	yes	Tom was doodling a stick man again.
doodle	yes	yes	no	Sean had doodled a stick man.
doodle	yes	yes	yes	Tom had doodled a stick man again.
drink	yes	no	no	Mary was drinking juice.
drink	yes	no	yes	Tom was drinking juice again.
drink	yes	yes	no	Sarah had drunk juice.
drink	yes	yes	yes	Sam had drunk juice again.
eat	yes	no	no	John was eating pizza.
eat	yes	no	yes	Sean was eating pizza again.
eat	yes	yes	no	Sean had eaten pizza.
eat	yes	yes	yes	John had eaten pizza again.
embroider	yes	no	no	Sam was embroidering a tapestry.
embroider	yes	no	yes	Sam was embroidering a tapestry again.
embroider	yes	yes	no	Clara had embroidered a tapestry.
embroider	yes	yes	yes	Tom had embroidered a tapestry again.

verb	dObj	perfective	iterative	sentence
hum	yes	no	no	Sam was humming a lullaby.
hum	yes	no	yes	Sean was humming a lullaby again.
hum	yes	yes	no	Sam had hummed a lullaby.
hum	yes	yes	yes	John had hummed a lullaby again.
kill	yes	no	no	Sarah was killing pests.
kill	yes	no	yes	Sam was killing pests again.
kill	yes	yes	no	Betty had killed pests.
kill	yes	yes	yes	Diana had killed pests again.
knife	yes	no	no	Betty was knifing a man.
knife	yes	no	yes	Sean was knifing a man again.
knife	yes	yes	no	Clara had knifed a man.
knife	yes	yes	yes	Tom had knifed a man again.
poison	yes	no	no	Paul was poisoning a plant.
poison	yes	no	yes	John was poisoning a plant again.
poison	yes	yes	no	Diana had poisoned a plant.
poison	yes	yes	yes	Paul had poisoned a plant again.
polish	yes	no	no	Sean was polishing a sword.
polish	yes	no	yes	Betty was polishing a sword again.
polish	yes	yes	no	Sam had polished a sword.
polish	yes	yes	yes	John had polished a sword again.
pour	yes	no	no	Sam was pouring wine.
pour	yes	no	yes	Betty was pouring wine again.
pour	yes	yes	no	Betty had poured wine.
pour	yes	yes	yes	Sam had poured wine again.
sew	yes	no	no	John was sewing a curtain.
sew	yes	no	yes	Mary was sewing a curtain again.
sew	yes	yes	no	Betty had sewn a curtain.
sew	yes	yes	yes	John had sewn a curtain again.
sign	yes	no	no	Sam was signing a paper.
sign	yes	no	yes	Betty was signing a paper again.
sign	yes	yes	no	Sean had signed a paper.
sign	yes	yes	yes	Sarah had signed a paper again.
sing	yes	no	no	Paul was singing a carol.
sing	yes	no	yes	Sarah was singing a carol again.
sing	yes	yes	no	Mary had sung a carol.
sing	yes	yes	yes	Mary had sung a carol again.
sip	yes	no	no	Sean was sipping water.
sip	yes	no	yes	Mary was sipping water again.
sip	yes	yes	no	Tom had sipped water.
sip	yes	yes	yes	John had sipped water again.
slice	yes	no	no	Sarah was slicing some pie.
slice	yes	no	yes	Mary was slicing some pie again.
slice	yes	yes	no	Sean had sliced some pie.
slice	yes	yes	yes	Mary had sliced some pie again.
smoke	yes	no	no	Sarah was smoking a cigarette.
smoke	yes	no	yes	Paul was smoking a cigarette again.
smoke	yes	yes	no	John had smoked a cigarette.
smoke	yes	yes	yes	Diana had smoked a cigarette again.
steal	yes	no	no	Sean was stealing money.

verb	dObj	perfective	iterative	sentence
steal	yes	no	yes	Mary was stealing money again.
steal	yes	yes	no	Tom had stolen money.
steal	yes	yes	yes	Betty had stolen money again.
swig	yes	no	no	Sarah was swigging beer.
swig	yes	no	yes	Clara was swigging beer again.
swig	yes	yes	no	John had swigged beer.
swig	yes	yes	yes	Mary had swigged beer again.
teach	yes	no	no	John was teaching linguistics.
teach	yes	no	yes	Sarah was teaching linguistics again.
teach	yes	yes	no	Mary had taught linguistics.
teach	yes	yes	yes	Sean had taught linguistics again.
wash	yes	no	no	Clara was washing a car.
wash	yes	no	yes	Tom was washing a car again.
wash	yes	yes	no	Betty had washed a car.
wash	yes	yes	yes	Sam had washed a car again.
watch	yes	no	no	Sean was watching a movie.
watch	yes	no	yes	Sarah was watching a movie again.
watch	yes	yes	no	Clara had watched a movie.
watch	yes	yes	yes	Tom had watched a movie again.
write	yes	no	no	Paul was writing a letter.
write	yes	no	yes	Betty was writing a letter again.
write	yes	yes	no	Sam had written a letter.
write	yes	yes	yes	Clara had written a letter again.

D.1.3 Filler sentences

verb	dObj	perfective	iterative	sentence
clap	no	no	no	Diana was clapping.
clap	no	no	yes	Paul was clapping again.
clap	no	yes	no	Mary had clapped.
clap	no	yes	yes	Sarah had clapped again.
clap	yes	no	no	Mary was clapping a show.
clap	yes	no	yes	Sean was clapping a show again.
clap	yes	yes	no	Tom had clapped a show.
clap	yes	yes	yes	Mary had clapped a show again.
fast	no	no	no	Betty was fasting.
fast	no	no	yes	Sarah was fasting again.
fast	no	yes	no	Sean had fasted.
fast	no	yes	yes	John had fasted again.
fast	yes	no	no	Clara was fasting sushi.
fast	yes	no	yes	Sam was fasting sushi again.
fast	yes	yes	no	Diana had fasted sushi.
fast	yes	yes	yes	Sarah had fasted sushi again.
knock	no	no	no	Betty was knocking.
knock	no	no	yes	Sarah was knocking again.
knock	no	yes	no	Mary had knocked.
knock	no	yes	yes	Sean had knocked again.
knock	yes	no	no	Clara was knocking a door.
knock	yes	no	yes	Sarah was knocking a door again.
knock	yes	yes	no	Diana had knocked a door.

verb	dObj	perfective	iterative	sentence
knock	yes	yes	yes	Betty had knocked a door again.
laugh	no	no	no	Sam was laughing.
laugh	no	no	yes	Tom was laughing again.
laugh	no	yes	no	Sam had laughed.
laugh	no	yes	yes	John had laughed again.
laugh	yes	no	no	John was laughing a joke.
laugh	yes	no	yes	John was laughing a joke again.
laugh	yes	yes	no	Betty had laughed a joke.
laugh	yes	yes	yes	Sean had laughed a joke again.
limp	no	no	no	Clara was limping.
limp	no	no	yes	Betty was limping again.
limp	no	yes	no	John had limped.
limp	no	yes	yes	Mary had limped again.
limp	yes	no	no	Mary was limping a road.
limp	yes	no	yes	Paul was limping a road again.
limp	yes	yes	no	Paul had limped a road.
limp	yes	yes	yes	John had limped a road again.
rest	no	no	no	Sean was resting.
rest	no	no	yes	Betty was resting again.
rest	no	yes	no	John had rested.
rest	no	yes	yes	Sam had rested again.
rest	yes	no	no	Betty was resting a bed.
rest	yes	no	yes	Mary was resting a bed again.
rest	yes	yes	no	Betty had rested a bed.
rest	yes	yes	yes	Tom had rested a bed again.
scream	no	no	no	Diana was screaming.
scream	no	no	yes	Tom was screaming again.
scream	no	yes	no	Paul had screamed.
scream	no	yes	yes	Sarah had screamed again.
scream	yes	no	no	Paul was screaming a spider.
scream	yes	no	yes	Paul was screaming a spider again.
scream	yes	yes	no	Tom had screamed a spider.
scream	yes	yes	yes	Tom had screamed a spider again.
sleep	no	no	no	Diana was sleeping.
sleep	no	no	yes	Clara was sleeping again.
sleep	no	yes	no	John had slept.
sleep	no	yes	yes	Sam had slept again.
sleep	yes	no	no	Sarah was sleeping a pillow.
sleep	yes	no	yes	Mary was sleeping a pillow again.
sleep	yes	yes	no	Sarah had slept a pillow.
sleep	yes	yes	yes	Mary had slept a pillow again.
smile	no	no	no	Clara was smiling.
smile	no	no	yes	Sean was smiling again.
smile	no	yes	no	Sam had smiled.
smile	no	yes	yes	Paul had smiled again.
smile	yes	no	no	Mary was smiling a friend.
smile	yes	no	yes	Sean was smiling a friend again.
smile	yes	yes	no	Clara had smiled a friend.
smile	yes	yes	yes	John had smiled a friend again.
stagger	no	no	no	Clara was staggering.
stagger	no	no	yes	Clara was staggering again.

verb	dObj	perfective	iterative	sentence
stagger	no	yes	no	Mary had staggered.
stagger	no	yes	yes	Betty had staggered again.
stagger	yes	no	no	Betty was staggering the pavement.
stagger	yes	no	yes	Tom was staggering the pavement again.
stagger	yes	yes	no	Paul had staggered the pavement.
stagger	yes	yes	yes	John had staggered the pavement again.

D.2 Italian

D.2.1 Target sentences

verb	dObj	perfective	iterative	sentence
accoltellare	no	no	no	Marta stava accoltellando.
accoltellare	no	no	yes	Sara stava accoltellando di nuovo.
accoltellare	no	yes	no	Sara aveva accoltellato.
accoltellare	no	yes	yes	Paolo aveva accoltellato di nuovo.
affettare	no	no	no	Gianni stava affettando.
affettare	no	no	yes	Luca stava affettando di nuovo.
affettare	no	yes	no	Giulia aveva affettato.
affettare	no	yes	yes	Paolo aveva affettato di nuovo.
avvelenare	no	no	no	Luca stava avvelenando.
avvelenare	no	no	yes	Giulia stava avvelenando di nuovo.
avvelenare	no	yes	no	Maria aveva avvelenato.
avvelenare	no	yes	yes	Franco aveva avvelenato di nuovo.
bere	no	no	no	Giulia stava bevendo.
bere	no	no	yes	Giulia stava bevendo di nuovo.
bere	no	yes	no	Maria aveva bevuto.
bere	no	yes	yes	Sara aveva bevuto di nuovo.
cantare	no	no	no	Maria stava cantando.
cantare	no	no	yes	Franco stava cantando di nuovo.
cantare	no	yes	no	Franco aveva cantato.
cantare	no	yes	yes	Franco aveva cantato di nuovo.
canticchiare	no	no	no	Marta stava canticchiando.
canticchiare	no	no	yes	Paolo stava canticchiando di nuovo.
canticchiare	no	yes	no	Sara aveva canticchiato.
canticchiare	no	yes	yes	Franco aveva canticchiato di nuovo.
costruire	no	no	no	Sara stava costruendo.
costruire	no	no	yes	Franco stava costruendo di nuovo.
costruire	no	yes	no	Marta aveva costruito.
costruire	no	yes	yes	Gianni aveva costruito di nuovo.
cucinare	no	no	no	Maria stava cucinando.
cucinare	no	no	yes	Franco stava cucinando di nuovo.
cucinare	no	yes	no	Piero aveva cucinato.
cucinare	no	yes	yes	Franco aveva cucinato di nuovo.
cucire	no	no	no	Gianni stava cucendo.

verb	dObj	perfective	iterative	sentence
cucire	no	no	yes	Marta stava cucendo di nuovo.
cucire	no	yes	no	Franco aveva cucito.
cucire	no	yes	yes	Luca aveva cucito di nuovo.
decapitare	no	no	no	Piero stava decapitando.
decapitare	no	no	yes	Paolo stava decapitando di nuovo.
decapitare	no	yes	no	Giulia aveva decapitato.
decapitare	no	yes	yes	Paolo aveva decapitato di nuovo.
divorare	no	no	no	Carla stava divorando.
divorare	no	no	yes	Marta stava divorando di nuovo.
divorare	no	yes	no	Franco aveva divorato.
divorare	no	yes	yes	Sara aveva divorato di nuovo.
firmare	no	no	no	Maria stava firmando.
firmare	no	no	yes	Piero stava firmando di nuovo.
firmare	no	yes	no	Sara aveva firmato.
firmare	no	yes	yes	Marta aveva firmato di nuovo.
fumare	no	no	no	Gianni stava fumando.
fumare	no	no	yes	Carla stava fumando di nuovo.
fumare	no	yes	no	Luca aveva fumato.
fumare	no	yes	yes	Marta aveva fumato di nuovo.
guardare	no	no	no	Franco stava guardando.
guardare	no	no	yes	Luca stava guardando di nuovo.
guardare	no	yes	no	Gianni aveva guardato.
guardare	no	yes	yes	Gianni aveva guardato di nuovo.
insegnare	no	no	no	Marta stava insegnando.
insegnare	no	no	yes	Sara stava insegnando di nuovo.
insegnare	no	yes	no	Paolo aveva insegnato.
insegnare	no	yes	yes	Franco aveva insegnato di nuovo.
lavare	no	no	no	Giulia stava lavando.
lavare	no	no	yes	Luca stava lavando di nuovo.
lavare	no	yes	no	Giulia aveva lavato.
lavare	no	yes	yes	Franco aveva lavato di nuovo.
lucidare	no	no	no	Carla stava lucidando.
lucidare	no	no	yes	Sara stava lucidando di nuovo.
lucidare	no	yes	no	Giulia aveva lucidato.
lucidare	no	yes	yes	Luca aveva lucidato di nuovo.
mangiare	no	no	no	Sara stava mangiando.
mangiare	no	no	yes	Franco stava mangiando di nuovo.
mangiare	no	yes	no	Sara aveva mangiato.
mangiare	no	yes	yes	Paolo aveva mangiato di nuovo.
pulire	no	no	no	Luca stava pulendo.
pulire	no	no	yes	Paolo stava pulendo di nuovo.
pulire	no	yes	no	Sara aveva pulito.
pulire	no	yes	yes	Piero aveva pulito di nuovo.
ricamare	no	no	no	Luca stava ricamando.
ricamare	no	no	yes	Franco stava ricamando di nuovo.
ricamare	no	yes	no	Giulia aveva ricamato.
ricamare	no	yes	yes	Carla aveva ricamato di nuovo.
rompere	no	no	no	Paolo stava rompendo.
rompere	no	no	yes	Luca stava rompendo di nuovo.
rompere	no	yes	no	Piero aveva rotto.
rompere	no	yes	yes	Paolo aveva rotto di nuovo.
rubare	no	no	no	Marta stava rubando.

verb	dObj	perfective	iterative	sentence
rubare	no	no	yes	Maria stava rubando di nuovo.
rubare	no	yes	no	Sara aveva rubato.
rubare	no	yes	yes	Carla aveva rubato di nuovo.
scarabocchiare	no	no	no	Franco stava scarabocchiando.
scarabocchiare	no	no	yes	Paolo stava scarabocchiando di nuovo.
scarabocchiare	no	yes	no	Marta aveva scarabocchiato.
scarabocchiare	no	yes	yes	Franco aveva scarabocchiato di nuovo.
scrivere	no	no	no	Sara stava scrivendo.
scrivere	no	no	yes	Carla stava scrivendo di nuovo.
scrivere	no	yes	no	Giulia aveva scritto.
scrivere	no	yes	yes	Marta aveva scritto di nuovo.
sorseggiare	no	no	no	Sara stava sorseggiando.
sorseggiare	no	no	yes	Maria stava sorseggiando di nuovo.
sorseggiare	no	yes	no	Piero aveva sorseggiato.
sorseggiare	no	yes	yes	Marta aveva sorseggiato di nuovo.
spaccare	no	no	no	Giulia stava spaccando.
spaccare	no	no	yes	Gianni stava spaccando di nuovo.
spaccare	no	yes	no	Gianni aveva spaccato.
spaccare	no	yes	yes	Marta aveva spaccato di nuovo.
tagliare	no	no	no	Marta stava tagliando.
tagliare	no	no	yes	Maria stava tagliando di nuovo.
tagliare	no	yes	no	Luca aveva tagliato.
tagliare	no	yes	yes	Piero aveva tagliato di nuovo.
trangugiare	no	no	no	Franco stava trangugiando.
trangugiare	no	no	yes	Sara stava trangugiando di nuovo.
trangugiare	no	yes	no	Giulia aveva trangugiato.
trangugiare	no	yes	yes	Sara aveva trangugiato di nuovo.
uccidere	no	no	no	Maria stava uccidendo.
uccidere	no	no	yes	Carla stava uccidendo di nuovo.
uccidere	no	yes	no	Giulia aveva ucciso.
uccidere	no	yes	yes	Giulia aveva ucciso di nuovo.
versare	no	no	no	Marta stava versando.
versare	no	no	yes	Franco stava versando di nuovo.
versare	no	yes	no	Maria aveva versato.
versare	no	yes	yes	Paolo aveva versato di nuovo.

D.2.2 Control sentences

verb	dObj	perfective	iterative	sentence
accoltellare	yes	no	no	Sara stava accoltellando un uomo.
accoltellare	yes	no	yes	Carla stava accoltellando un uomo di nuovo.
accoltellare	yes	yes	no	Sara aveva accoltellato un uomo.
accoltellare	yes	yes	yes	Gianni aveva accoltellato un uomo di nuovo.
affettare	yes	no	no	Paolo stava affettando una torta.
affettare	yes	no	yes	Piero stava affettando una torta di nuovo.
affettare	yes	yes	no	Carla aveva affettato una torta.

verb	dObj	perfective	iterative	sentence
affettare	yes	yes	yes	Piero aveva affettato una torta di nuovo.
avvelenare	yes	no	no	Marta stava avvelenando una pianta.
avvelenare	yes	no	yes	Marta stava avvelenando una pianta di nuovo.
avvelenare	yes	yes	no	Carla aveva avvelenato una pianta.
avvelenare	yes	yes	yes	Carla aveva avvelenato una pianta di nuovo.
bere	yes	no	no	Carla stava bevendo del succo.
bere	yes	no	yes	Paolo stava bevendo del succo di nuovo.
bere	yes	yes	no	Maria aveva bevuto del succo.
bere	yes	yes	yes	Carla aveva bevuto del succo di nuovo.
cantare	yes	no	no	Marta stava cantando un inno.
cantare	yes	no	yes	Gianni stava cantando un inno di nuovo.
cantare	yes	yes	no	Sara aveva cantato un inno.
cantare	yes	yes	yes	Marta aveva cantato un inno di nuovo.
canticchiare	yes	no	no	Paolo stava canticchiando una ninna-nanna.
canticchiare	yes	no	yes	Maria stava canticchiando una ninna-nanna di nuovo.
canticchiare	yes	yes	no	Gianni aveva canticchiato una ninna-nanna.
canticchiare	yes	yes	yes	Franco aveva canticchiato una ninna-nanna di nuovo.
costruire	yes	no	no	Luca stava costruendo una casa.
costruire	yes	no	yes	Marta stava costruendo una casa di nuovo.
costruire	yes	yes	no	Luca aveva costruito una casa.
costruire	yes	yes	yes	Paolo aveva costruito una casa di nuovo.
cucinare	yes	no	no	Maria stava cucinando la cena.
cucinare	yes	no	yes	Giulia stava cucinando la cena di nuovo.
cucinare	yes	yes	no	Marta aveva cucinato la cena.
cucinare	yes	yes	yes	Franco aveva cucinato la cena di nuovo.
cucire	yes	no	no	Sara stava cucendo una tenda.
cucire	yes	no	yes	Franco stava cucendo una tenda di nuovo.
cucire	yes	yes	no	Maria aveva cucito una tenda.
cucire	yes	yes	yes	Franco aveva cucito una tenda di nuovo.
decapitare	yes	no	no	Luca stava decapitando un prigioniero.
decapitare	yes	no	yes	Luca stava decapitando un prigioniero di nuovo.
decapitare	yes	yes	no	Gianni aveva decapitato un prigioniero.

verb	dObj	perfective	iterative	sentence
decapitare	yes	yes	yes	Piero aveva decapitato un prigioniero di nuovo.
divorare	yes	no	no	Gianni stava divorando un pollo arrosto.
divorare	yes	no	yes	Gianni stava divorando un pollo arrosto di nuovo.
divorare	yes	yes	no	Carla aveva divorato un pollo arrosto.
divorare	yes	yes	yes	Franco aveva divorato un pollo arrosto di nuovo.
firmare	yes	no	no	Marta stava firmando un documento.
firmare	yes	no	yes	Gianni stava firmando un documento di nuovo.
firmare	yes	yes	no	Luca aveva firmato un documento.
firmare	yes	yes	yes	Sara aveva firmato un documento di nuovo.
fumare	yes	no	no	Franco stava fumando una sigaretta.
fumare	yes	no	yes	Piero stava fumando una sigaretta di nuovo.
fumare	yes	yes	no	Luca aveva fumato una sigaretta.
fumare	yes	yes	yes	Carla aveva fumato una sigaretta di nuovo.
guardare	yes	no	no	Paolo stava guardando un film.
guardare	yes	no	yes	Luca stava guardando un film di nuovo.
guardare	yes	yes	no	Giulia aveva guardato un film.
guardare	yes	yes	yes	Gianni aveva guardato un film di nuovo.
insegnare	yes	no	no	Carla stava insegnando linguistica.
insegnare	yes	no	yes	Gianni stava insegnando linguistica di nuovo.
insegnare	yes	yes	no	Marta aveva insegnato linguistica.
insegnare	yes	yes	yes	Franco aveva insegnato linguistica di nuovo.
lavare	yes	no	no	Marta stava lavando una macchina.
lavare	yes	no	yes	Piero stava lavando una macchina di nuovo.
lavare	yes	yes	no	Maria aveva lavato una macchina.
lavare	yes	yes	yes	Franco aveva lavato una macchina di nuovo.
lucidare	yes	no	no	Paolo stava lucidando una spada.
lucidare	yes	no	yes	Giulia stava lucidando una spada di nuovo.
lucidare	yes	yes	no	Franco aveva lucidato una spada.
lucidare	yes	yes	yes	Piero aveva lucidato una spada di nuovo.
mangiare	yes	no	no	Paolo stava mangiando della pizza.

verb	dObj	perfective	iterative	sentence
mangiare	yes	no	yes	Marta stava mangiando della pizza di nuovo.
mangiare	yes	yes	no	Luca aveva mangiato della pizza.
mangiare	yes	yes	yes	Gianni aveva mangiato della pizza di nuovo.
pulire	yes	no	no	Carla stava pulendo un tavolo.
pulire	yes	no	yes	Gianni stava pulendo un tavolo di nuovo.
pulire	yes	yes	no	Giulia aveva pulito un tavolo.
pulire	yes	yes	yes	Sara aveva pulito un tavolo di nuovo.
ricamare	yes	no	no	Paolo stava ricamando un arazzo.
ricamare	yes	no	yes	Paolo stava ricamando un arazzo di nuovo.
ricamare	yes	yes	no	Luca aveva ricamato un arazzo.
ricamare	yes	yes	yes	Franco aveva ricamato un arazzo di nuovo.
rompere	yes	no	no	Piero stava rompendo un vaso.
rompere	yes	no	yes	Marta stava rompendo un vaso di nuovo.
rompere	yes	yes	no	Sara aveva rotto un vaso.
rompere	yes	yes	yes	Giulia aveva rotto un vaso di nuovo.
rubare	yes	no	no	Paolo stava rubando dei soldi.
rubare	yes	no	yes	Paolo stava rubando dei soldi di nuovo.
rubare	yes	yes	no	Piero aveva rubato dei soldi.
rubare	yes	yes	yes	Marta aveva rubato dei soldi di nuovo.
scarabocchiare	yes	no	no	Luca stava scarabocchiando un omino.
scarabocchiare	yes	no	yes	Giulia stava scarabocchiando un omino di nuovo.
scarabocchiare	yes	yes	no	Franco aveva scarabocchiato un omino.
scarabocchiare	yes	yes	yes	Gianni aveva scarabocchiato un omino di nuovo.
scrivere	yes	no	no	Luca stava scrivendo una lettera.
scrivere	yes	no	yes	Giulia stava scrivendo una lettera di nuovo.
scrivere	yes	yes	no	Sara aveva scritto una lettera.
scrivere	yes	yes	yes	Sara aveva scritto una lettera di nuovo.
sorseggiare	yes	no	no	Giulia stava sorseggiando dell'acqua.
sorseggiare	yes	no	yes	Franco stava sorseggiando dell'acqua di nuovo.
sorseggiare	yes	yes	no	Gianni aveva sorseggiato dell'acqua.
sorseggiare	yes	yes	yes	Luca aveva sorseggiato dell'acqua di nuovo.
spaccare	yes	no	no	Luca stava spaccando un tronco.

verb	dObj	perfective	iterative	sentence
spaccare	yes	no	yes	Maria stava spaccando un tronco di nuovo.
spaccare	yes	yes	no	Carla aveva spaccato un tronco.
spaccare	yes	yes	yes	Carla aveva spaccato un tronco di nuovo.
tagliare	yes	no	no	Paolo stava tagliando della carta.
tagliare	yes	no	yes	Sara stava tagliando della carta di nuovo.
tagliare	yes	yes	no	Piero aveva tagliato della carta.
tagliare	yes	yes	yes	Luca aveva tagliato della carta di nuovo.
trangugiare	yes	no	no	Luca stava trangugiando una birra.
trangugiare	yes	no	yes	Sara stava trangugiando una birra di nuovo.
trangugiare	yes	yes	no	Carla aveva trangugiato una birra.
trangugiare	yes	yes	yes	Sara aveva trangugiato una birra di nuovo.
uccidere	yes	no	no	Marta stava uccidendo dei parassiti.
uccidere	yes	no	yes	Carla stava uccidendo dei parassiti di nuovo.
uccidere	yes	yes	no	Luca aveva ucciso dei parassiti.
uccidere	yes	yes	yes	Sara aveva ucciso dei parassiti di nuovo.
versare	yes	no	no	Sara stava versando del vino.
versare	yes	no	yes	Carla stava versando del vino di nuovo.
versare	yes	yes	no	Maria aveva versato del vino.
versare	yes	yes	yes	Carla aveva versato del vino di nuovo.

D.2.3 Filler sentences

verb	dObj	perfective	iterative	sentence
applaudire	no	no	no	Sara stava applaudendo.
applaudire	no	no	yes	Carla stava applaudendo di nuovo.
applaudire	no	yes	no	Paolo aveva applaudito.
applaudire	no	yes	yes	Piero aveva applaudito di nuovo.
applaudire	yes	no	no	Giulia stava applaudendo uno spettacolo.
applaudire	yes	no	yes	Marta stava applaudendo uno spettacolo di nuovo.
applaudire	yes	yes	no	Marta aveva applaudito uno spettacolo.
applaudire	yes	yes	yes	Marta aveva applaudito uno spettacolo di nuovo.
barcollare	no	no	no	Giulia stava barcollando.
barcollare	no	no	yes	Maria stava barcollando di nuovo.
barcollare	no	yes	no	Giulia aveva barcollato.
barcollare	no	yes	yes	Giulia aveva barcollato di nuovo.

verb	dObj	perfective	iterative	sentence
barcollare	yes	no	no	Giulia stava barcollando il marciapiede.
barcollare	yes	no	yes	Marta stava barcollando il marciapiede di nuovo.
barcollare	yes	yes	no	Maria aveva barcollato il marciapiede.
barcollare	yes	yes	yes	Paolo aveva barcollato il marciapiede di nuovo.
bussare	no	no	no	Giulia stava bussando.
bussare	no	no	yes	Luca stava bussando di nuovo.
bussare	no	yes	no	Gianni aveva bussato.
bussare	no	yes	yes	Franco aveva bussato di nuovo.
bussare	yes	no	no	Piero stava bussando una porta.
bussare	yes	no	yes	Piero stava bussando una porta di nuovo.
bussare	yes	yes	no	Franco aveva bussato una porta.
bussare	yes	yes	yes	Paolo aveva bussato una porta di nuovo.
digiunare	no	no	no	Paolo stava digiunando.
digiunare	no	no	yes	Gianni stava digiunando di nuovo.
digiunare	no	yes	no	Maria aveva digiunato.
digiunare	no	yes	yes	Piero aveva digiunato di nuovo.
digiunare	yes	no	no	Giulia stava digiunando del sushi.
digiunare	yes	no	yes	Sara stava digiunando del sushi di nuovo.
digiunare	yes	yes	no	Marta aveva digiunato del sushi.
digiunare	yes	yes	yes	Marta aveva digiunato del sushi di nuovo.
dormire	no	no	no	Piero stava dormendo.
dormire	no	no	yes	Maria stava dormendo di nuovo.
dormire	no	yes	no	Piero aveva dormito.
dormire	no	yes	yes	Carla aveva dormito di nuovo.
dormire	yes	no	no	Sara stava dormendo un cuscino.
dormire	yes	no	yes	Luca stava dormendo un cuscino di nuovo.
dormire	yes	yes	no	Giulia aveva dormito un cuscino.
dormire	yes	yes	yes	Carla aveva dormito un cuscino di nuovo.
ridere	no	no	no	Sara stava ridendo.
ridere	no	no	yes	Piero stava ridendo di nuovo.
ridere	no	yes	no	Sara aveva riso.
ridere	no	yes	yes	Paolo aveva riso di nuovo.
ridere	yes	no	no	Luca stava ridendo una barzelletta.
ridere	yes	no	yes	Giulia stava ridendo una barzelletta di nuovo.
ridere	yes	yes	no	Marta aveva riso una barzelletta.
ridere	yes	yes	yes	Giulia aveva riso una barzelletta di nuovo.
riposarsi	no	no	no	Luca si stava riposando.
riposarsi	no	no	yes	Gianni si stava riposando di nuovo.

verb	dObj	perfective	iterative	sentence
riposarsi	no	yes	no	Piero si è riposato.
riposarsi	no	yes	yes	Sara si è riposata di nuovo.
riposarsi	yes	no	no	Luca si stava riposando un letto.
riposarsi	yes	no	yes	Maria si stava riposando un letto di nuovo.
riposarsi	yes	yes	no	Piero si è riposato un letto.
riposarsi	yes	yes	yes	Luca si è riposato un letto di nuovo.
sorridere	no	no	no	Sara stava sorridendo.
sorridere	no	no	yes	Carla stava sorridendo di nuovo.
sorridere	no	yes	no	Paolo aveva sorriso.
sorridere	no	yes	yes	Maria aveva sorriso di nuovo.
sorridere	yes	no	no	Carla stava sorridendo un amico.
sorridere	yes	no	yes	Giulia stava sorridendo un amico di nuovo.
sorridere	yes	yes	no	Gianni aveva sorriso un amico.
sorridere	yes	yes	yes	Sara aveva sorriso un amico di nuovo.
urlare	no	no	no	Sara stava urlando.
urlare	no	no	yes	Franco stava urlando di nuovo.
urlare	no	yes	no	Luca aveva urlato.
urlare	no	yes	yes	Marta aveva urlato di nuovo.
urlare	yes	no	no	Gianni stava urlando un ragno.
urlare	yes	no	yes	Carla stava urlando un ragno di nuovo.
urlare	yes	yes	no	Paolo aveva urlato un ragno.
urlare	yes	yes	yes	Gianni aveva urlato un ragno di nuovo.
zoppicare	no	no	no	Luca stava zoppicando.
zoppicare	no	no	yes	Maria stava zoppicando di nuovo.
zoppicare	no	yes	no	Paolo aveva zoppicato.
zoppicare	no	yes	yes	Piero aveva zoppicato di nuovo.
zoppicare	yes	no	no	Luca stava zoppicando una strada.
zoppicare	yes	no	yes	Paolo stava zoppicando una strada di nuovo.
zoppicare	yes	yes	no	Maria aveva zoppicato una strada.
zoppicare	yes	yes	yes	Sara aveva zoppicato una strada di nuovo.

Stochastic OT models of object drop in English and Italian

E

This appendix collects all the 18 models of the indefinite object construction that I computed for this dissertation, as detailed in Section 9.1.1. The interested reader can browse my GitHub profile to find both the raw data* and the Python scripts† I used to compute these results.

E.1 English

E.1.1 Basic model

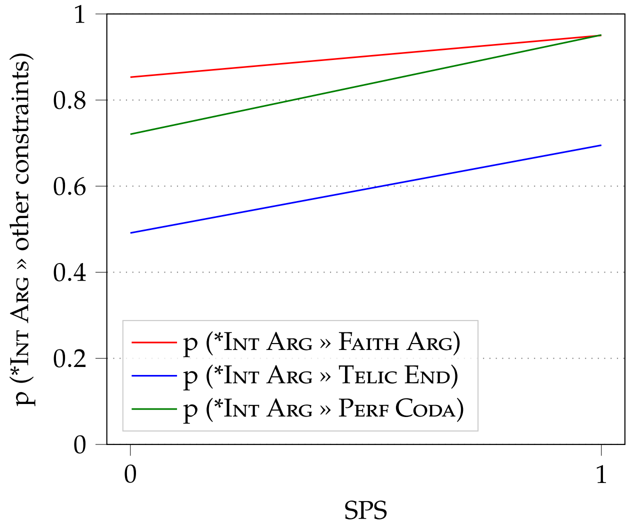

Figure E.1: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Resnik's SPS (English basic model).

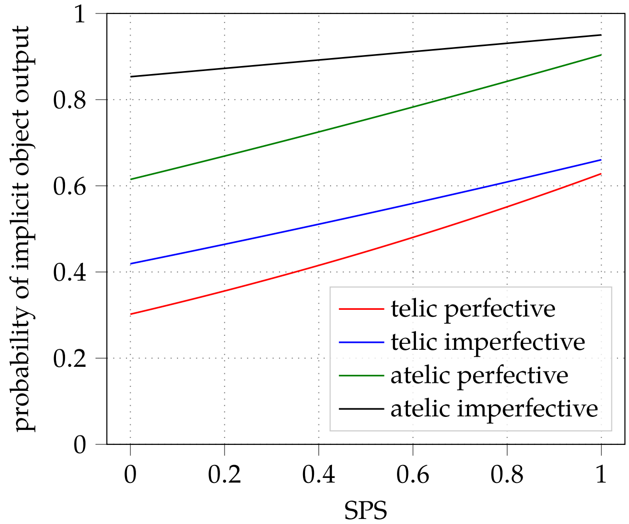

Figure E.2: Probability of an implicit object output for each aspectual type, as a function of Resnik's SPS (English basic model).

* https://github.com/giuliacappelli/dissertationData
† https://github.com/giuliacappelli/MedinaStochasticOptimalityTheory

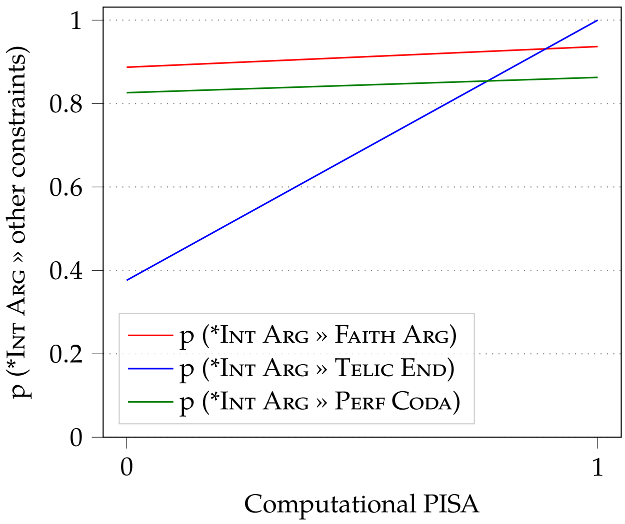

Figure E.3: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Computational PISA (English basic model).

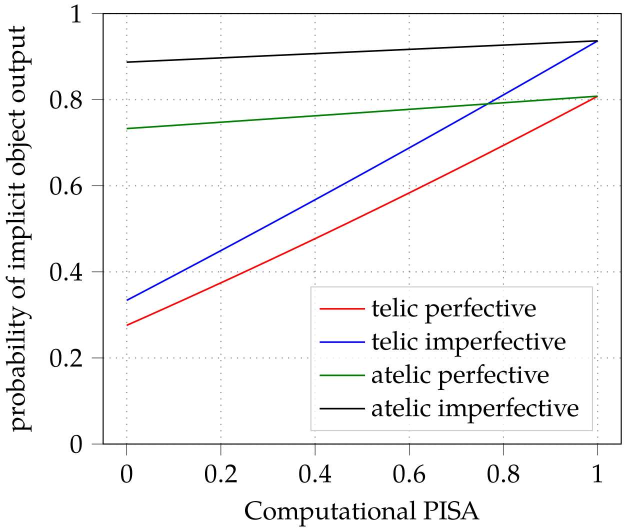

Figure E.4: Probability of an implicit object output for each aspectual type, as a function of Computational PISA (English basic model).

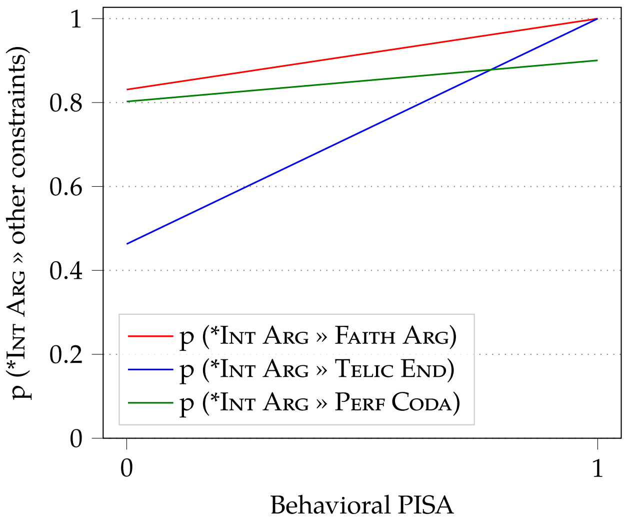

Figure E.5: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Behavioral PISA (English basic model).

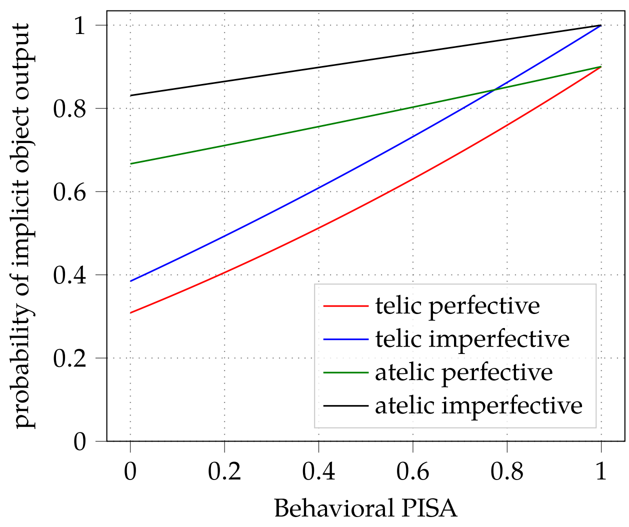

Figure E.6: Probability of an implicit object output for each aspectual type, as a function of Behavioral PISA (English basic model).

E.1.2 Intermediate model

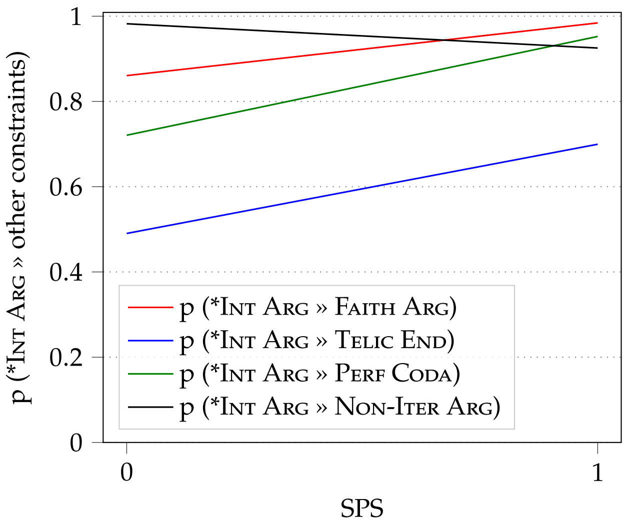

Figure E.7: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Resnik's SPS (English intermediate model).

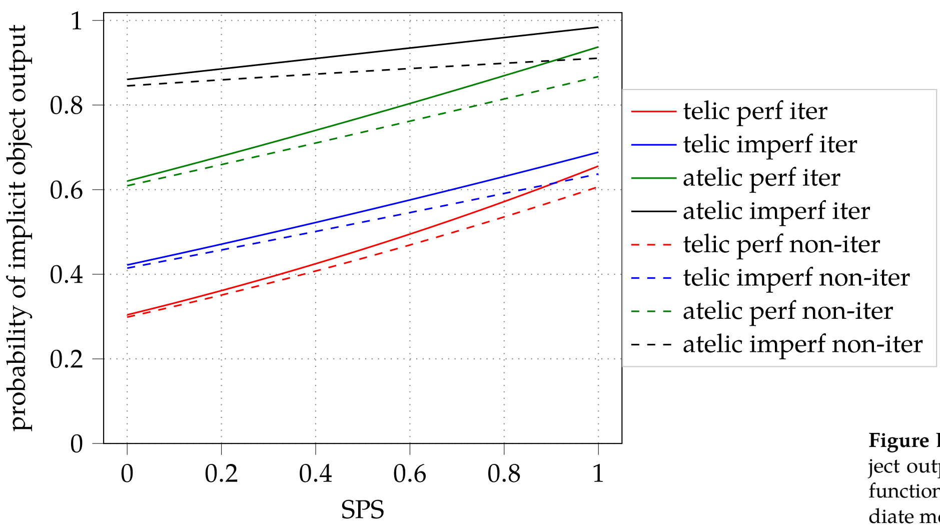

Figure E.8: Probability of an implicit object output for each aspectual type, as a function of Resnik's SPS (English intermediate model).

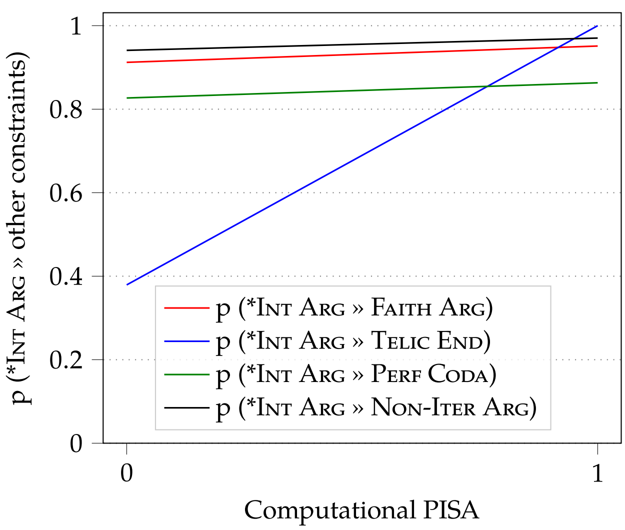

Figure E.9: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Computational PISA (English intermediate model).

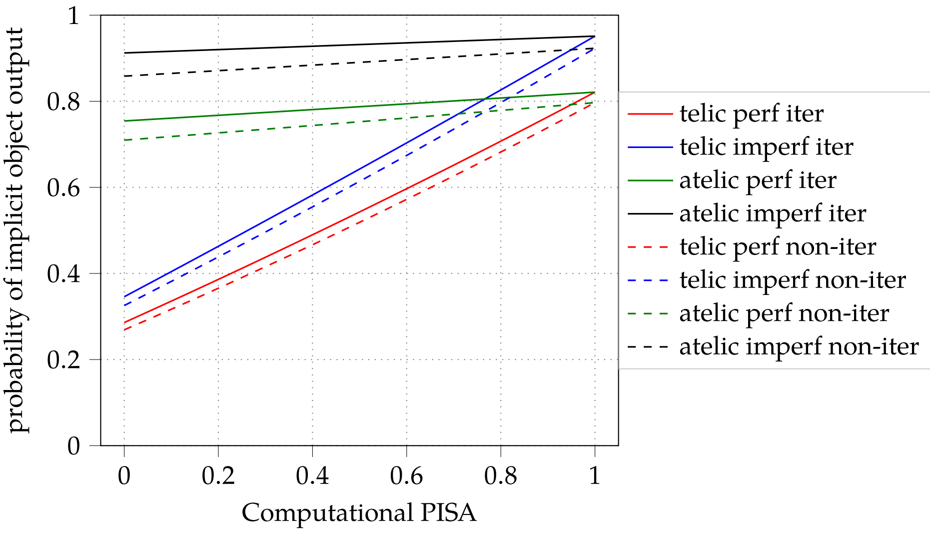

Figure E.10: Probability of an implicit object output for each aspectual type, as a function of Computational PISA (English intermediate model).

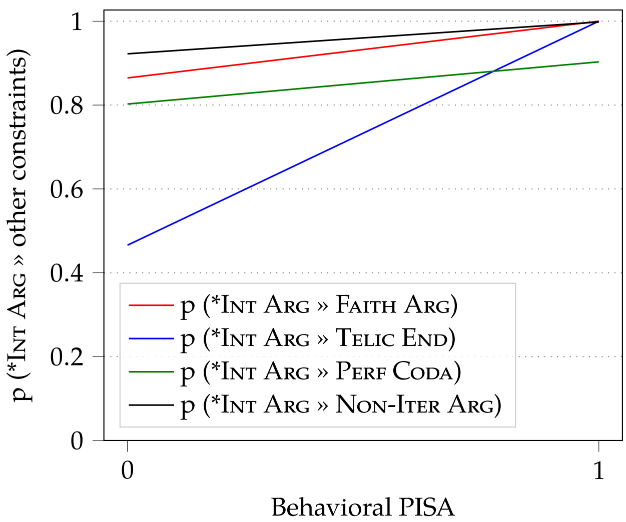

Figure E.11: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Behavioral PISA (English intermediate model).

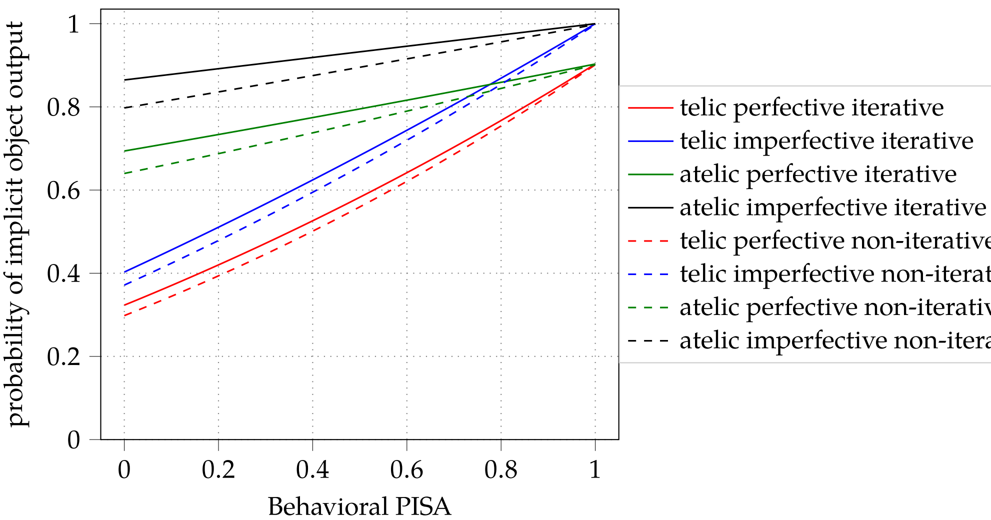

Figure E.12: Probability of an implicit object output for each aspectual type, as a function of Behavioral PISA (English intermediate model).

E.1.3 Full model

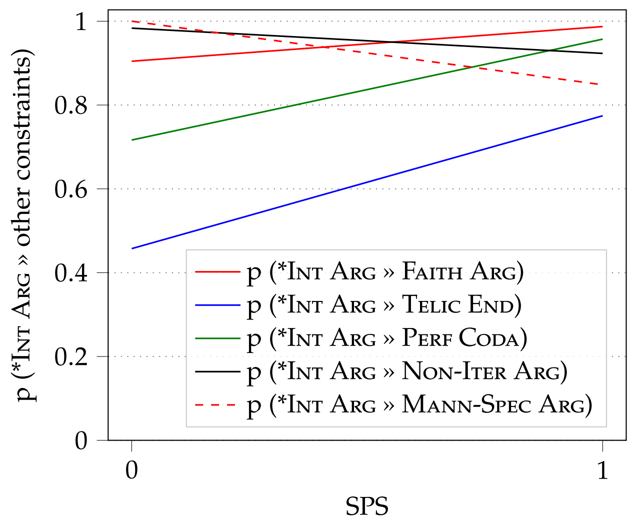

Figure E.13: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Resnik's SPS (English full model).

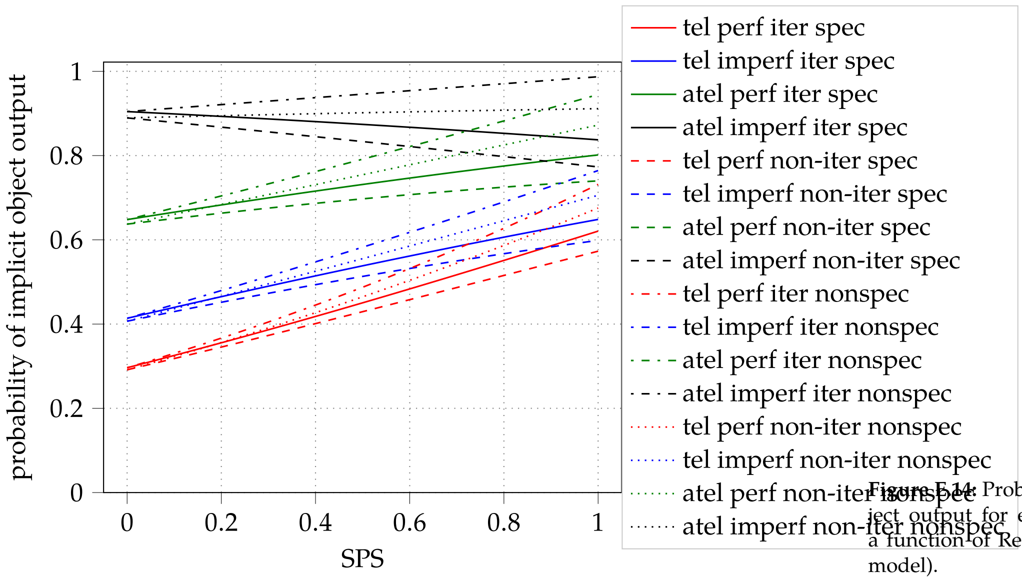

Figure E.14: Probability of an implicit object output for each aspectual type, as a function of Resnik's SPS (English full model).

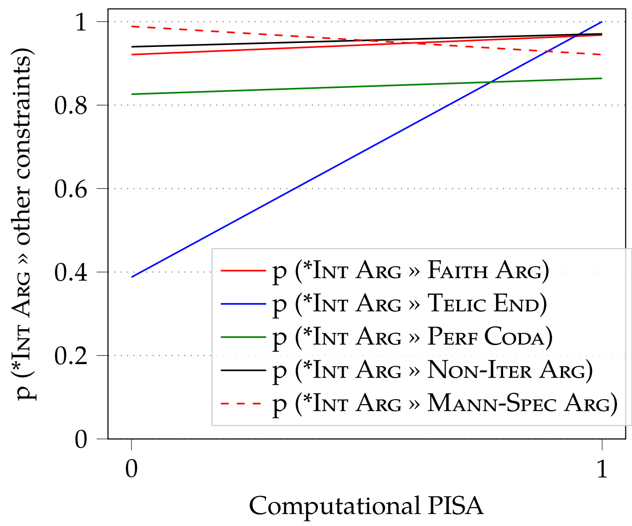

Figure E.15: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Computational PISA (English full model).

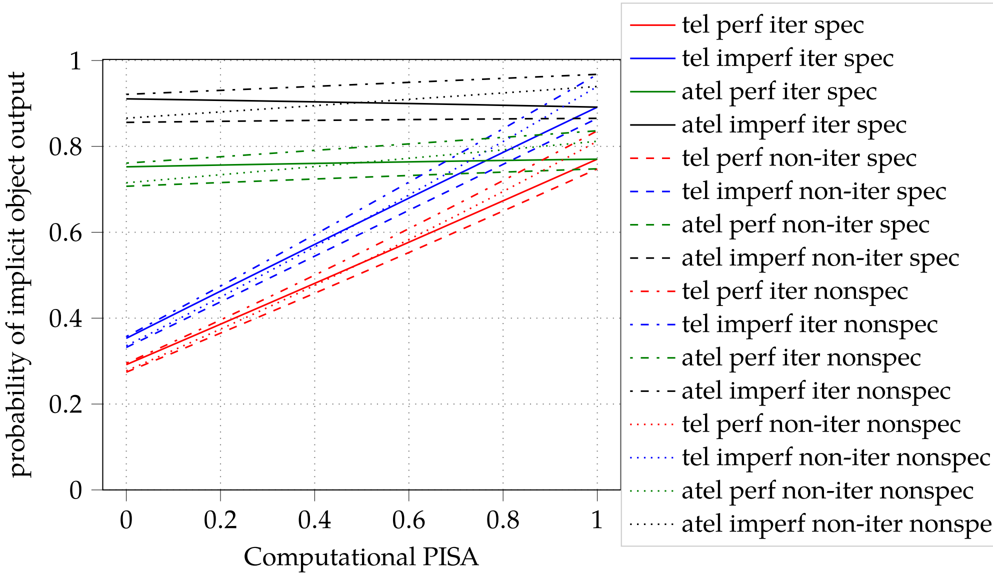

Figure E.16: Probability of an implicit object output for each aspectual type, as a function of Computational PISA (English full model).

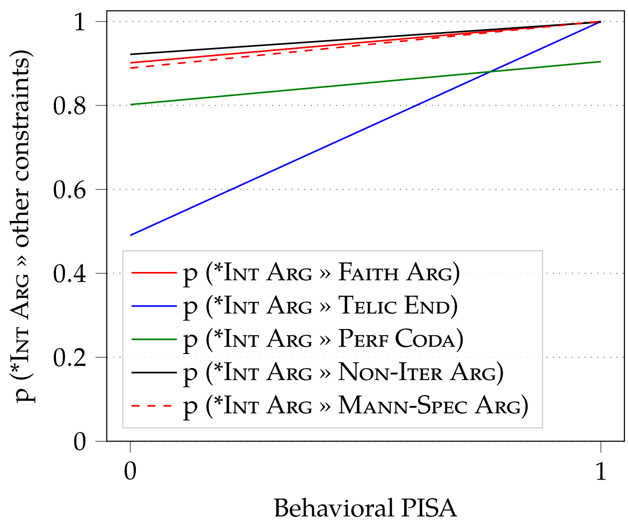

Figure E.17: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Behavioral PISA (English full model).

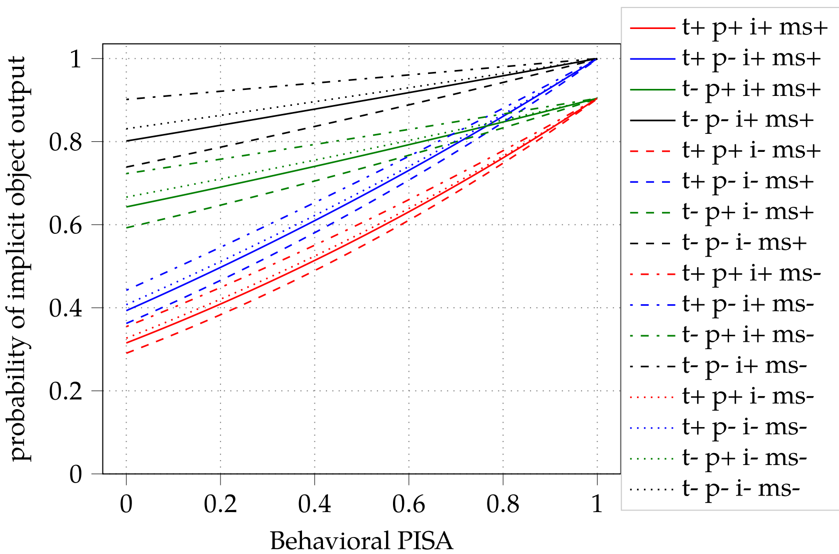

Figure E.18: Probability of an implicit object output for each aspectual type, as a function of Behavioral PISA (English full model).

E.2 Italian

E.2.1 Basic model

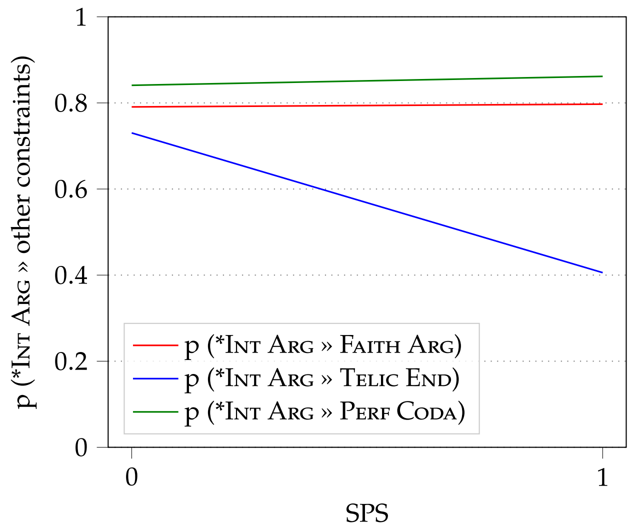

Figure E.19: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Resnik's SPS (Italian basic model).

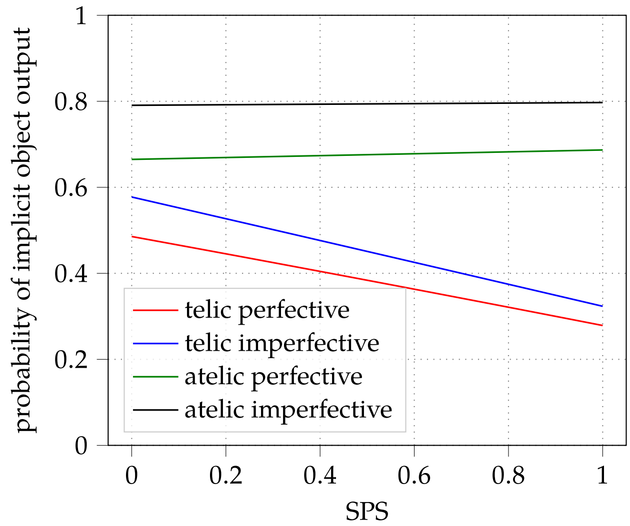

Figure E.20: Probability of an implicit object output for each aspectual type, as a function of Resnik's SPS (Italian basic model).

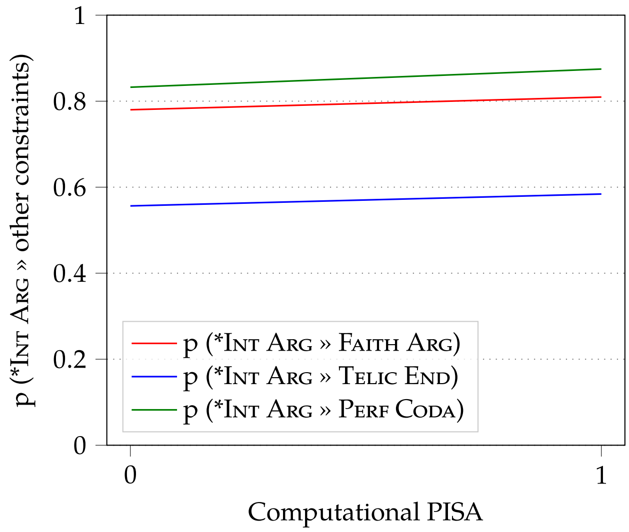

Figure E.21: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Computational PISA (Italian basic model).

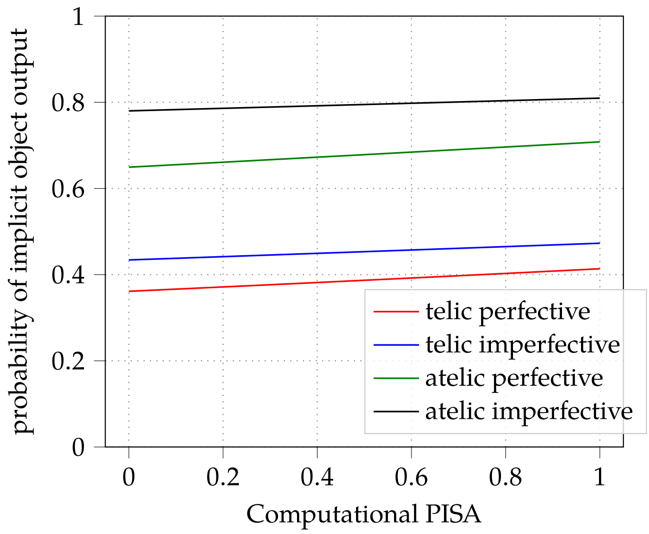

Figure E.22: Probability of an implicit object output for each aspectual type, as a function of Computational PISA (Italian basic model).

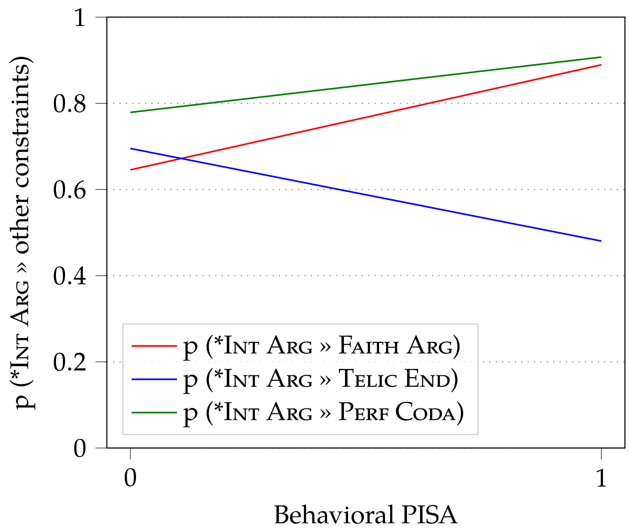

Figure E.23: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Behavioral PISA (Italian basic model).

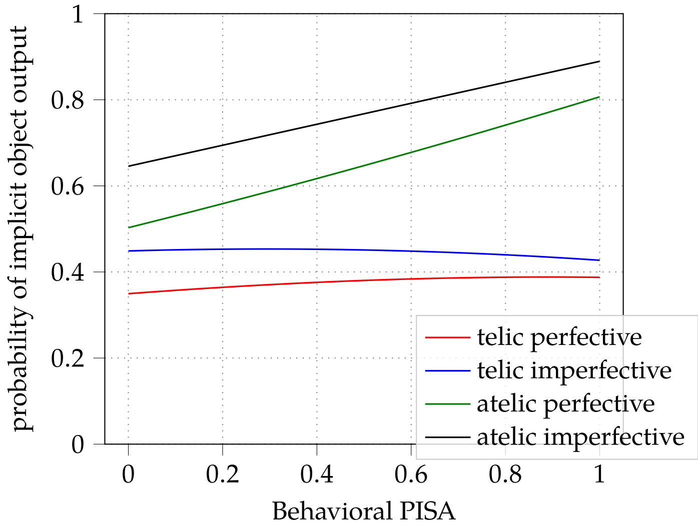

Figure E.24: Probability of an implicit object output for each aspectual type, as a function of Behavioral PISA (Italian basic model).

E.2.2 Intermediate model

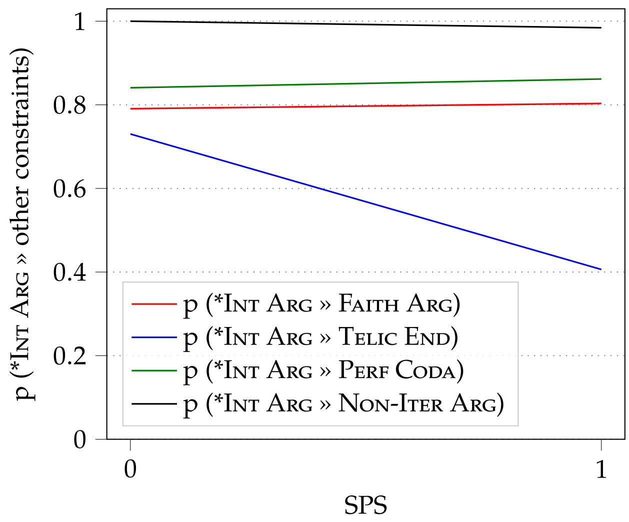

Figure E.25: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Resnik's SPS (Italian intermediate model).

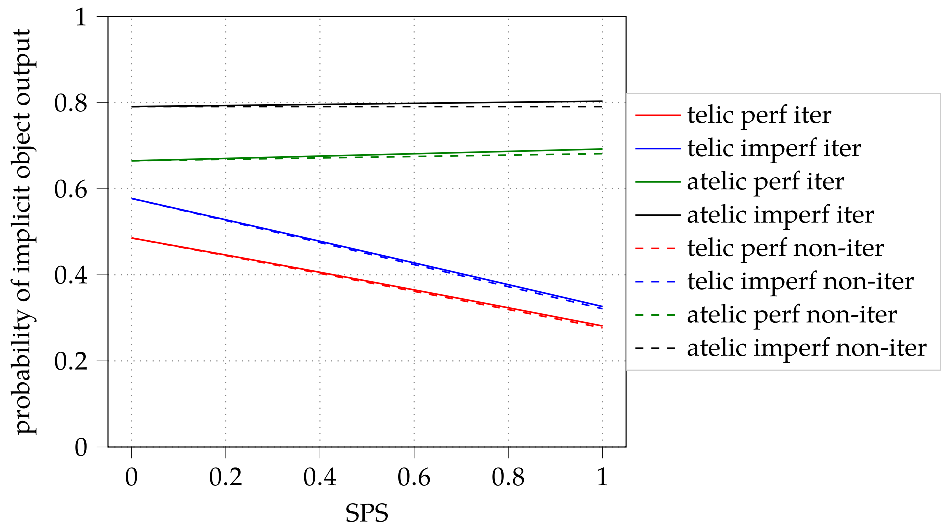

Figure E.26: Probability of an implicit object output for each aspectual type, as a function of Resnik's SPS (Italian intermediate model).

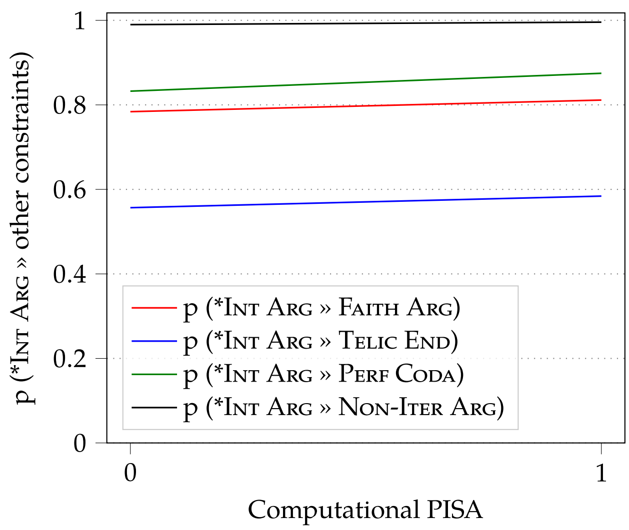

Figure E.27: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Computational PISA (Italian intermediate model).

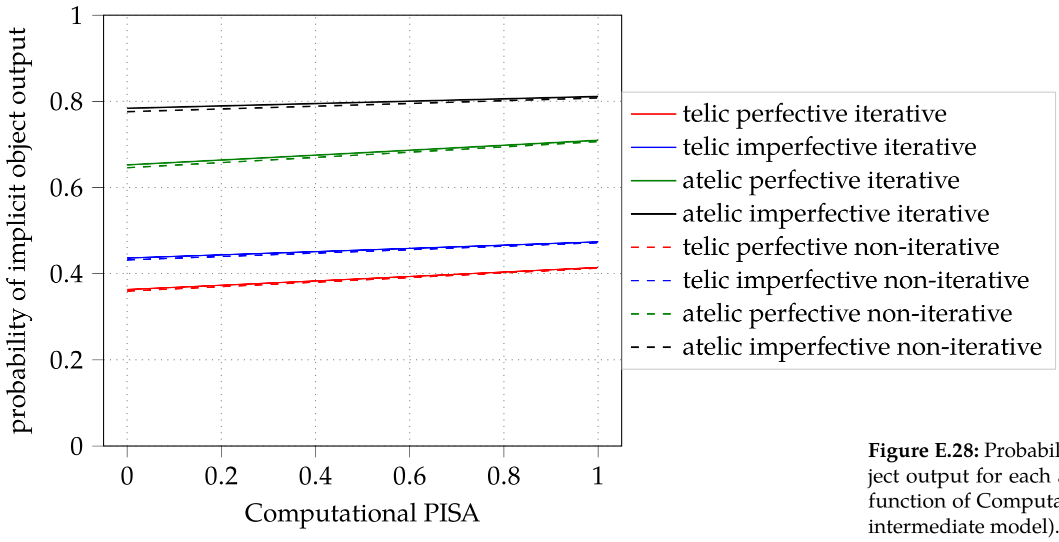

Figure E.28: Probability of an implicit object output for each aspectual type, as a function of Computational PISA (Italian intermediate model).

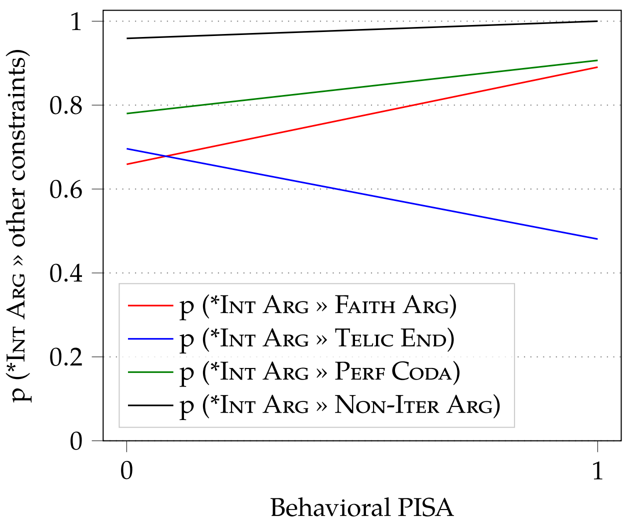

Figure E.29: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Behavioral PISA (Italian intermediate model).

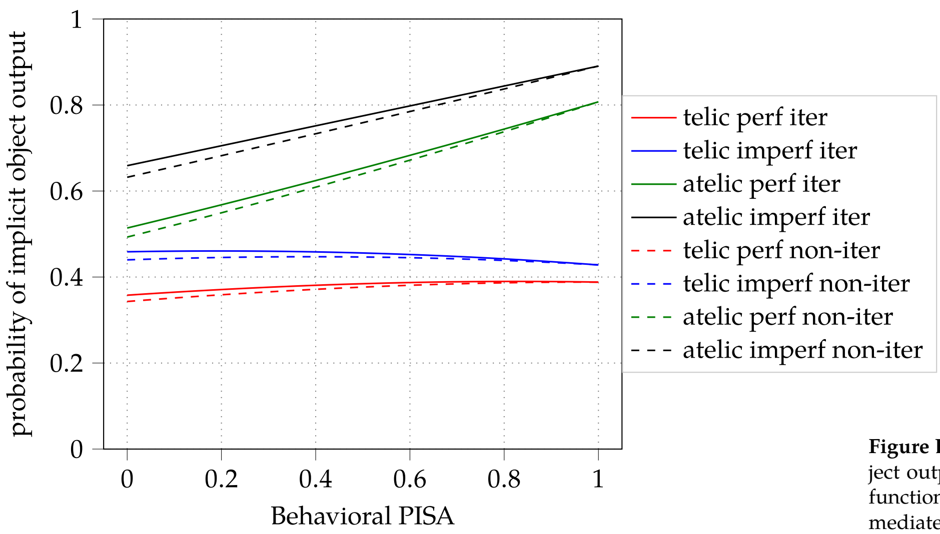

Figure E.30: Probability of an implicit object output for each aspectual type, as a function of Behavioral PISA (Italian intermediate model).

E.2.3 Full model

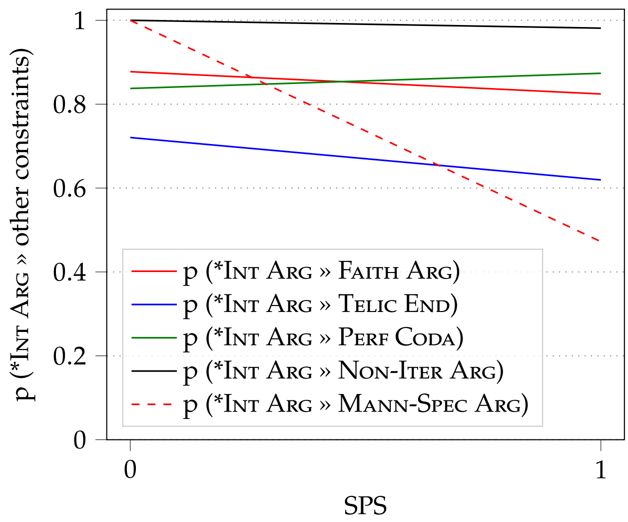

Figure E.31: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Resnik's SPS (Italian full model).

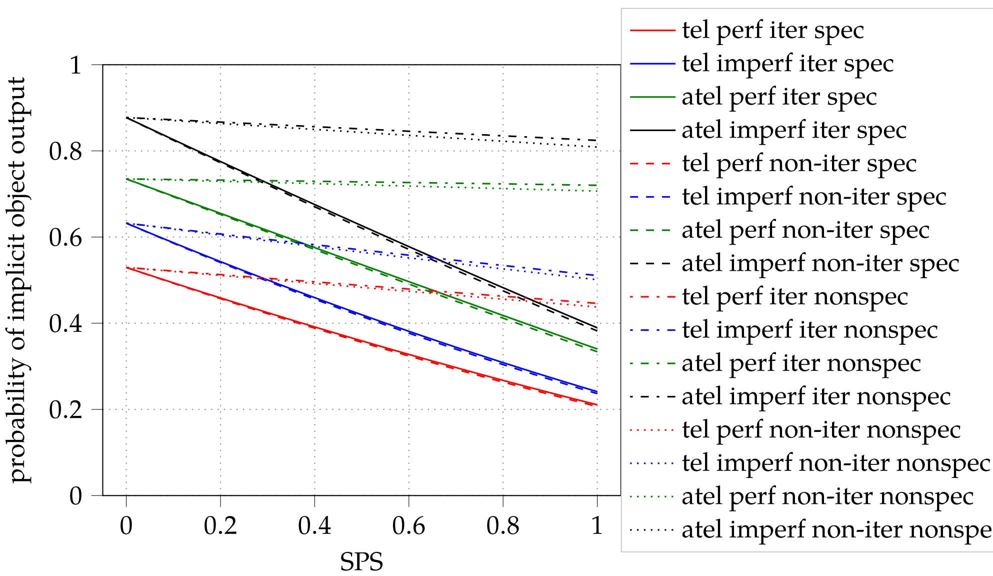

Figure E.32: Probability of an implicit object output for each aspectual type, as a function of Resnik's SPS (Italian full model).

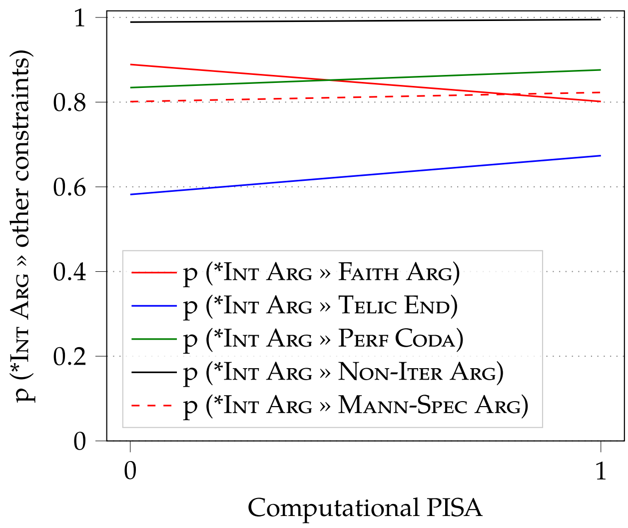

Figure E.33: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Computational PISA (Italian full model).

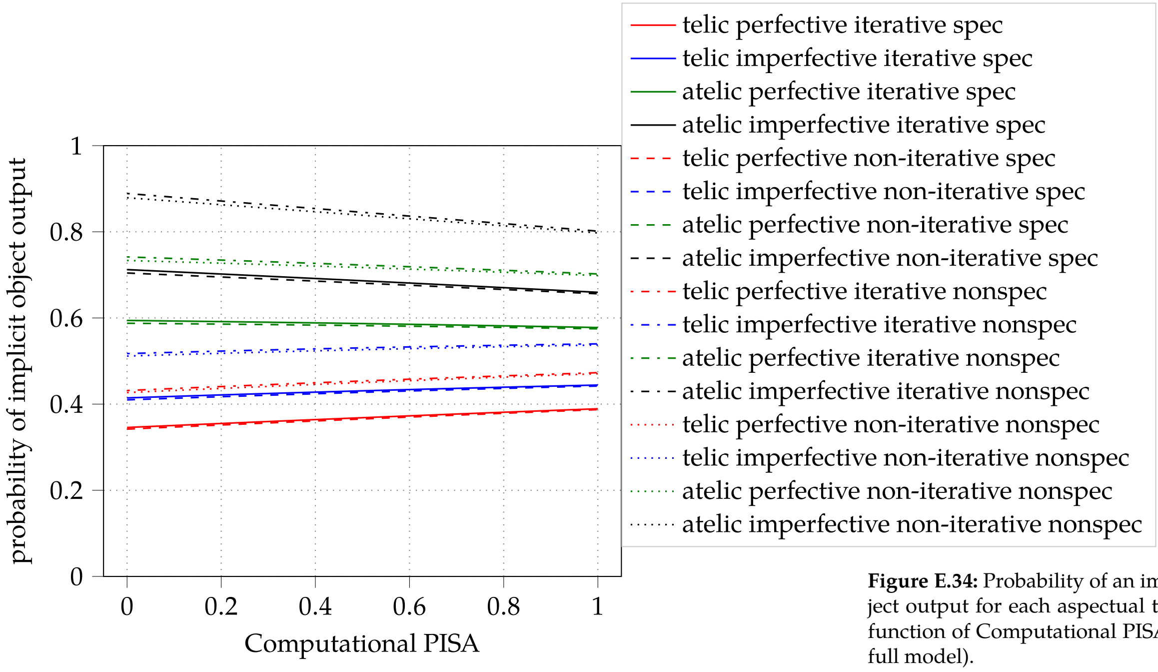

Figure E.34: Probability of an implicit object output for each aspectual type, as a function of Computational PISA (Italian full model).

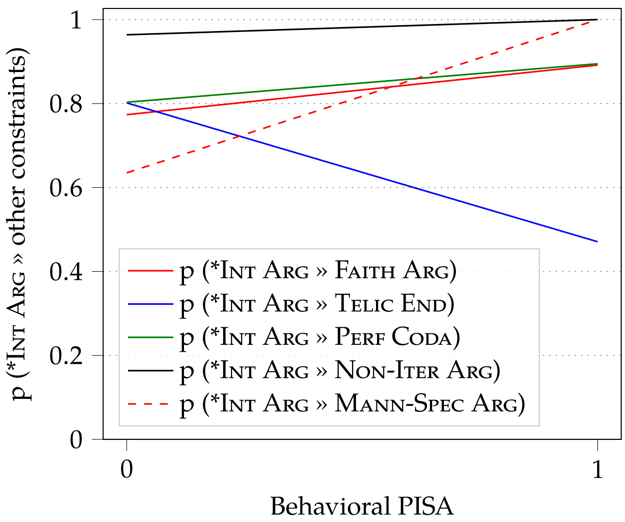

Figure E.35: Probability of *Int Arg being ranked above each of the other constraints, varying in accordance with Behavioral PISA (Italian full model).

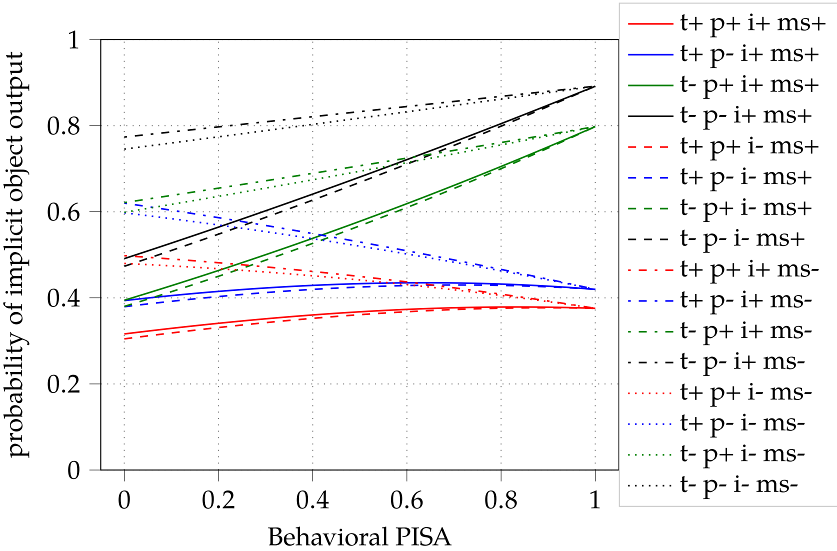

Figure E.36: Probability of an implicit object output for each aspectual type, as a function of Behavioral PISA (Italian full model).

Squared errors: distance between actual judgments and values predicted by the model

F

This appendix collects the actual acceptability judgments provided by human participants, the values predicted by my Stochastic Optimality Theoretic model (in the full version using Behavioral PISA, as explained in full in Chapter 9), and the squared error for each sentence in the set of stimuli for English and Italian (which the interested reader can peruse in Appendix D).
These data are also available here on my GitHub profile[1] , while the Python scripts I wrote to analyse the results and create a Stochastic Optimality Theoretic model of object drop are available on my Github profile[2] in another dedicated repository.

1: https://github.com/giuliacappelli/dissertationData

2: https://github.com/giuliacappelli/MedinaStochasticOptimalityTheory

F.1 English

verb	actual	predicted	squared error
behead	0.251	0.444	0.037
behead	0.507	0.475	0.001
behead	0.155	0.363	0.043
behead	0.567	0.389	0.032
break	0.050	0.445	0.156
break	0.038	0.480	0.195
break	0.000	0.360	0.130
break	0.039	0.388	0.122
build	0.616	0.551	0.004
build	0.883	0.584	0.089
build	0.124	0.457	0.111
build	0.450	0.485	0.001
chop	0.768	0.541	0.052
chop	0.850	0.571	0.078
chop	0.447	0.452	0.000
chop	0.629	0.477	0.023
clean	0.994	0.852	0.020
clean	0.994	0.914	0.006
clean	0.893	0.694	0.039
clean	0.910	0.745	0.027
cook	0.994	0.889	0.011
cook	0.995	0.937	0.003
cook	0.926	0.746	0.032
cook	0.915	0.786	0.017
cut	0.744	0.889	0.021
cut	0.678	0.937	0.067
cut	0.286	0.745	0.211
cut	0.519	0.785	0.071
devour	0.133	0.429	0.087
devour	0.217	0.460	0.059
devour	0.064	0.349	0.081

verb	actual	predicted	squared error
devour	0.057	0.375	0.101
doodle	0.984	0.803	0.033
doodle	0.975	0.852	0.015
doodle	0.951	0.666	0.081
doodle	0.930	0.707	0.050
drink	0.967	0.931	0.001
drink	0.977	0.961	0.000
drink	0.461	0.805	0.119
drink	0.766	0.831	0.004
eat	0.994	0.916	0.006
eat	1.000	0.953	0.002
eat	0.984	0.784	0.040
eat	0.976	0.815	0.026
embroider	0.959	0.835	0.015
embroider	0.930	0.877	0.003
embroider	0.836	0.704	0.018
embroider	0.765	0.739	0.001
hum	0.987	0.829	0.025
hum	0.985	0.872	0.013
hum	0.923	0.696	0.051
hum	0.874	0.733	0.020
kill	0.549	0.641	0.008
kill	0.830	0.671	0.025
kill	0.835	0.543	0.086
kill	0.939	0.568	0.138
knife	0.240	0.467	0.051
knife	0.225	0.498	0.075
knife	0.124	0.384	0.068
knife	0.304	0.410	0.011
poison	0.316	0.419	0.011
poison	0.479	0.451	0.001
poison	0.250	0.341	0.008
poison	0.572	0.367	0.042
polish	0.795	0.803	0.000
polish	0.830	0.852	0.000
polish	0.660	0.666	0.000
polish	0.454	0.707	0.064
pour	0.562	0.888	0.106
pour	0.527	0.936	0.167
pour	0.493	0.743	0.063
pour	0.463	0.784	0.103
sew	0.994	0.897	0.009
sew	0.952	0.941	0.000
sew	0.761	0.756	0.000
sew	0.821	0.794	0.001
sign	0.743	0.767	0.001
sign	0.759	0.787	0.001
sign	0.706	0.669	0.001
sign	0.679	0.687	0.000
sing	1.000	0.999	0.000
sing	0.993	1.000	0.000
sing	0.910	0.903	0.000

verb	actual	predicted	squared error
sing	0.935	0.904	0.001
sip	0.699	0.864	0.027
sip	0.693	0.899	0.042
sip	0.426	0.737	0.097
sip	0.588	0.768	0.032
slice	0.740	0.591	0.022
slice	0.616	0.620	0.000
slice	0.457	0.499	0.002
slice	0.495	0.523	0.001
smoke	0.987	0.953	0.001
smoke	0.992	0.974	0.000
smoke	0.924	0.836	0.008
smoke	0.894	0.854	0.002
steal	0.927	0.407	0.270
steal	0.962	0.442	0.270
steal	0.830	0.327	0.253
steal	0.945	0.355	0.348
swig	0.536	0.508	0.001
swig	0.531	0.539	0.000
swig	0.431	0.422	0.000
swig	0.485	0.447	0.001
teach	0.994	0.931	0.004
teach	0.994	0.961	0.001
teach	0.700	0.805	0.011
teach	0.839	0.831	0.000
wash	0.859	0.856	0.000
wash	0.882	0.893	0.000
wash	0.842	0.728	0.013
wash	0.855	0.760	0.009
watch	0.903	0.843	0.004
watch	0.795	0.909	0.013
watch	0.738	0.682	0.003
watch	0.599	0.736	0.019
write	0.990	0.929	0.004
write	1.000	0.960	0.002
write	0.839	0.802	0.001
write	0.941	0.829	0.013

F.2 Italian

verb	actual	predicted	squared error
accoltellare	0.211	0.414	0.041
accoltellare	0.445	0.424	0.000
accoltellare	0.337	0.345	0.000
accoltellare	0.358	0.353	0.000
affettare	0.537	0.421	0.013
affettare	0.493	0.430	0.004
affettare	0.227	0.355	0.016
affettare	0.431	0.362	0.005
avvelenare	0.138	0.385	0.061
avvelenare	0.210	0.399	0.036

verb	actual	predicted	squared error
avvelenare	0.013	0.311	0.088
avvelenare	0.243	0.322	0.006
bere	0.916	0.892	0.001
bere	0.939	0.892	0.002
bere	0.933	0.798	0.018
bere	0.876	0.798	0.006
cantare	0.959	0.889	0.005
cantare	0.942	0.890	0.003
cantare	0.868	0.795	0.005
cantare	0.785	0.795	0.000
canticchiare	0.890	0.681	0.044
canticchiare	0.915	0.693	0.049
canticchiare	0.843	0.580	0.069
canticchiare	0.825	0.590	0.055
costruire	0.555	0.582	0.001
costruire	0.564	0.602	0.001
costruire	0.422	0.474	0.003
costruire	0.418	0.489	0.005
cucinare	0.919	0.866	0.003
cucinare	0.928	0.872	0.003
cucinare	0.926	0.762	0.027
cucinare	0.835	0.767	0.005
cucire	0.878	0.816	0.004
cucire	0.901	0.831	0.005
cucire	0.771	0.691	0.006
cucire	0.598	0.704	0.011
decapitare	0.155	0.410	0.065
decapitare	0.294	0.421	0.016
decapitare	0.155	0.340	0.034
decapitare	0.040	0.349	0.095
divorare	0.103	0.424	0.103
divorare	0.223	0.432	0.044
divorare	0.000	0.359	0.129
divorare	0.007	0.366	0.128
firmare	0.747	0.428	0.102
firmare	0.760	0.434	0.106
firmare	0.789	0.365	0.179
firmare	0.687	0.371	0.099
fumare	0.966	0.857	0.012
fumare	0.887	0.864	0.001
fumare	0.855	0.749	0.011
fumare	0.954	0.755	0.039
guardare	0.613	0.747	0.018
guardare	0.660	0.775	0.013
guardare	0.523	0.600	0.006
guardare	0.549	0.623	0.005
insegnare	0.853	0.830	0.001
insegnare	0.827	0.842	0.000
insegnare	0.545	0.711	0.028
insegnare	0.503	0.721	0.048
lavare	0.818	0.659	0.025
lavare	0.661	0.672	0.000

verb	actual	predicted	squared error
lavare	0.650	0.558	0.008
lavare	0.575	0.569	0.000
lucidare	0.657	0.546	0.012
lucidare	0.658	0.562	0.009
lucidare	0.509	0.448	0.004
lucidare	0.415	0.462	0.002
mangiare	0.912	0.862	0.002
mangiare	0.934	0.868	0.004
mangiare	0.876	0.755	0.015
mangiare	0.832	0.761	0.005
pulire	0.913	0.829	0.007
pulire	0.892	0.841	0.003
pulire	0.853	0.709	0.021
pulire	0.873	0.720	0.023
ricamare	0.880	0.663	0.047
ricamare	0.784	0.676	0.012
ricamare	0.805	0.563	0.059
ricamare	0.756	0.573	0.033
rompere	0.450	0.502	0.003
rompere	0.391	0.510	0.014
rompere	0.419	0.431	0.000
rompere	0.372	0.437	0.004
rubare	0.888	0.598	0.084
rubare	0.859	0.620	0.057
rubare	0.850	0.480	0.137
rubare	0.892	0.498	0.155
scarabocchiare	0.820	0.620	0.040
scarabocchiare	0.828	0.634	0.038
scarabocchiare	0.710	0.520	0.036
scarabocchiare	0.719	0.532	0.035
scrivere	0.895	0.877	0.000
scrivere	0.853	0.880	0.001
scrivere	0.661	0.777	0.013
scrivere	0.842	0.779	0.004
sorseggiare	0.640	0.707	0.005
sorseggiare	0.559	0.718	0.025
sorseggiare	0.378	0.606	0.052
sorseggiare	0.521	0.616	0.009
spaccare	0.378	0.399	0.000
spaccare	0.262	0.412	0.022
spaccare	0.157	0.327	0.029
spaccare	0.281	0.337	0.003
tagliare	0.563	0.807	0.059
tagliare	0.579	0.824	0.060
tagliare	0.408	0.679	0.074
tagliare	0.461	0.694	0.054
trangugiare	0.357	0.419	0.004
trangugiare	0.438	0.428	0.000
trangugiare	0.229	0.351	0.015
trangugiare	0.349	0.359	0.000
uccidere	0.593	0.534	0.003
uccidere	0.679	0.546	0.018

verb	actual	predicted	squared error
uccidere	0.604	0.449	0.024
uccidere	0.754	0.459	0.087
versare	0.299	0.859	0.313
versare	0.432	0.866	0.188
versare	0.165	0.751	0.343
versare	0.183	0.757	0.330

References

References

A

Abe, Naoki and Hang Li (1996). *Learning Word Association Norms Using Tree Cut Pair Models*. arXiv: cmp-lg/9605029 (cited on page 92).

Ahringberg, Johanna (2015). '"We Come up All This Way to Visit" A Case Study of Null Instantiation in English with the Verbs Visit and Destroy'. Term Paper. Lund University (cited on pages 14, 18, 23, 48).

Alatrash, Reem, Dominik Schlechtweg, Jonas Kuhn, and Sabine Schulte Im Walde (2020). 'CCOHA: Clean Corpus Of Historical American English'. In: *Proceedings of The 12th Language Resources and Evaluation Conference*, pp. 6958–6966 (cited on page 171).

Aldezabal, Izaskun, Koldo Gojenola, Kepa Sarasola, Aitziber Atutxa, et al. (2003). 'Learning argument/adjunct distinction for Basque'. In: *Anuario del Seminario de Filologıa Vasca "Julio de Urquijo"*, pp. 75–93. DOI: 10.1387/asju.9711 (cited on page 63).

Alexopoulou, Theodora and Frank Keller (2006). 'Gradience and Parametric Variation'. In: *ExLing 2006: 1st Tutorial and Research Workshop on Experimental Linguistics*, pp. 69–72 (cited on pages 62, 63).

Alishahi, Afra and Suzanne Stevenson (2007). 'A Cognitive Model for the Representation and Acquisition of Verb Selectional Preferences'. In: *Proceedings of the Workshop on Cognitive Aspects of Computational Language Acquisition*. Prague, Czech Republic: Association for Computational Linguistics, pp. 41–48 (cited on page 92).

Allen, Shanley E. M. (2000). 'A discourse-pragmatic explanation for argument representation in child Inuktitut'. In: *Linguistics* 38.3, pp. 483–521. DOI: 10.1515/ling.38.3.483 (cited on page 31).

Allerton, David J. (1975). 'Deletion and Proform Reduction'. In: *Journal of Linguistics* 11.2, pp. 213–237. DOI: 10.1017/S0022226700004540 (cited on pages 11, 12, 40, 49).

Almeida, Francisco Alonso (2009). 'Null Objects in Middle English Medical Texts'. In: *Textual Healing: Studies in Medieval English Medical, Scientific and Technical Texts*. Ed. by Javier E. Díaz-Vera and Rosario Caballero. Peter Lang, pp. 1–27 (cited on page 14).

Amberber, Mengistu (1996). 'Transitivity Alternations, Event-Types and Light Verbs'. PhD thesis. McGill University (cited on page 26).

– (2009). 'Quirky Alternation of Transitivity: The Case of Ingestive Predicates'. In: *The Linguistics of Eating and Drinking*. John Benjamins Publishing (cited on page 26).

AnderBois, Scott (2012). 'Indefiniteness and the Typology of Implicit Arguments'. In: *Proceedings of the 30th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project, pp. 43–53 (cited on pages 12, 15).

Antinucci, Francesco and Ruth Miller (1976). 'How children talk about what happened'. In: *Journal of child language* 3.2, pp. 167–189 (cited on page 44).

Armstrong, Grant Warren (2011). 'Two classes of transitive verbs: Evidence from Spanish'. PhD thesis. Georgetown University (cited on page 17).

Arunachalam, Sudha (2013). 'Experimental methods for linguists'. In: *Language and Linguistics Compass* 7.4, pp. 221–232 (cited on page 106).

B

Bach, Emmon (1986). 'The algebra of events'. In: *Linguistics and philosophy* 9.1, pp. 5–16 (cited on page 71).

Bader, Markus and Jana Häussler (2010). 'Toward a Model of Grammaticality Judgments'. In: *Journal of Linguistics* 46.2, pp. 273–330. DOI: 10.1017/S0022226709990260 (cited on pages 62, 169).

Baker, Mark (1988). 'Theta theory and the syntax of applicatives in Chichewa'. In: *Natural Language & Linguistic Theory* 6.3, pp. 353–389 (cited on page 24).

Bard, Ellen Gurman, Dan Robertson, and Antonella Sorace (1996). 'Magnitude estimation of linguistic acceptability'. In: *Language* 72.1, pp. 32–68. DOI: 10.2307/416793 (cited on pages 60, 106).

Baroni, Marco, Silvia Bernardini, Adriano Ferraresi, and Eros Zanchetta (2009). 'The WaCky wide web: a collection of very large linguistically processed web-crawled corpora'. In: *Language resources and evaluation* 43.3, pp. 209–226. DOI: 10.1007/s10579-009-9081-4 (cited on pages 96, 107).

Basile, Pierpaolo, Annalina Caputo, Tommaso Caselli, Pierluigi Cassotti, and Rossella Varvara (2020). 'A Diachronic Italian Corpus based on" L'Unità"'. In: *Proceedings of the 7th Italian Conference on Computational Linguistics, CLiC-it 2020*. Ed. by Felice Dell'Orletta, Johanna Monti, and Fabio Tamburini. Collana dell'Associazione Italiana di Linguistica Computazionale. Torino: Accademia University Press, pp. 31–36. DOI: 10.4000/books.aaccademia.8203 (cited on pages 171, 172).

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker (2015). 'Fitting Linear Mixed-Effects Models Using lme4'. In: *Journal of Statistical Software* 67.1, pp. 1–48. DOI: 10.18637/jss.v067.i01 (cited on pages 121, 161).

Beavers, John (2013). 'Aspectual Classes and Scales of Change'. In: *Linguistics* 51.4. DOI: 10.1515/ling-2013-0024 (cited on page 34).

Beavers, John and Andrew Koontz-Garboden (2012). 'Manner and Result in the Roots of Verbal Meaning'. In: *Linguistic Inquiry* 43.3, pp. 331–369 (cited on pages 34, 35).

– (2017). 'Result Verbs, Scalar Change, and the Typology of Motion Verbs'. In: *Language* 93.4, pp. 842–876. DOI: 10.1353/lan.2017.0060 (cited on page 34).

Beavers, John, Beth Levin, and Shiao Wei Tham (2010). 'The Typology of Motion Expressions Revisited'. In: *Journal of Linguistics* 46.2, pp. 331–377. DOI: 10.1017/S0022226709990272 (cited on page 34).

Beckwith, Richard, Christiane Fellbaum, Derek Gross, and George A. Miller (1991). 'WordNet: A lexical database organized on psycholinguistic principles'. In: *Lexical acquisition: Exploiting on-line resources to build a lexicon* 211 (cited on pages 68, 91).

Bellik, Jennifer and Nick Kalivoda (2019). 'Automated tableau generation using SPOT (Syntax Prosody in Optimality Theory)'. In: *Linguistics Vanguard* 5.1 (cited on page 53).

Bender, Emily (1999). 'Constituting Context: Null Objects in English Recipes Revisited'. In: *University of Pennsylvania Working Papers in Linguistics* 6.1 (cited on page 14).

Bergh, Gunnar and Sölve Ohlander (2016). 'Iniesta Passed and Messi Finished Clinically: Football Verbs and Transitivity'. In: *Nordic Journal of English Studies* 15.2, p. 19. DOI: 10.35360/njes.359 (cited on pages 14, 15, 31).

Bergsma, Shane, Dekang Lin, and Randy Goebel (2008). 'Discriminative Learning of Selectional Preference from Unlabeled Text'. In: *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*. EMNLP 2008. Honolulu, Hawaii: Association for Computational Linguistics, pp. 59–68. DOI: 10.5555/1613715.1613725 (cited on page 92).

Bermel, Neil and Luděk Knittl (2012). 'Corpus Frequency and Acceptability Judgments: A Study of Morphosyntactic Variants in Czech'. In: *Corpus Linguistics and Linguistic Theory* 8.2, pp. 241–275. DOI: 10.1515/cllt-2012-0010 (cited on page 169).

Bertinetto, Pier Marco (1992). 'Le strutture tempo-aspettuali dell'italiano e dell'inglese'. In: *L'Europa linguistica: contatti, contrasti, e affinità di lingue*. Ed. by Antonia G. Mocciaro and Giulio Soravia. Roma: Bulzoni, pp. 49–68 (cited on page 103).

– (2001). 'On a frequent misunderstanding in the temporal-aspectual domain: The 'perfective-telic confusion''. In: *Semantic Interfaces: Reference, Anaphora and Aspect, Stanford: CSLI Publications*, pp. 177–210 (cited on pages 44, 102, 173).

Bertinetto, Pier Marco and Denis Delfitto (2000). 'Aspect vs. Actionality: Why they should be kept apart'. In: *Empirical approaches to language typology* 6, pp. 189–226 (cited on pages 44, 100, 173).

Bertinetto, Pier Marco, Eva Maria Freiberger, Alessandro Lenci, Sabrina Noccetti, and Maddalena Agonigi (2015). 'The acquisition of tense and aspect in a morphology-sensitive framework: Data from Italian and Austrian-German children'. In: *Linguistics* 53.5, pp. 1113–1168. DOI: 10.1515/ling-2015-0030 (cited on page 45).

Bertinetto, Pier Marco and Alessandro Lenci (2012). 'Habituality, pluractionality and imperfectivity'. In: *Oxford Handbook of Tense and Aspect*. Ed. by Robert I. Binnick. Oxford: Oxford University Press, pp. 852–880. DOI: 10.1093/oxfordhb/9780195381979.013.0030 (cited on page 47).

Bertinetto, Pier Marco and Anna Lentovskaya (2012). 'A diachronic view of the actional/aspectual properties of Russian verbs'. In: *Russian linguistics* 36.1, pp. 1–19 (cited on page 173).

Bertinetto, Pier Marco, Clémentine Talaato Pacmogda, and Alessandro Lenci (2021). 'On the acquisition of verbal tenses in Mòoré (Gur): a morphology-based approach'. In: *Lingue e linguaggio* 20.1, pp. 111–160. DOI: 10.1418/101115 (cited on page 45).

Bertinetto, Pier Marco and Mario Squartini (1996). 'La distribuzione del Perfetto Semplice e del Perfetto Composto nelle diverse varietà di italiano'. In: *Romance Philology* 49.4, pp. 383–419 (cited on page 103).

Bhatt, Rajesh and Roumyana Pancheva (2017). 'Implicit arguments'. In: *The Wiley Blackwell Companion to Syntax, Second Edition*, pp. 1–35 (cited on page 170).

Bod, Rens, Jennifer Hay, and Stefanie Jannedy (2003). *Probabilistic Linguistics*. Mit Press (cited on page 62).

Boersma, Paul (2004). *A Stochastic OT Account of Paralinguistic Tasks Such as Grammaticality and Prototypicality Judgments* (cited on pages 63, 168, 169).

Boersma, Paul et al. (1997). 'How we learn variation, optionality, and probability'. In: *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam*. Vol. 21, pp. 43–58 (cited on pages 63, 64).

Boersma, Paul and Bruce Hayes (2001). 'Empirical tests of the gradual learning algorithm'. In: *Linguistic inquiry* 32.1, pp. 45–86 (cited on pages 63–65, 168).

Borik, Olga (2006). 'Main Theories of Aspect (I): The Telicity Approach'. In: *Aspect and Reference Time*. Oxford University Press. DOI: 10.1093/acprof:oso/9780199291298.001.0001 (cited on page 101).

Bornkessel-Schlesewsky, Ina and Matthias Schlesewsky (2007). 'The Wolf in Sheep's Clothing: Against a New Judgement-Driven Imperialism'. In: *Theoretical Linguistics* 33.3. DOI: 10.1515/TL.2007.021 (cited on pages 60, 62).

Bourmayan, Anouch and François Recanati (2013). 'Transitive Meanings for Intransitive Verbs'. In: *Brevity*. Ed. by Laurence Goldstein. Oxford University Press, pp. 122–142. DOI: 10.1093/acprof:oso/9780199664986.003.0008 (cited on pages 10, 21, 23).

Brehm, Laurel and Matthew Goldrick (2017). 'Distinguishing Discrete and Gradient Category Structure in Language: Insights from Verb-Particle Constructions.' In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 43.10, pp. 1537–1556. DOI: 10.1037/xlm0000390 (cited on page 62).

Bresnan, Joan (1978). 'A realistic transformational grammar'. In: *Linguistic theory and psychological reality*. Ed. by Morris Halle, Joan Bresnan, and George A. Miller. Cambridge, MA: MIT Press, pp. 1–59 (cited on page 19).

Bresnan, Joan, Sam Featherston, and Wolfgang Sternefeld (2007). 'Is syntactic knowledge probabilistic? Experiments with the English dative alternation'. In: *Roots: Linguistics in search of its evidential base* 96, pp. 77–96 (cited on page 62).

Bresnan, Joan and Jennifer Hay (2008). 'Gradient Grammar: An Effect of Animacy on the Syntax of Give in New Zealand and American English'. In: *Lingua* 118.2, pp. 245–259. DOI: 10.1016/j.lingua.2007.02.007 (cited on page 62).

Bresnan, Joan and Tatiana Nikitina (2008). 'The gradience of the dative alternation'. In: *Reality exploration and discovery: Pattern interaction in language and life*, pp. 161–184 (cited on page 62).

Brisson, Christine (1994). 'The licensing of unexpressed objects in English verbs'. In: *28th Regional Meeting of the Chicago Linguistic Society (CLS)*. Vol. 1, pp. 90–102 (cited on page 21).

Brockmann, Carsten and Mirella Lapata (2003). 'Evaluating and Combining Approaches to Selectional Preference Acquisition'. In: *Proceedings of the Tenth Conference on European Chapter of the Association for Computational Linguistics - EACL '03*. Vol. 1. Budapest, Hungary: Association for Computational Linguistics, p. 27. DOI: 10.3115/1067807.1067813 (cited on page 92).

Bross, Fabian (2019). *Acceptability ratings in linguistics: a practical guide to grammaticality judgments, data collection, and statistical analysis* (cited on page 121).

Brysbaert, Marc, Paweł Mandera, and Emmanuel Keuleers (2018). 'The word frequency effect in word processing: An updated review'. In: *Current Directions in Psychological Science* 27.1, pp. 45–50 (cited on page 106).
Buchwald, Adam, Oren Schwartz, Amanda Seidl, and Paul Smolensky (2002). 'Recoverability Optimality Theory: Discourse anaphora in a bidirectional framework'. In: *Proceedings of the 6th International workshop on formal semantics and pragmatics of dialogue*, pp. 37–44 (cited on page 72).

C

Cappelli, Giulia, Pier Marco Bertinetto, and Alessandro Lenci (2019). 'On the argumenthood of optional PPs with Italian motion verbs'. In: *Proceedings of 10th International Conference of Experimental Linguistics*. Ed. by Antonis Botinis. Vol. 25. Athens, Greece: ExLing Society, pp. 45–48 (cited on page 38).
Cappelli, Giulia and Alessandro Lenci (2020). 'PISA: A measure of Preference In Selection of Arguments to model verb argument recoverability'. In: *Proceedings of the Ninth Joint Conference on Lexical and Computational Semantics*. Barcelona, Spain (Online): Association for Computational Linguistics, pp. 131–136 (cited on pages vii, 2, 4, 6, 33, 68, 74, 89, 90, 92, 94–97, 129, 167, 170).
Carnie, Andrew (2012). *Syntax: A generative introduction*. John Wiley & Sons (cited on pages 18, 24).
Cennamo, Michela (2017). 'Object Omission and the Semantics of Predicates in Italian in a Comparative Perspective'. In: *Linguistik Aktuell/Linguistics Today*. Ed. by Lars Hellan, Andrej L. Malchukov, and Michela Cennamo. Vol. 237. Amsterdam: John Benjamins Publishing Company, pp. 252–273. DOI: 10.1075/la.237.08cen (cited on pages 15, 43).
Cennamo, Michela and Alessandro Lenci (2019). 'Gradience in Subcategorization? Locative Phrases with Italian Verbs of Motion'. In: *Studia Linguistica* 73.2, pp. 369–397. DOI: 10.1111/stul.12095 (cited on pages 34, 38, 63).
Chomsky, Noam (1957). *Syntactic Structures*. Walter de Gruyter (cited on page 62).
– (1981). *Lectures on government and binding: The Pisa lectures*. Walter de Gruyter (cited on page 55).
– (1982). *Some concepts and consequences of the theory of government and binding*. MIT press (cited on pages 22, 54).
– (1991). 'Some Notes on Economy of Derivation and Representation'. In: *Principles and Parameters in Comparative Grammar*. Ed. by Robert Freidin. Cambridge: The MIT Press (cited on page 54).
– (1993). 'A minimalist program for linguistic theory'. In: *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger* (cited on page 61).
Ciaramita, Massimiliano and Mark Johnson (2000). 'Explaining Away Ambiguity: Learning Verb Selectional Preference with Bayesian Networks'. In: *COLING 2000 Volume 1: The 18th International Conference on Computational Linguistics*. COLING 2000 (cited on page 92).

Civardi, Eugenio and Pier Marco Bertinetto (2015). 'The semantics of degree verbs and the telicity issue'. In: *Borealis–An International Journal of Hispanic Linguistics* 4.1, pp. 57–77 (cited on page 44).

Clark, Stephen and David Weir (2001). 'Class-Based Probability Estimation Using a Semantic Hierarchy'. In: *Second Meeting of the North American Chapter of the Association for Computational Linguistics*. NAACL 2001 (cited on page 92).

Comrie, Bernard (1976). *Aspect: An introduction to the study of verbal aspect and related problems*. Vol. 2. Cambridge university press (cited on pages 38, 42, 45, 102).

– (1989). *Language universals and linguistic typology: Syntax and morphology*. University of Chicago press (cited on page 9).

Condoravdi, Cleo and Jean Mark Gawron (1996). 'The Context-Dependency of Implicit Arguments'. In: *Quantifiers, Deduction, and Context*. Stanford, CA: CSLI Publications, pp. 1–32 (cited on page 18).

Copley, Bridget and Heidi Harley (2015). 'A force-theoretic framework for event structure'. In: *Linguistics and Philosophy* 38.2, pp. 103–158 (cited on page 41).

Cote, Sharon Ann (1996). 'Grammatical and Discourse Properties of Null Arguments in English'. PhD thesis. University of Pennsylvania (cited on pages 13, 14, 16, 21, 30, 31, 42, 48).

Crocker, Matthew and Frank Keller (2006). 'Probabilistic Grammars as Models of Gradience in Language Processing'. In: *Gradience in Grammar: Generative Perspectives* (cited on page 62).

Culy, Christopher (1996). 'Null Objects in English Recipes'. In: *Language Variation and Change* 8.1, pp. 91–124. DOI: 10.1017/S0954394500001083 (cited on pages 13, 14).

Cummins, Sarah and Yves Roberge (2004). 'Null Objects in French and English'. In: *Current Issues in Linguistic Theory*. Ed. by Julie Auger, J. Clancy Clements, and Barbara Vance. Vol. 258. Amsterdam: John Benjamins Publishing Company, pp. 121–138. DOI: 10.1075/cilt.258.07cum (cited on pages 12, 22).

– (2005). 'A Modular Account of Null Objects in French'. In: *Syntax* 8.1, pp. 44–64. DOI: 10.1111/j.1467-9612.2005.00074.x (cited on pages 15, 22, 49).

Cunnings, Ian (2012). 'An overview of mixed-effects statistical models for second language researchers'. In: *Second Language Research* 28.3, pp. 369–382 (cited on page 121).

D

David, Oana Alexandra (2016). 'Metaphor in the Grammar of Argument Realization'. PhD thesis. University of California, Berkeley. 207 pp. (cited on page 11).

Davidson, Lisa and Matthew Goldrick (2003). 'Tense, agreement and defaults in child Catalan: An Optimality Theoretic analysis'. In: *Linguistic theory and language development in Hispainic languages*, pp. 193–211 (cited on page 63).

Davies, Mark (2010). 'The Corpus of Contemporary American English as the first reliable monitor corpus of English'. In: *Literary and linguistic computing* 25.4, pp. 447–464. DOI: 10.1093/llc/fqq018 (cited on page 171).

– (2012). 'Expanding horizons in historical linguistics with the 400-million word Corpus of Historical American English'. In: *Corpora* 7.2, pp. 121–157 (cited on page 171).

De Smet, Hendrik (2005). 'A corpus of Late Modern English texts'. In: *Icame Journal* 29, pp. 69–82 (cited on page 171).

De Smet, Hendrik, Susanne Flach, Jukka Tyrkkö, and Hans-Jürgen Diller (2015). *Corpus of Late Modern English texts (version 3.1)* (cited on page 171).

DeLancey, Scott (1987). 'Transitivity in grammar and cognition'. In: *Coherence and grounding in discourse*. Typological studies in language. Amsterdam: John Benjamins, pp. 53–68 (cited on page 45).

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. DOI: 10.48550/ARXIV.1810.04805. URL: https://arxiv.org/abs/1810.04805 (cited on page 98).

Diesing, Molly (1992). *Indefinites*. Vol. 20. Linguistic Inquiry Monographs. MIT Press, pp. xiv + 175 (cited on page 61).

Dixon, Robert M. W. (1992). *A New Approach to English Grammar, on Semantic Principles*. Clarendon Press (cited on pages 45, 48).

Dowty, David Roach (1979 [2012]). *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*. Vol. 7. Studies in Linguistics and Philosophy. Springer Science & Business Media (cited on pages 41, 69, 70).

– (1981). 'Quantification and the lexicon: A reply to Fodor and Fodor'. In: *The scope of lexical rules*, pp. 79–106 (cited on page 21).

– (1991). 'Thematic proto-roles and argument selection'. In: *Language* 67.3, pp. 547–619 (cited on pages 17, 39).

– (2003). 'The Dual Analysis of Adjuncts/Complements in Categorial Grammar'. In: *Modifying Adjuncts*. Ed. by Ewald Lang, Claudia Maienborn, and Cathrine Fabricius-Hansen. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110894646.33 (cited on page 63).

Dvořák, Věra (2017a). 'A Syntactic Approach to Indefinite Null Objects'. In: *NELS 48* (cited on pages 17, 18).

– (2017b). 'Generic and Indefinite Null Objects'. PhD Thesis. Rutgers University New Brunswick (cited on pages 14, 19, 23, 41, 42, 44).

E

Ebeling, Signe Oksefjell (2021). 'To Score or to Score a Goal: Transitivity in Football Match Reports'. In: *English Studies* 102.2, pp. 243–266 (cited on page 14).

Endresen, Anna and Laura A. Janda (2017). 'Five Statistical Models for Likert-Type Experimental Data on Acceptability Judgments'. In: *Journal of Research Design and Statistics in Linguistics and Communication Science* 3.2, pp. 217–250. DOI: 10.1558/jrds.30822 (cited on page 121).

Engelberg, Stefan (2002). 'Intransitive Accomplishments and the Lexicon: The Role of Implicit Arguments, Definiteness, and Reflexivity in Aspectual Composition'. In: *Journal of Semantics* 19.4, pp. 369–416. DOI: 10.1093/jos/19.4.369 (cited on page 8).

Erk, Katrin (2007). 'A Simple, Similarity-Based Model for Selectional Preferences'. In: *Proceedings of the 45th annual meeting of the Association of Computational Linguistics*, pp. 216–223 (cited on page 94).

Erk, Katrin, Sebastian Padó, and Ulrike Padó (2010). 'A Flexible, Corpus-Driven Model of Regular and Inverse Selectional Preferences'. In: *Computational Linguistics* 36.4, pp. 723–763. DOI: 10.1162/coli_a_00017 (cited on page 94).

Erlewine, Michael Yoshitaka and Hadas Kotek (2016). 'A Streamlined Approach to Online Linguistic Surveys'. In: *Natural Language & Linguistic Theory* 34.2, pp. 481–495. DOI: 10.1007/s11049-015-9305-9 (cited on page 105).

Eu, Jinseung (2018). 'On the Nature of Object Omission: Indefiniteness as Indeterminacy'. In: *English Language and Linguistics* 22.3, pp. 523–530. DOI: 10.1017/S1360674317000296 (cited on pages 11, 13, 46).

Evert, Stefan (2005). 'The statistics of word cooccurrences: word pairs and collocations'. PhD thesis. Universität Stuttgart (cited on page 95).

F

Fellbaum, Christiane and Judy Kegl (1989). 'Taxonomic structures and cross-category linking in the lexicon'. In: *Proceedings of the Sixth Eastern State Conference on Linguistics*. Ohio State University. Columbus, Ohio, pp. 93–104 (cited on pages 21, 34–36, 46, 49).

Ferraresi, Adriano, Eros Zanchetta, Marco Baroni, and Silvia Bernardini (2008). 'Introducing and Evaluating Ukwac, a Very Large Web-Derived Corpus of English'. In: *In Proceedings of the 4th Web as Corpus Workshop (WAC-4* (cited on page 96).

Filip, Hana (2004). 'The telicity parameter revisited'. In: *Semantics and Linguistic Theory*. Vol. 14, pp. 92–109 (cited on page 39).

Fillmore, Charles J. (1969). 'Types of lexical information'. In: *Studies in syntax and semantics*. Springer, pp. 109–137 (cited on pages 11, 17, 20, 157).

– (1986). 'Pragmatically Controlled Zero Anaphora'. In: *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society (1986)*, pp. 95–107 (cited on pages vii, 1, 8, 11–13, 15–17, 19, 23, 27, 48, 49).

Firth, John R. (1957). 'A synopsis of linguistic theory, 1930-1955'. In: *Studies in linguistic analysis* (cited on page 92).

Fodor, Jerry and Janet Fodor (1980). 'Functional structure, quantifiers, and meaning postulates'. In: *Linguistic Inquiry* 11.4, pp. 759–770 (cited on pages 18, 21).

Fraser, Bruce and John Robert Ross (1970). 'Idioms and unspecified NP deletion'. In: *Linguistic Inquiry* 1.2, pp. 264–265 (cited on page 17).

G

Gaeta, Livio, Claudio Iacobini, Davide Ricca, Marco Angster, Aurelio De Rosa, and Giovanna Schirato (2013). 'Midia: a balanced diachronic corpus of italian'. In: *21st International Conference on Historical Linguistics, Oslo* (cited on page 171).

García-Velasco, Daniel and Carmen Portero Muñoz (2002). 'Understood Objects in Functional Grammar'. In: *Working papers in functional grammar* 76, p. 25 (cited on pages 12, 14, 23, 32, 35, 36, 45, 49).

Gibson, Edward, Steve Piantadosi, and Kristina Fedorenko (2011). 'Using Mechanical Turk to obtain and analyze English acceptability judgments'. In: *Language and Linguistics Compass* 5.8, pp. 509–524 (cited on pages 105, 121).

Gillon, Brendan S. (2006). *English Relational Words, Context Sensitivity and Implicit Arguments*. Manuscript. URL: https://semanticsarchive.net/Archive/jk5ZjU1O/implicit-argument.pdf (cited on pages 18, 170).

– (2011). 'French relational words, context sensitivity and implicit arguments'. In: *Current Research in the Semantics-Pragmatics Interface (Making Semantics Pragmatic)* 24, pp. 143–164 (cited on page 18).

– (2012). 'Implicit Complements: A Dilemma for Model Theoretic Semantics'. In: *Linguistics and Philosophy* 35.4, pp. 313–359. DOI: 10.1007/s10988-012-9120-2 (cited on pages 16, 18, 21).

Glass, Lelia (2013). 'What Does It Mean for an Implicit Object to Be Recoverable?' In: *University of Pennsylvania Working Papers in Linguistics* 20.1 (cited on pages 3, 15, 16, 30–32, 48, 89, 103).

– (2020). 'Verbs Describing Routines Facilitate Object Omission in English'. In: *Proceedings of the Linguistic Society of America* 5.1, p. 44. DOI: 10.3765/plsa.v5i1.4663 (cited on pages 3, 28, 30–32, 45, 46, 49, 50, 103, 170).

– (2022). 'English verbs can omit their objects when they describe routines'. In: *English Language & Linguistics* 26.1, pp. 49–73 (cited on pages 3, 30, 31).

Goldberg, Adele E. (2001). 'Patient Arguments of Causative Verbs Can Be Omitted: The Role of Information Structure in Argument Distribution'. In: *Language Sciences* 23, p. 22 (cited on pages 13, 34, 35, 40, 46, 48, 49, 170, 171).

– (2005a). 'Argument Realization: The Role of Constructions, Lexical Semantics and Discourse Factors'. In: *Constructional Approaches to Language*. Ed. by Jan-Ola Östman and Mirjam Fried. Vol. 3. Amsterdam: John Benjamins Publishing Company, pp. 17–43. DOI: 10.1075/cal.3.03gol (cited on pages 19, 23, 31, 49, 152).

– (2005b). 'Constructions, Lexical Semantics and the Correspondence Principle: Accounting for Generalizations and Subregularities in the Realization of Arguments'. In: *The Syntax of Aspect: Deriving Thematic and Aspectual Interpretation*. Erteschik-Shir, Nomi and Rapoport, Tova (cited on pages 45, 49, 50, 170).

– (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford Linguistics. Oxford University Press (cited on pages 48, 49).

Greene, Stephan and Philip Resnik (2009). 'More than words: Syntactic packaging and implicit sentiment'. In: *Proceedings of human language technologies: The 2009 nnual conference of the North American chapter of the Association for Computational Linguistics*, pp. 503–511 (cited on page 23).

Grimshaw, Jane and Vieri Samek-Lodovici (1998). 'Optimal subjects and subject universals'. In: *Is the best good enough? Optimality and Competition in Syntax*. Ed. by Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky. MIT Press, pp. 193–219 (cited on page 53).

Grishman, Ralph and John Sterling (1992). 'Acquisition of Selectional Patterns'. In: *COLING 1992 Volume 2: The 15th International Conference on Computational Linguistics*. COLING 1992 (cited on page 92).

Groefsema, Marjolein (1995). 'Understood arguments: A semantic/pragmatic approach'. In: *Lingua* 96.2-3, pp. 139–161 (cited on pages 13, 23, 49).

Grootswagers, Tijl (2020). 'A primer on running human behavioural experiments online'. In: *Behavior Research Methods* 52.6, pp. 2283–2286 (cited on pages 105, 106).

H

Haagsma, Hessel and Johannes Bjerva (2016). 'Detecting Novel Metaphor Using Selectional Preference Information'. In: *Proceedings of the Fourth Workshop on Metaphor in NLP*. Proceedings of the Fourth Workshop on Metaphor in NLP. San Diego, California: Association for Computational Linguistics, pp. 10–17. DOI: 10.18653/v1/W16-1102 (cited on page 92).

Haegeman, Liliane (1987). 'Register Variation in English: Some Theoretical Observations'. In: *Journal of English Linguistics* 20.2, pp. 230–248. DOI: 10.1177/007542428702000207 (cited on pages 13, 16).

Hall, Alison Margaret (2009). 'Free pragmatic processes and explicit utterance content'. PhD thesis. University of London (cited on page 23).

Hartkemeyer, Dale (2000). 'An OT approach to atonic vowel loss patterns in Old French and Old Spanish'. In: *New Approaches to Old Problems: Issues in Romance Historical Linguistics*. Ed. by Steven Norman Dworkin and Dieter Wanner, pp. 65–84 (cited on page 72).

Haspelmath, Martin (1994). 'Passive participles across languages'. In: *Voice: Form and function* 27 (cited on page 26).

Hayes, Bruce, Bruce Tesar, and Kie Zuraw (2003). *OTSoft 2.1, Software package*. URL: http://www.linguistics.ucla.edu/people/hayes/otsoft/ (cited on page 53).

Hickman, Louis, Julia Taylor, and Victor Raskin (2016). 'Direct Object Omission as a Sign of Conceptual Defaultness'. In: *Proceedings of the Twenty-Ninth International Florida Artificial Intelligence Research Society Conference*, pp. 516–521 (cited on pages 30, 31).

Hilpert, Martin and Stefan Th. Gries (2016). 'Quantitative approaches to diachronic corpus linguistics'. In: *The Cambridge handbook of English historical linguistics*. Ed. by Merja Kytö and Päivi Pahta. Cambridge University Press, pp. 36–53. DOI: 10.1017/CBO9781139600231.003 (cited on page 171).

Hopper, Paul J. and Sandra A. Thompson (1980). 'Transitivity in Grammar and Discourse'. In: *Language* 56.2, p. 251. DOI: 10.2307/413757 (cited on pages vii, 1, 8–10, 15, 17, 25, 27, 33, 38, 40, 43, 44, 110, 154, 168).

Huddleston, Rodney, Rodney D. Huddleston, Geoffrey K. Pullum, and Laurie Bauer (2002). *The Cambridge Grammar of the English Language*. Cambridge University Press (cited on pages 8, 12, 19, 20).

I

Iacobini, Claudio, Aurelio De Rosa, and Giovanna Schirato (2014). 'Part-of-Speech tagging strategy for MIDIA: a diachronic corpus of the Italian language'. In: *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 & and of the Fourth International Workshop*

EVALITA 2014: 9-11 December 2014, Pisa. Pisa: Pisa University Press, pp. 213–218 (cited on page 171).

Ikegami, Yoshihiko (1991). ''DO-language' and 'BECOME-language': Two contrasting types of linguistic representation'. In: *The empire of signs: Semiotic essays on Japanese culture* 8, pp. 285–327 (cited on page 172).

Ingham, Richard (1993). 'Input and Learnability: Direct-Object Omissibility in English'. In: *Language Acquisition* 3.2, pp. 95–120. DOI: 10.1207/s15327817la0302_1 (cited on page 31).

Isingoma, Bebwa (2020). 'Implicit Arguments in English and Rutooro'. In: *Linguistik Online* 101.1, pp. 19–47. DOI: 10.13092/LO.101.6671 (cited on page 26).

Iten, Corinne, Marie-Odile Junker, Aryn Pyke, Robert Stainton, and Catherine Wearing (2005). 'Null complements: Licensed by syntax or by semantics-pragmatics?' In: *Proceedings of the Annual Conference of the Canadian Linguistic Association*, pp. 1–15 (cited on pages 20, 23).

Iwata, Seizi (2002). 'Does MANNER Count or Not? Manner-of-Motion Verbs Revisited'. In: *Linguistics* 40.1. DOI: 10.1515/ling.2002.008 (cited on page 34).

J

Jackendoff, Ray (2003). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford University Press (cited on pages 16, 37, 172).

Jespersen, Otto (1927). *A Modern English Grammar on Historical Principles. Part III. Syntax*. Vol. 2. Routledge (cited on page 30).

Juzek, Tom S. (2016). 'Acceptability Judgement Tasks and Grammatical Theory'. PhD thesis. University of Oxford (cited on pages 112, 113).

Juzek, Tom S. and Jana Häussler (2019). 'Semantic Influences on Syntactic Acceptability Ratings'. In: *Proceedings of Linguistic Evidence 2018: Experimental Data Drives Linguistic Theory*. Ed. by Anstatt Gattnar, Robin Hörnig, Melanie Störzer, and Sam Featherston. Tübingen: University of Tübingen, pp. 341–355 (cited on page 60).

K

Kardos, Eva (2010). 'The Argument Expression of Change-of-State Verbs and Pseudo-Transitive Verbs'. In: *Bergen Language and Linguistics Studies* 1. DOI: 10.15845/bells.v1i1.50 (cited on pages 9, 17, 31).

Katz, Jerrold J. and Paul M. Postal (1967). 'An integrated theory of linguistic description'. In: *Synthese* 17.1 (cited on page 17).

Keller, Frank (1997). 'Extraction, gradedness, and optimality'. In: *University of Pennsylvania Working Papers in Linguistics* 4.2, p. 12 (cited on pages 60, 61, 64, 167).

– (1998a). 'Gradient grammaticality as an effect of selective constraint re-ranking'. In: *Papers from the 34th meeting of the Chicago Linguistic Society*. Vol. 2, pp. 95–109 (cited on page 63).

– (1998b). 'Grammaticality Judgments and Linguistic Methodology'. In: *Research Paper EUCCS-RP-1998-3*, p. 16 (cited on page 62).

– (2000). 'Gradience in Grammar: Experimental and Computational Aspects of Degrees of Grammaticality'. PhD Thesis. University of Edinburgh (cited on pages 59, 63, 161).

Keller, Frank (2006). 'Linear Optimality Theory as a Model of Gradience in Grammar'. In: *Gradience in Grammar: Generative Perspectives* (cited on pages 59, 63, 161).

Keller, Frank and Ash Asudeh (2002). 'Probabilistic Learning Algorithms and Optimality Theory'. In: *Linguistic Inquiry* 33.2, pp. 225–244. DOI: 10.1162/002438902317406704 (cited on page 169).

Keller, Frank and Maria Lapata (1998). 'Object Drop and Discourse Accessibility'. In: *Proceedings of the 17th West Coast Conference on Formal Linguistics*. Stanford, CA: CSLI Publications, pp. 362–374 (cited on page 12).

Keller, Frank and Antonella Sorace (2003). 'Gradient Auxiliary Selection and Impersonal Passivization in German: An Experimental Investigation'. In: *Journal of Linguistics* 39.1, pp. 57–108 (cited on page 62).

Kemmer, Suzanne (1993). *The Middle Voice*. John Benjamins Publishing Company (cited on page 9).

Kempen, Gerard and Karin Harbusch (2005). 'The Relationship between Grammaticality Ratings and Corpus Frequencies: A Case Study into Word Order Variability in the Midfield of German Clauses'. In: *Linguistic Evidence*. Ed. by Henk van Riemsdijk, Harry van der Hulst, Jan Koster, Stephan Kepser, and Marga Reis. Vol. 85. Berlin, New York: Mouton de Gruyter, pp. 329–350. DOI: 10.1515/9783110197549.329 (cited on pages 50, 169, 170).

– (2008). 'Comparing Linguistic Judgments and Corpus Frequencies as Windows on Grammatical Competence: A Study of Argument Linearization in German Clauses'. In: *The Discourse Potential of Underspecified Structures*. Ed. by Anita Steube. Language, context, and cognition. Berlin: Walter de Gruyter. DOI: 10.1515/9783110209303.3.179 (cited on page 169).

Kim, Najoung, Kyle Rawlins, and Paul Smolensky (2018). *A Gradient Blend Analysis of English PP Verbal Dependents* (cited on pages 63, 99, 116).

– (2019). *The Complement-Adjunct Distinction As Gradient Blends: The Case Of English Prepositional Phrases* (cited on pages 63, 99, 116).

Kim, Najoung, Kyle Rawlins, Benjamin Van Durme, and Paul Smolensky (2019). *Predicting the Argumenthood of English Prepositional Phrases*. arXiv: 1809.07889 [cs] (cited on pages 63, 99, 116).

Kizach, Johannes (2014). *Analyzing Likert-Scale Data with Mixed-Effects Linear Models - a Simulation Study*. Poster Presented at Linguistic Evidence (cited on page 121).

Koenig, Jean-Pierre, Gail Mauner, and Breton Bienvenue (2002). 'Class Specificity and the Lexical Encoding of Participant Information'. In: *Brain and Language* 81.1-3, pp. 224–235. DOI: 10.1006/brln.2001.2519 (cited on page 170).

– (2003). 'Arguments for Adjuncts'. In: *Cognition* 89.2, pp. 67–103. DOI: 10.1016/S0010-0277(03)00082-9 (cited on page 170).

Koenig, Jean-Pierre, Gail Mauner, Breton Bienvenue, and Kathy Conklin (2007). 'What with? The Anatomy of a (Proto)-Role'. In: *Journal of Semantics* 25.2, pp. 175–220. DOI: 10.1093/jos/ffm013 (cited on page 170).

Korkiakangas, Timo (2018). 'Verso l'analisi della transitività dei generi testuali latini: il caso del latino notarile'. In: *Studi e Saggi Linguistici* 56.1, pp. 9–41 (cited on page 14).

Kučera, Henry and Winthrop Nelson Francis (1967). *Computational analysis of present-day American English*. University Press of New England (cited on page 69).

Kuhn, Jonas (2002). 'Corpus-based Learning in Stochastic OT-LFG–Experiments with a Bidirectional Bootstrapping Approach'. In: *The LFG 02 Conference*. Citeseer, pp. 239–257 (cited on pages 55, 58, 59, 167).

L

Landau, Idan (2010). 'The Explicit Syntax of Implicit Arguments'. In: *Linguistic Inquiry* 41.3, pp. 357–388. DOI: 10.1162/LING_a_00001 (cited on page 22).

Langsford, Steven, Amy Perfors, Andrew T. Hendrickson, Lauren A. Kennedy, and Danielle J. Navarro (2018). 'Quantifying Sentence Acceptability Measures: Reliability, Bias, and Variability'. In: *Glossa: A journal of general linguistics* 3.1, p. 37. DOI: 10.5334/gjgl.396 (cited on page 106).

Larson, Pär, Elena Artale, and Diego Dotto (2005). *Corpus OVI dell'Italiano antico*. Firenze, Istituto Opera del Vocabolario Italiano. URL: http://gattoweb.ovi.cnr.it/ (cited on page 171).

Lasersohn, Peter (1995 [2013]). *Plurality, conjunction and events*. Vol. 55. Springer Science & Business Media (cited on page 47).

– (1993). 'Lexical Distributivity and Implicit Arguments'. In: *Semantics and Linguistic Theory* 3, p. 145. DOI: 10.3765/salt.v3i0.2751 (cited on pages 18, 170).

Lau, Jey Han, Alexander Clark, and Shalom Lappin (2017). 'Grammaticality, Acceptability, and Probability: A Probabilistic View of Linguistic Knowledge'. In: *Cognitive Science* 41.5, pp. 1202–1241. DOI: 10.1111/cogs.12414 (cited on page 60).

Lavidas, Nikolaos (2013). 'Null and Cognate Objects and Changes in (in)Transitivity: Evidence from the History of English'. In: *Acta Linguistica Hungarica* 60.1, pp. 69–106. DOI: 10.1556/ALing.60.2013.1.2 (cited on page 47).

Lazard, Gilbert (2002). 'Transitivity Revisited as an Example of a More Strict Approach in Typological Research'. In: *Folia Linguistica* 36.3-4, pp. 141–190. DOI: 10.1515/flin.2002.36.3-4.141 (cited on page 43).

Legendre, Géraldine (2001). 'An introduction to Optimality Theory in syntax'. In: *Optimality-Theoretic Syntax*. Ed. by Géraldine Legendre, Jane Grimshaw, and Sten Vikner. MIT Press. Chap. 1, pp. 1–27 (cited on pages 53, 55, 72).

– (2019). 'Optimality-Theoretic Syntax'. In: *Current Approaches to Syntax. A Comparative Handbook*. Ed. by András Kertész, Edith Moravcsik, and Csilla Rákosi. De Gruyter Mouton. Chap. 10, pp. 263–290. DOI: 10.1515/9783110540253-010 (cited on page 53).

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky (1990). 'Can Connectionism Contribute to Syntax? Harmonic Grammar, with an Application'. In: *Proceedings of the 26th Meeting of the Chicago Linguistic Society*. Citeseer (cited on page 58).

– (1991). 'Unifying syntactic and semantic approaches to unaccusativity: A connectionist approach'. In: *Annual Meeting of the Berkeley Linguistics Society*. Vol. 17. 1, pp. 156–167 (cited on page 58).

Legendre, Géraldine, Paul Smolensky, and Colin Wilson (1998). 'When is Less More? Faithfulness and Minimal Links in wh -Chains'. In: *Is the Best Good Enough? Optimality and Competition in Syntax*. Ed. by Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky. Cambridge, MA: MIT Press and MIT Working Papers in Linguistics, pp. 249–289 (cited on page 61).

Legendre, Géraldine, Antonella Sorace, and Paul Smolensky (2006). 'The Optimality Theory-Harmonic Grammar connection'. In: *The harmonic mind: From neural computation to Optimality Theoretic grammar, Volume 2: Linguistic and philosophical implications*. Ed. by Paul Smolensky and Géraldine Legendre. MIT Press, pp. 339–402 (cited on pages 58, 59).

Legendre, Géraldine, Colin Wilson, Paul Smolensky, Kristin Homer, and William Raymond (1995). 'Optimality and Wh-Extraction'. In: *Papers in Optimality Theory*. Ed. by Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk. Vol. 18. Occasional Papers in Linguistics. University of Massachusetts, pp. 607–636 (cited on page 61).

Lemmens, Maarten (2006). 'More on objectless transitives and ergativization patterns in English'. In: *Constructions* 1. Ed. by Doris Schönefeld. DOI: 10.24338/cons-447 (cited on pages 17, 37).

Lenci, Alessandro (2008). 'Distributional semantics in linguistic and cognitive research'. In: *Italian Journal of Linguistics* 20.1, pp. 1–31 (cited on page 92).

– (2018). 'Distributional Models of Word Meaning'. In: *Annual Review of Linguistics* 4, pp. 151–171 (cited on page 92).

Levin, Beth (1993). *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press. 348 pp. (cited on pages 17, 19, 23, 35, 46, 106, 108).

Levin, Beth and Malka Rappaport Hovav (2008). 'Lexicalized manner and result are in complementary distribution'. In: *Handout of talk given at the 24th meeting of the Israeli Association for Theoretical Linguistics*, pp. 26–27 (cited on pages 34–36).

Linzen, Tal and Yohei Oseki (2018). 'The Reliability of Acceptability Judgments across Languages'. In: *Glossa: A journal of general linguistics* 3.1, p. 100. DOI: 10.5334/gjgl.528 (cited on page 62).

Liu, Dilin (2008). 'Intransitive or Object Deleting?: Classifying English Verbs Used without an Object'. In: *Journal of English Linguistics* 36.4, pp. 289–313. DOI: 10.1177/0075424208317128 (cited on pages 10, 12, 14, 23, 32).

Liu, Lei (2014). 'Reconsidering the End-Point Approach: (A)Telicity and (Un)Boundedness Distinction'. In: *Theory and Practice in Language Studies* 4.1, pp. 137–142. DOI: 10.4304/tpls.4.1.137-142 (cited on page 101).

Lorenzetti, Maria Ivana (2008). 'The Null Instantiation of Objects as a Polysemy-Trigger. A Study on the English Verb See'. In: *Lexis Journal in English Lexicology* 1. DOI: 10.4000/lexis.769 (cited on pages 10, 19, 22, 23, 42, 49, 50, 170).

M

MacWhinney, Brian (2000). *The CHILDES Project: Tools for analyzing talk*. Vol. 1. Psychology Press (cited on page 69).

Magri, Giorgio (2018). 'Implicational Universals in Stochastic Constraint-Based Phonology'. In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 3265–3274 (cited on page 63).

Makowski, Dominique, Mattan S. Ben-Shachar, Indrajeet Patil, and Daniel Lüdecke (2021). *Automated Results Reporting as a Practical Tool to Improve Reproducibility and Methodological Best Practices Adoption*. URL: https://github.com/easystats/report (cited on page 121).

Malchukov, Andrej L. (2006). 'Transitivity parameters and transitivity alternations'. In: *Case, valency and transitivity*. Ed. by Leonid Kulikov, Andrej L. Malchukov, and Peter de Swart. Vol. 77. Amsterdam: John Benjamins Publishing Company, pp. 329–357 (cited on page 33).

Manning, Christopher D. (2003). *Probabilistic Syntax*. Ed. by Rens Bod, Jennifer Hay, and Stefanie Jannedy. Mit Press (cited on pages 62, 63, 167, 169).

Maouene, Josita, Aarre Laakso, and Linda B. Smith (2011). 'Object Associations of Early-Learned Light and Heavy English Verbs'. In: *First Language* 31.1, pp. 109–132. DOI: 10.1177/0142723710380528 (cited on page 32).

Marantz, Alec (1981). 'On the nature of grammatical relations'. PhD thesis. Massachusetts Institute of Technology (cited on page 26).

Martí, Luisa (2010). 'Implicit Indefinite Objects: The Barest of the Bare'. In: *Occasional Papers Advancing Linguistics* 15 (cited on page 23).

– (2015). 'Grammar versus Pragmatics: Carving Nature at the Joints: Grammar versus Pragmatics'. In: *Mind & Language* 30.4, pp. 437–473. DOI: 10.1111/mila.12086 (cited on pages 18, 23–25).

Massam, Diane (1992). 'Null Objects and Non-Thematic Subjects'. In: *Journal of Linguistics* 28.1, pp. 115–137 (cited on page 14).

Massam, Diane and Yves Roberge (1989). 'Recipe Context Null Objects in English'. In: *Linguistic Inquiry* 20.1, pp. 134–139 (cited on page 14).

Mastrofini, Roberta (2013). 'English manner of speaking verbs and their Italian translations'. In: *Athens Journal of Philology* 1.2, pp. 83–98. DOI: 10.30958/ajp.1-2-1 (cited on page 136).

McShane, Marjorie J. (2005). *A Theory of Ellipsis*. Oxford University Press, USA (cited on page 21).

Medina, Tamara Nicol (2007). 'Learning Which Verbs Allow Object Omission: Verb Semantic Selectivity and the Implicit Object Construction'. PhD thesis. Johns Hopkins University (cited on pages vii, 1, 2, 4, 6, 12, 31–33, 35, 39–45, 50–52, 57, 63, 64, 67, 68, 70–73, 75–78, 81–83, 86, 89, 98, 102, 103, 109, 112, 116, 121, 122, 128–130, 132, 133, 135, 137, 139, 141, 149, 150, 155, 158, 160, 164, 167, 168, 170, 173).

Megitt, Maria (2019). *"When Top Coals Are Partially Covered with Ash, Pour Evenly over Grill." A Study of Clause-Initial Adverbials and Ellipsis in Recipes*. Term Paper (cited on page 14).

Melchin, Paul (2019). 'The Semantic Basis for Selectional Restrictions'. PhD thesis. University of Ottawa. 188 pp. (cited on pages 12, 18, 19, 34–36, 39, 41).

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean (2013). 'Efficient estimation of word representations in vector space'. arXiv preprint arXiv:1301.3781 (cited on page 96).

Miller, George A. (1995). 'WordNet: A Lexical Database for English'. In: *Communications of the ACM* 38.11, pp. 39–41. DOI: 10.1145/219717.219748 (cited on pages 36, 68, 91, 106).

Mithun, Marianne (2009). 'Polysynthesis in the Arctic'. In: *Variations on polysynthesis: The Eskaleut languages*. Ed. by Marc-Antoine Mahieu and Nicole Tersis, pp. 3–18 (cited on page 24).

Mittwoch, Anita (1982). 'On the Difference between Eating and Eating Something: Activities versus Accomplishments'. In: *Linguistic Inquiry* 13.1, pp. 113–122. JSTOR: 4178263 (cited on pages 18, 21, 38, 40, 69).

– (2005). 'Unspecified Arguments in Episodic and Habitual Sentences'. In: *The Syntax of Aspect*. Oxford Studies in Theoretical Linguistics. Oxford University Press, pp. 237–254 (cited on pages 16, 18, 19, 25, 32, 46–48).

Myers, James (2009). 'Syntactic Judgment Experiments'. In: *Language and Linguistics Compass* 3.1, pp. 406–423. DOI: 10.1111/j.1749-818X.2008.00113.x (cited on page 113).

N

Nadejde, Maria, Alexandra Birch, and Philipp Koehn (2016). 'Modeling Selectional Preferences of Verbs and Nouns in String-to-Tree Machine Translation'. In: *Proceedings of the First Conference on Machine Translation: Volume 1, Research Papers*. Berlin, Germany: Association for Computational Linguistics, pp. 32–42. DOI: 10.18653/v1/W16-2204 (cited on page 92).

Næss, Åshild (2007). 'Prototypical Transitivity'. In: *Typological Studies in Language* 72 (cited on pages 8, 9, 19, 20, 22, 25–27, 33, 34, 36, 39, 40, 42–44).

– (2009). 'How Transitive Are EAT and DRINK Verbs?' In: *The Linguistics of Eating and Drinking*. John Benjamins Publishing (cited on page 9).

– (2011). 'The Grammar of Eating and Drinking Verbs'. In: *Language and Linguistics Compass* 5.6, pp. 413–423. DOI: 10.1111/j.1749-818X.2011.00279.x (cited on pages 20, 26, 27, 39, 46).

Nagy, Naomi and Bill Reynolds (1997). 'Optimality Theory and variable word-final deletion in Faetar'. In: *Language variation and change* 9.1, pp. 37–55 (cited on page 65).

Nedjalkov, Vladimir P. and Sergej J. Jaxontov (1988). 'Typology of resultative constructions'. In: *Typology of causative constructions*. Ed. by Vladimir P. Nedjalkov. Vol. 48. Typological Studies in Language. Amsterdam: John Benjamins Publishing, pp. 3–62 (cited on page 26).

Németh, Enikő (2014). 'Implicit Arguments at the Grammar-Pragmatics Interface: Some Methodological Considerations'. In: *Argumentum* 10, pp. 679–694 (cited on pages 48, 170).

Newman, John and Sally Rice (2006). 'Transitivity Schemas of English EAT and DRINK in the BNC'. In: *Corpora in Cognitive Linguistics: Corpus-Based Approaches to Syntax and Lexis*. De Gruyter Mouton. Chap. Corpora in Cognitive Linguistics, pp. 225–260 (cited on pages 19, 20, 39).

Nicolas, Aline (2019). 'Transitive Structures with Generic or Indefinite Object-Arguments in English as Functionally Antipassive Constructions'. Master's Thesis. Université de Liège (cited on page 26).

O

O'Grady, William, Yoshie Yamashita, and Sookeun Cho (2008). 'Object Drop in Japanese and Korean'. In: *Language Acquisition* 15.1, pp. 58–68. DOI: 10.1080/10489220701774278 (cited on page 31).

Ohlander, Urban (1943). 'Omission of the Object in English'. In: *Studia Neophilologica* 16.1, pp. 105–127. DOI: 10.1080/00393274308586940 (cited on page 30).

Olsen, Mari Broman (1997 [2014]). *A Semantic and Pragmatic Model of Lexical and Grammatical Aspect*. Routledge (cited on pages 35, 40–42, 44, 69–71, 83, 100, 102, 103).

Olsen, Mari Broman and Philip Resnik (1997). 'Implicit Object Constructions and the (In)Transitivity Continuum'. In: *In Proceedings of the 33rd Regional Meeting of the Chicago Linguistics Society*, pp. 327–336 (cited on pages 30, 32, 40, 69).

Olsen, Mari Broman, Amy Weinberg, Jeffrey P. Lilly, and John E. Drury (1998). 'Acquiring grammatical aspect via lexical aspect: The continuity hypothesis'. In: *University of Maryland Working Papers in Linguistics* 6, pp. 122–151 (cited on page 44).

Onelli, Corinna, Domenico Proietti, Corrado Seidenari, and Fabio Tamburini (2006). 'The DiaCORIS project: a diachronic corpus of written Italian'. In: *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)* (cited on page 171).

Onozuka, Hiromi (2007). 'Remarks on Causative Verbs and Object Deletion in English'. In: *Language Sciences* 29.4, pp. 538–553. DOI: 10.1016/j.langsci.2006.01.002 (cited on page 35).

P

Padó, Ulrike, Matthew W. Crocker, and Frank Keller (2009). 'A Probabilistic Model of Semantic Plausibility in Sentence Processing'. In: *Cognitive Science* 33.5, pp. 794–838. DOI: 10.1111/j.1551-6709.2009.01033.x (cited on page 92).

Paesani, Kate (2006). 'Extending the Nonsentential Analysis. The Case of Special Registers'. In: *The Syntax of Nonsententials: Multidisciplinary Perspectives*. Ed. by Ljiljana Progovac. Linguistik Aktuell = Linguistics Today v. 93. Amsterdam ; Philadelphia: John Benjamins Publishing, pp. 147–182 (cited on page 14).

Pater, Joe (2009). 'Weighted constraints in generative linguistics'. In: *Cognitive science* 33.6, pp. 999–1035 (cited on pages 58, 59).

Pater, Joe, Christopher Potts, and Rajesh Bhatt (2006). *Harmonic Grammar with Linear Programming* (cited on page 161).

Paul, Ileana and Diane Massam (2021). 'Licensing null arguments in recipes across languages'. In: *Journal of Linguistics* 57.4, pp. 815–839 (cited on page 14).

Peirce, Jonathan, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv (2019). 'PsychoPy2: Experiments in behavior made easy'. In: *Behavior research methods* 51.1, pp. 195–203 (cited on page 105).

Pérez-Leroux, Ana Teresa, Mihaela Pirvulescu, and Yves Roberge (2011). 'Topicalization and Object Omission in Child Language'. In: *First Language* 31.3, pp. 280–299. DOI: 10.1177/0142723710394384 (cited on page 31).

Pérez-Leroux, Ana Teresa, Mihaela Pirvulescu, and Yves Roberge (2018). *Direct Objects and Language Acquisition*. Cambridge Studies in Linguistics. Cambridge: Cambridge University Press (cited on page 31).

Pérez-Leroux, Ana Teresa, Mihaela Pirvulescu, Yves Roberge, and Anny Castilla (2013). 'On the Development of Null Implicit Objects in L1 English'. In: *Canadian Journal of Linguistics/Revue canadienne de linguistique* 58.3, pp. 443–464. DOI: 10.1017/S0008413100002656 (cited on page 31).

Perlmutter, David M. (1978). 'Impersonal passives and the unaccusative hypothesis'. In: *Annual meeting of the Berkeley Linguistics Society*. Vol. 4, pp. 157–190 (cited on page 8).

Permuth-Wey, Jennifer and Amy R. Borenstein (2009). 'Financial Remuneration for Clinical and Behavioral Research Participation: Ethical and Practical Considerations'. In: *Annals of Epidemiology* 19.4, pp. 280–285. DOI: https://doi.org/10.1016/j.annepidem.2009.01.004 (cited on page 113).

Pethõ, Gergely and Eva Kardos (2006). 'A Cross-Linguistic Investigation of the Licensing and Interpretation of Implicit Object Arguments'. In: *Pre-Proceedings of the SPRIK Conference 2006* (cited on pages 12, 16, 18, 21, 26, 46, 49).

Piñón, Christopher (2008). 'Aspectual composition with degrees'. In: *Adjectives and Adverbs: Syntax, Semantics, and Discourse*. Ed. by Louise McNally and Chris Kennedy. Oxford: Oxford University Press, pp. 183–219 (cited on pages 35, 39).

Prince, Alan and Paul Smolensky (1993 [2008]). *Optimality Theory: Constraint Interaction in Generative Grammar*. Wiley-Blackwell (cited on pages 52, 60).

– (1997). 'Optimality: From neural networks to universal grammar'. In: *Science* 275.5306, pp. 1604–1610 (cited on page 52).

Prytz, Johanna (2016). 'Optional Rhemes and Omitted Undergoers : An Event Structure Approach to Implicit Objects in Swedish'. PhD thesis. Stockholm University (cited on page 45).

Q

Quirk, Randolph, Sidney Greenbaum, Geoffrey Neil Leech, and Jan Svartvik (1985). *A grammar of contemporary English*. Longman London (cited on pages 14, 19).

R

Radden, Günter and Ken-ichi Seto (2003). 'Metonymic construals of shopping requests in HAVE-and BE-languages'. In: *Metonymy and Pragmatic Inferencing*. Ed. by Klaus-Uwe Panther and Linda L. Thornburg. Pragmatics and beyond. Amsterdam: John Benjamins Publishing Company, pp. 223–240 (cited on page 172).

Rappaport Hovav, Malka and Beth Levin (1998). 'Building verb meanings'. In: *The projection of arguments: Lexical and compositional factors*, pp. 97–134 (cited on pages 34–36).

– (2005). 'Change-of-State Verbs: Implications for Theories of Argument Projection'. In: *The Syntax of Aspect* (cited on pages 17, 34, 35).

– (2010). 'Reflections on Manner/Result Complementarity*'. In: *Lexical Semantics, Syntax, and Event Structure*. Ed. by Malka Rappaport Hovav, Edit Doron, and Ivy Sichel. Oxford University Press, pp. 21–38. DOI: 10.1093/acprof:oso/9780199544325.003.0002 (cited on pages 34, 35).

Rasetti, Lucienne (2003). 'Optional categories in early French syntax: a developmental study of root infinitives and null arguments'. PhD thesis. Université de Genève. DOI: 10.13097/archive-ouverte/unige:561 (cited on page 31).

Ratitamkul, Theeraporn, Adele E. Goldberg, and Cynthia Fisher (2004). 'The Role of Discourse Context in Determining the Argument Structure of Novel Verbs with Omitted Arguments'. In: *Proceedings of the 32nd Stanford Child Language Research Forum*. Stanford, CA (cited on page 31).

Recanati, François (2002). 'Unarticulated constituents'. In: *Linguistics and Philosophy* 25.3, pp. 299–345 (cited on pages 23, 170).

Resnik, Philip (1993). *Selection and Information: A Class-Based Approach to Lexical Relationships*. IRCS Technical Reports Series. University of Pennsylvania. 177 pp. (cited on pages vii, 2, 4, 15, 30, 32, 33, 67, 68, 74, 75, 78, 83, 89–91, 106, 108, 116, 129, 160, 166–168).

– (1996). 'Selectional Constraints: An Information-Theoretic Model and Its Computational Realization'. In: *Cognition* 61.1-2, pp. 127–159. DOI: 10.1016/S0010-0277(96)00722-6 (cited on pages vii, 15, 30, 32, 49, 50, 67–69, 74, 75, 78, 83, 89, 91, 92, 116, 129, 160, 168).

Reynolds, William Thomas (1994). 'Variation and phonological theory'. PhD thesis. University of Pennsylvania (cited on page 65).

Rice, Sally (1988). 'Unlikely Lexical Entries'. In: *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*. Vol. 14, pp. 202–212 (cited on pages 16, 19, 32, 34–37, 174).

Rimmer, Wayne (2006). 'Grammaticality judgment tests: Trial by error'. In: *Journal of Language and Linguistics* 5.2, pp. 246–261 (cited on page 60).

Rissanen, Matti (1993). 'The Helsinki Corpus of English Texts'. In: *Corpora Across the Centuries: Proceedings of the First International Colloquium on English Diachronic Corpora*. Ed. by Merja Kytö, Matti Rissanen, and Susan Wright, pp. 73–81 (cited on page 171).

Rissman, Lilia (2010). 'Instrumental with, Locatum with and the Argument/Adjunct Distinction'. In: *LSA Annual Meeting Extended Abstracts* 1, p. 23. DOI: 10.3765/exabs.v0i0.502 (cited on page 170).

– (2016). 'Cinderella Broke and Broke: Object Deletion and Manner-Result Complementarity'. In: *Proceedings of CLS 51*. Chicago Linguistic Society, pp. 425–439 (cited on pages 35, 48).

Rissman, Lilia and Kyle Rawlins (2017). 'Ingredients of Instrumental Meaning'. In: *Journal of Semantics* 34.3, pp. 507–537. DOI: 10.1093/jos/ffx003 (cited on page 170).

Rissman, Lilia, Kyle Rawlins, and Barbara Landau (2015). 'Using Instruments to Understand Argument Structure: Evidence for Gradient Representation'. In: *Cognition* 142, pp. 266–290. DOI: 10.1016/j.cognition.2015.05.015 (cited on page 170).

Rizzi, Luigi (1986). 'Null Objects in Italian and the Theory of Pro'. In: *Linguistic Inquiry* 17.3, pp. 501–557 (cited on page 16).

Roberge, Yves (2002). 'Transitivity requirement effects and the EPP'. In: *Western Conference on Linguistics* (cited on page 22).

Ruda, Marta (2014). 'Missing Objects in Special Registers: The Syntax of Null Objects in English'. In: *Canadian Journal of Linguistic-*

s/Revue canadienne de linguistique 59.3, pp. 339–372. doi: 10.1017/S0008413100000396 (cited on page 14).

Ruda, Marta (2017). *On the Syntax of Missing Objects: A Study with Special Reference to English, Polish, and Hungarian*. Vol. 244. Linguistik Aktuell/Linguistics Today. Amsterdam: John Benjamins Publishing Company (cited on pages 39, 103).

Ruppenhofer, Josef Karl (2004). 'The interaction of valence and information structure'. PhD thesis. University of California, Berkeley (cited on page 50).

– (2005). 'Regularities in null instantiation'. Manuscript. University of Colorado (cited on pages 12, 170).

Ruppenhofer, Josef Karl, Philip Gorinski, and Caroline Sporleder (2011). 'In search of missing arguments: A linguistic approach'. In: *Proceedings of the international conference Recent advances in natural language processing 2011*, pp. 331–338 (cited on page 170).

Ruppenhofer, Josef Karl and Laura A. Michaelis (2010). 'A Constructional Account of Genre-Based Argument Omissions'. In: *Constructions and Frames* 2.2, pp. 158–184. doi: 10.1075/cf.2.2.02rup (cited on pages 13–15).

– (2014). 'Frames and the Interpretation of Omitted Arguments in English'. In: *Pragmatics & Beyond New Series*. Ed. by Stacey Katz Bourns and Lindsy L. Myers. Vol. 244. Amsterdam: John Benjamins Publishing Company, pp. 57–86. doi: 10.1075/pbns.244.04rup (cited on pages 16, 170).

Rutherford, William E. (1998). *A workbook in the structure of English: Linguistic principles and language acquisition*. Blackwell Malden, MA (cited on page 10).

S

Schulte im Walde, Sabine, Christian Hying, Christian Scheible, and Helmut Schmid (2008). 'Combining EM Training and the MDL Principle for an Automatic Verb Classification Incorporating Selectional Preferences'. In: *Proceedings of ACL-08: HLT*. ACL-HLT 2008. Columbus, Ohio: Association for Computational Linguistics, pp. 496–504 (cited on page 92).

Schütze, Carson T. (1996 [2016]). 'The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology'. In: *Classics in Linguistics*, 1.01 MB. doi: 10.17169/LANGSCI.B89.100 (cited on pages 60, 169).

Scott, Kate (2006). 'When less is more: Implicit arguments and relevance theory'. In: *UCL Working Papers in Linguistics* 18, pp. 139–170 (cited on page 13).

Shannon, Claude E. (1948). 'A mathematical theory of communication'. In: *The Bell system technical journal* 27.3, pp. 379–423 (cited on page 95).

Shutova, Ekaterina, Niket Tandon, and Gerard de Melo (2015). 'Perceptually Grounded Selectional Preferences'. In: *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Beijing, China: Association for Computational Linguistics, pp. 950–960. doi: 10.3115/v1/P15-1092 (cited on page 92).

Sigurðsson, Halldór Ármann and Joan Maling (2008). 'Argument Drop and the Empty Left Edge Condition'. In: *Working papers in Scandinavian syntax*. Vol. 81, pp. 1–27 (cited on page 13).

Smith, Carlota S. (1991). *The parameter of aspect*. Vol. 43. Springer Science & Business Media (cited on pages 38, 102).

Smolensky, Paul (2006). 'Harmony in Linguistic Cognition'. In: *Cognitive Science* 30.5, pp. 779–801 (cited on page 59).

Smolensky, Paul, Géraldine Legendre, and Yoshiro Miyata (1993). 'Integrating connectionist and symbolic computation for the theory of language'. In: *Current Science*, pp. 381–391 (cited on page 58).

Smolensky, Paul and Alan Prince (1993). 'Optimality Theory: Constraint interaction in generative grammar'. In: *Optimality Theory in phonology* 3 (cited on pages 52, 53, 59, 60).

Smollett, Rebecca (2005). 'Quantized direct objects don't delimit after all'. In: *Perspectives on aspect*. Ed. by Henk J. Verkuyl, Henriette de Swart, and Angeliek van Hout. Vol. 32. Studies in Theoretical Psycholinguistics. Springer, pp. 41–59. DOI: 10.1007/1-4020-3232-3_3 (cited on pages 35, 39, 42).

Somers, Harold L. (1984). 'On the Validity of the Complement-Adjunct Distinction in Valency Grammar'. In: *Linguistics* 22.4. DOI: 10.1515/ling.1984.22.4.507 (cited on page 10).

Sopata, Aldona (2016). 'Null objects in adult and child Polish: Syntax, discourse and pragmatics'. In: *Lingua* 183, pp. 86–106 (cited on pages 31, 43, 173).

Sorace, Antonella and Frank Keller (2005). 'Gradience in Linguistic Data'. In: *Lingua* 115.11, pp. 1497–1524. DOI: 10.1016/j.lingua.2004.07.002 (cited on pages 62, 63).

Spärck Jones, Karen (1973). 'Index term weighting'. In: *Information storage and retrieval* 9.11, pp. 619–633 (cited on page 95).

Sprouse, Jon (2007). 'Continuous acceptability, categorical grammaticality, and experimental syntax'. In: *Biolinguistics* 1, pp. 123–134 (cited on pages 60, 62).

– (2015). 'Three Open Questions in Experimental Syntax'. In: *Linguistics Vanguard* 1.1. DOI: 10.1515/lingvan-2014-1012 (cited on page 62).

– (2018). 'Acceptability Judgments and Grammaticality, Prospects and Challenges'. In: *Syntactic Structures after 60 Years*. Ed. by Norbert Hornstein, Howard Lasnik, Pritty Patel-Grosz, and Charles Yang. Berlin, Boston: De Gruyter, pp. 195–224. DOI: 10.1515/9781501506925-199 (cited on page 62).

Stark, Elisabeth and Petra Meier (2017). 'Argument Drop in Swiss WhatsApp Messages'. In: *Zeitschrift für französische Sprache und Literatur* 127.3, p. 30 (cited on pages 11, 14).

Starosta, Stanley (1978). 'The one per sent solution'. In: *Valence, Semantic Case, and Grammatical Relations: Workshop Studies Prepared for the 12th International Congress of Linguists, Vienna, August 29th to September 3rd, 1977*. Ed. by Werner Abraham. Vol. 1. Amsterdam: John Benjamins Publishing, pp. 459–576 (cited on page 26).

Stoica, Irina (2017). 'Parametric Variation in the Role of Viewpoint Aspect in the Omission of Direct Objects'. In: *13th Conference on British and American Studies*. Cambridge Scholars Publishing (cited on page 44).

Stosic, Dejan (2019). 'Manner as a Cluster Concept: What Does Lexical Coding of Manner of Motion Tell Us about Manner?' In: *Human Cognitive Processing*. Ed. by Michel Aurnague and Dejan Stosic. Vol. 66.

Amsterdam: John Benjamins Publishing Company, pp. 142–177. DOI: 10.1075/hcp.66.04sto (cited on page 34).

Stosic, Dejan (2020). 'Defining the concept of manner: An attempt to order chaos'. In: *Testi e linguaggi* 14. DOI: 10.14273/unisa-3440 (cited on page 34).

Stuntebeck, Franziska (2018). 'Annotating argument drop in the Swiss Whatsapp corpus'. In: *Proceedings of Generative Grammar in Geneva 11*, pp. 1–13 (cited on page 14).

T

Talmy, Leonard (1991). 'Path to realization. A typology of event conflation'. In: *Annual Meeting of the Berkeley Linguistics Society*. Vol. 17. 1, pp. 480–519 (cited on page 136).

– (2000). *Toward a cognitive semantics*. Vol. 2. Cambridge: MIT Press (cited on page 136).

Taylor, John R. (1995). *Linguistic categorization*. OUP Oxford (cited on page 8).

Tenny, Carol L. (1994). *Aspectual roles and the syntax-semantics interface*. Vol. 52. Studies in Linguistics and Philosophy. Springer (cited on pages 25, 39, 42).

Tesar, Bruce and Paul Smolensky (1993). *The Learnability of Optimality Theory: An Algorithm and Some Basic Complexity Results*. DOI: 10.7282/T34Q7SB7 (Tech report) (cited on page 64).

Tesnière, Lucien (1959 [2015]). *Elements of structural syntax*. John Benjamins Publishing Company (cited on page 63).

Tonelli, Sara and Rodolfo Delmonte (2011). 'Desperately Seeking Implicit Arguments in Text'. In: *Proceedings of the ACL 2011 Workshop on Relational Models of Semantics*. Portland, Oregon, USA: Association for Computational Linguistics, pp. 54–62 (cited on pages 12, 13, 16).

Tsimpli, Ianthi Maria and Despina Papadopoulou (2006). 'Aspect and Argument Realization: A Study on Antecedentless Null Objects in Greek'. In: *Lingua* 116.10, pp. 1595–1615. DOI: 10.1016/j.lingua.2005.07.011 (cited on pages 42–44, 173).

Tsunoda, Tasaku (1999). 'Aspect and Transitivity of Iterative Constructions in Warrungu'. In: *Tense-Aspect, Transitivity and Causativity: Essays in Honor of Vladimir Nedjalkov*. Ed. by Werner Abraham and Leonid Kulikov, pp. 3–19 (cited on page 48).

Turney, Peter D. and Patrick Pantel (2010). 'From Frequency to Meaning: Vector Space Models of Semantics'. In: *Journal of Artificial Intelligence Research* 37, pp. 141–188. DOI: 10.1613/jair.2934 (cited on page 62).

V

Van de Cruys, Tim (2014). 'A Neural Network Approach to Selectional Preference Acquisition'. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, pp. 26–35. DOI: 10.3115/v1/D14-1004 (cited on page 92).

Van Heuven, Walter J. B., Pawel Mandera, Emmanuel Keuleers, and Marc Brysbaert (2014). 'SUBTLEX-UK: A new and improved word frequency database for British English'. In: *Quarterly journal of experimental psychology* 67.6, pp. 1176–1190. DOI: 10.1080/17470218.2013.850521 (cited on page 107).

Van Valin, Robert and Randy LaPolla (1997). *Syntax: Structure, meaning, and function*. Cambridge University Press (cited on pages 18, 21, 39, 63).

Vater, Heinz (1978). 'On the possibility of distinguishing between complements and adjuncts'. In: *Valence, semantic case and grammatical relations* 1, pp. 21–45 (cited on page 63).

Vendler, Zeno (1957). 'Verbs and Times'. In: *The Philosophical Review* 66.2, pp. 143–160. DOI: 10.2307/2182371 (cited on pages 38, 39, 69, 70).

Verkuyl, Henk J. (1972). *On the Compositional Nature of the Aspects*. Vol. Foundations of Language Supplementary Series. 15. Dordrecht, Netherlands: Reidel Publishing Company (cited on page 39).

– (1989). 'Aspectual classes and aspectual composition'. In: *Linguistics and philosophy*, pp. 39–94 (cited on page 39).

Villavicencio, Aline (2002). 'Learning to Distinguish PP Arguments from Adjuncts'. In: *Proceeding of the 6th Conference on Natural Language Learning - COLING-02*. Vol. 20. Not Known: Association for Computational Linguistics, pp. 1–7. DOI: 10.3115/1118853.1118886 (cited on page 63).

Virtanen, Pauli et al. (2020). 'SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python'. In: *Nature Methods* 17, pp. 261–272. DOI: 10.1038/s41592-019-0686-2 (cited on page 141).

W

Wagner, Laura (2001). 'Aspectual influences on early tense comprehension'. In: *Journal of Child Language* 28.3, p. 661 (cited on pages 44, 45, 102).

Wasow, Thomas (2007). 'Gradient data and gradient grammars'. In: *Proceedings from the Annual Meeting of the Chicago Linguistic Society*. Vol. 43. 1. Chicago Linguistic Society, pp. 255–271 (cited on page 62).

Weir, Andrew (2017). 'Object Drop and Article Drop in Reduced Written Register'. In: *Linguistic Variation* 17.2, pp. 157–185. DOI: 10.1075/lv.14016.wei (cited on page 14).

Weskott, Thomas and Gisbert Fanselow (2011). 'On the Informativity of Different Measures of Linguistic Acceptability'. In: *Language* 87.2, pp. 249–273. DOI: 10.1353/lan.2011.0041 (cited on pages 106, 112).

White, Michael (1993). 'The imperfective paradox and trajectory-of-motion events'. In: *31st Annual Meeting of the Association for Computational Linguistics*, pp. 283–285 (cited on page 41).

Wierzbicka, Anna (1982). 'Why Can You Have a Drink When You Can't *Have an Eat?' In: *Language* 58.4, pp. 753–799. DOI: 10.2307/413956 (cited on page 26).

Willim, Ewa (2006). *Event, individuation and countability: A study with special reference to English and Polish*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego (cited on page 39).

Wilson, Deirdre and Dan Sperber (2000). 'Truthfulness and relevance'. In: *UCL Working Papers in Linguistics* 12, pp. 215–254 (cited on page 23).

Y

Yankes, Andrew (2021 [2022]). 'Object drop in English: A statistical and Optimality Theoretical analysis'. MA thesis. Carnegie Mellon University (cited on pages 56, 57, 73, 74, 166).

Yasutake, Tomoko (1987). 'Objectless Transitives in English'. In: *The Bulletin of Aichi University of Education* 36, pp. 43–55 (cited on pages 10, 16, 20, 23, 24).

Yousefi, Moslem and Fatemeh Mardian (2019). 'Analyzing Meaning: An Introduction to Semantics and Pragmatics: Paul L. Kroeger Berlin: Language Science Press, 2018, Xiv+482 Pp.' In: *Australian Journal of Linguistics*, pp. 1–3. DOI: 10.1080/07268602.2019.1680090 (cited on page 43).

Z

Zapirain, Beñat, Eneko Agirre, Lluís Màrquez, and Mihai Surdeanu (2013). 'Selectional Preferences for Semantic Role Classification'. In: *Computational Linguistics* 39.3, pp. 631–663. DOI: 10.1162/COLI_a_00145 (cited on page 92).

Zipf, George K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley (cited on page 15).